Essays on Public Policies and Uncertainty

Inaugural-Dissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft

Fachbereich Wirtschaftswissenschaft
Freie Universität Berlin

vorgelegt von

Sebastian Schmitz, M.Sc.
aus Düsseldorf

Berlin, 2017

Erstgutachter
Prof. Dr. Viktor Steiner, Freie Universität Berlin

Zweitgutachter
Prof. Dr. Helmut Bester, Freie Universität Berlin

Tag der Disputation: 22.11.2017

Zusammenarbeit mit Koautoren und Vorveröffentlichungen

Kapitel 1: Is Our Income Too Risky to Have A(nother) Baby? Evidence from German Micro Data.

- Kapitel 1 basiert auf einem unveröffentlichten Artikel der zu gleichen Teilen mit Davud Rostam-Afschar verfasst wurde.
- Rostam-Afschar und Schmitz (mimeo) *Is Our Income Too Risky to Have A(nother) Baby? Evidence from German Micro Data.*

Kapitel 2: How Important is Precautionary Labor Supply?

- Kapitel 2 basiert auf einem Artikel der zu gleichen Teilen mit Davud Rostam-Afschar und Robin Jessen verfasst wurde. Teile von Kapitel 2 sind gegenüber dem zitierten Artikel verändert worden. Diese Anpassungen wurden von den Autoren gemeinsam vorgenommen.
- Jessen, Rostam-Afschar und Schmitz (2018) *How Important is Precautionary Labor Supply? Oxford Economic Papers*, online first `https://doi.org/10.1093/oep/gpx053`

Kapitel 3: The Effects of Germany's New Minimum Wage on Employment and Welfare Dependency

- Kapitel 3 basiert auf einem Artikel der eigenständig verfasst wurde. Teile von Kapitel 3 sind gegenüber dem zitierten Arbeitspapier verändert worden. Der Artikel wurde bei *German Economic Review* eingereicht
- Schmitz (2017) *The Effects of Germany's New Minimum Wage on Employment and Welfare Dependency.* School of Business & Economics Discussion Papers, Freie Universität Berlin, 2017/21.

Kapitel 4: Congestion Pricing: A Mechanism Design Approach

- Kapitel 4 basiert auf einem Artikel der zu gleichen Teilen mit Philipp Heller und Johannes Johnen verfasst wurde. Teile von Kapitel 4 sind gegenüber dem zitierten Arbeitspapier verändert worden. Diese Anpassungen wurden von den Autoren gemeinsam vorgenommen. Der Artikel wurde bei *Journal of Transport, Economics and Policy* eingereicht.
- Heller, Johnen und Schmitz (2015) *Congestion Pricing: A Mechanism Design Approach.* School of Business & Economics Discussion Papers, Freie Universität Berlin, 2015/20.

Acknowledgements

First and foremost, I am deeply grateful to my supervisor Viktor Steiner for excellent supervision during my time as a PhD student. I benefited a lot from discussions with him and his thoughtful suggestions. I also thank my second supervisor Helmut Bester for agreeing to co-supervise my dissertation, which made it possible to have such a methodically and thematically diverse collection of research projects.

My special thanks also to my coauthors Philipp Heller, Robin Jessen, Johannes Johnen and Davud Rostam-Afschar, with whom I wrote parts of this dissertation together. It was a pleasure working with you.

I want to thank Jochen Kluve for developing together with me a paper from my master thesis, which is now my first published article and not part of this thesis. I also want to thank Bas van der Klaauw and the people from the Vrije Universiteit Amsterdam for allowing me to spent some time at their institution.

In my life as an economist, I am deeply indebted to all professors, fellow students and colleague I met on my way. Here in Berlin, but also in the past years in Amsterdam, Toulouse, Ankara, Florence and Bolzano. I was so fortunate to find at all these places some truly outstanding people. My understanding about economics was formed and shaped not only in classes and my own research, but also in discussions and debates with all of you.

All chapters in this dissertation benefited from comments and suggestions from conference and seminar participants, fellow PhD students and researchers with whom I talked about my research projects. The papers improved considerably, due to their feedback.

Last but not least, I want to thank Anna, my family and my friends for their support during the time of this work.

Sebastian Schmitz
Berlin, July 2017

Contents

Preface: Public Policies & Uncertainty

In this dissertation I study a wide range of topics and policies, ranging from the impact of wage and income risk on fertility or labor supply, over the evaluation of a new minimum wage to design features for optimal traffic regulation. While these studies differ in key aspects and methods, they have in common that uncertainty plays a crucial role in understanding the impact of the respective policies. Thus, I decided to name this dissertation *Essays on Public Policies and Uncertainty*. I use the term *uncertainty* in a wider sense than the classic Knightian distinction (Knight, 1921) between risk, i.e. something quantifiable, opposed to an unmeasurable uncertainty. The first two chapters in this dissertation deal with the former, the last chapter with the latter concept of uncertainty. In the third chapter, neither uncertainty nor risk plays a major role, but the new minimum wage is an important policy to evaluate.

The first chapter, *Is Our Income Too Risky to Have A(nother) Baby? Evidence from German Micro Data*, studies women's fertility transitions and the timing of fertility, and how it is affected by household income, the female earnings potential, and the associated uninsured, idiosyncratic measures of risk. Uncertainty takes the form of properly quantified wage and income risk measures. The underlying policy question is whether households postpone founding a family or having another child because of income risk. If households experience increases in the riskiness of their income sources, they might wish to postpone family formation to increase their savings or to establish themselves on the labor market. The analysis in the first chapter is an empirical investigation using data from the German Socio-Economic Panel for the years 1984 to 2014. Fertility transitions, i.e. having a(nother) child, are modeled as the time until the first or subsequent births and estimated using discrete time duration models, taking into account unobserved heterogeneity.

The second chapter, *How Important is Precautionary Labor Supply?*, also studies the impact of wage risk, but this time on the labor supply decisions of German prime-age men. I study the question whether people chose to work longer hours in order to self-insure against their wage risk and the possibility of a negative wage shock. Thus, this chapter quantifies the importance of *precautionary labor supply* defined as the difference between hours supplied in the presence of wage risk and hours under perfect foresight. From a policy perspective a thorough understanding of labor supply incentives is key, e.g., for the design of the tax and transfer system or the unemployment insurance. Precautionary labor supply could also explain differences in hours worked across occupations or why self-employed work more hours than employees for a given wage. Also the second chapter is an exploratory empirical analysis based on data from the German Socio-Economic Panel. The estimation is based on an application of a dynamic labor supply model.

The third chapter, *The Effects of Germany's New Minimum Wage on Employment and Welfare Dependency*, deals with an actual public policy, Germany's new statutory minimum wage of € 8.50 per hour. The questions are whether the minimum wage caused job losses, or whether the introduction of the new wage floor reduced welfare dependency by lifting the income of the working poor. This chapter is an ex-post evaluation, studying the effects of the new minimum wage on regular and marginal employment and on welfare dependency, the so-called *Aufstocker*. This chapter uses the difference-in-differences technique on county-level administrative data, exploiting regional variation in the *bite* of the minimum wage. The bite is the county-specific share of employees paid less than € 8.50 *before* the introduction of the new policy. The idea is to exploit the fact that a uniformly set minimum wage is felt very differently across the country. Hence, also the effects of such a policy should be larger in more heavily affected counties. Neither uncertainty nor risk play a major or direct role in this chapter; however, as in any evaluation study, the goal is to remove some uncertainty about the true repercussions of the new policy.

The fourth chapter, *Congestion Pricing: A Mechanism Design Approach*, is a theoretical analysis of congestion pricing, using mechanism design, in order to overcome the regulator's uncertainty. This chapter does not deal with an actual policy, but analyses the optimal design of road pricing schemes for traffic regulation. From an economist's point of view, congestion is an externality problem since an individual driver does not take the effect of her journey on other drivers' travel time into account. The textbook solution is a simple corrective Pigouvian tax, set at a level to ensure that each driver internalizes the marginal cost of the increased travel time of other drivers (Pigou, 1920; Knight, 1924). In order to find the correct level of such a tax or price, a regulator needs knowledge about the distribution of drivers' time values. This chapter shows that Pigouvian prices are not efficient in the presence of *aggregate uncertainty* that arises in case of a finite number of drivers. Thus, setting a single Pigouvian price is generally not optimal. Nevertheless, mechanism design can be used to implement the efficient allocation. The suggested solution involves direct communication between the drivers and the regulator, which might have seemed impractical in the past. However, modern technology, such as smart phones, GPS, and the advent of self-driving cars imply that these practical problems may soon be overcome and that such a form of price discrimination could actually be used in future traffic regulation.

In sum, all four studies highlight the importance of risk and uncertainty for policy making, either in the form of agents' responses to risk, i.e. quantifiable uncertainty, or in the form of uncertainty about the exact repercussions of a new policy, or the efficient allocation of drivers on congested roads. Uncertainty is a fundamental trait of policy making and all four studies in this dissertation strive to reveal some of it.

1 Is Our Income Too Risky to Have A(nother) Baby? Evidence from German Micro Data

1.1 Introduction

Persistently low fertility rates are a growing challenge for many developed countries since they will put ever larger pressure on public budgets and the sustainability of the welfare state. Germany has a particularly low fertility rate[1] and faces one of the most severe demographic transitions in the world. According to the Deutsche Bundesbank (2017), the potential growth rate of the German economy will fall to well below 1% in the 2020s due to demographic change, which not only affects the working-age population, but also dampens the accumulation of capital stock (Deutsche Bundesbank, 2017) and hampers productivity growth (Bloom et al., 2001). Fertility rates well below replacement level lead to an ever older society, in which a shrinking work force needs to sustain an unprecedented number of pensioners. The German social security system with its pay-as-you-go pension scheme will be under pressure in the future, if it does not cut back the levels of benefit or increases the social security contributions considerably (Werding, 2014).

In the 1950s and 60s many high-income countries experienced a so-called baby boom, which was followed by persistent decrease in birth rates. Since the 1980s, all developed countries have fertility rates below the replacement level of 2.1 children per women, even though the variation across countries is large (Feyrer et al., 2008). Previous economic literature tried to explain these patterns in various ways. Following Gary Becker's seminal work (Becker, 1960)[2], fertility choices of households are frequently analyzed in light of the disposable family income and the "prices" of children, namely the opportunity cost in the form of foregone earnings. Galor and Weil (1996) stress the role of a narrowing gender wage gap and increases in the demand for human capital due to technological change (Galor and Weil, 2000). Likewise, Greenwood et al. (2005) argue that the baby boom was a result of an atypical burst of technological progress and attributes the baby bust to increases in the female wage rate. Goldin and Katz (2000, 2002) argue that the advent of "career women" in the 1970s was only possible because of the introduction of oral contraception as well as liberalizations in divorce

[1] The total fertility rate for Germany peaked at 1.5 in 2015, which is a rather high value for Germany. Since 1982, the German fertility rate has been well below 1.5 (Statistisches Bundesamt, 2016).

[2] For an extensive discussion of the classic economic fertility studies see Hotz et al. (1997) as well as Arroyo and Zhang (1997).

and abortion laws (Van de Kaa, 1997). In cross-country comparisons there used to be a negative correlation between female labor force participation and fertility across developed countries; however, this relationship switched signs in recent decades (cf. among others Apps and Rees, 2004), indicating the importance of family-friendly labor market institutions.

We extend this literature by studying a previously unexplored determinant of fertility choice, the impact of income risk. Founding a family or having another baby are rather fundamental decisions because they have an irreversible and permanent character. Thus, if households experience increases in the riskiness of their income sources, they might wish to postpone family formation, for instance in order to accumulate savings to insure themselves against negative shocks. Additionally, the prospective parents might want to work more to establish themselves on the labor market, in order to resolve some of the uncertainty about the future income trajectory. Income risk is an important concept in economics to describe, model and analyze intertemporal decisions. It has a prominent role in the economic literature on the permanent income hypothesis with precautionary saving, but is also related to other economic areas such as labor supply (Parker et al., 2005; Rostam-Afschar et al., 2016). Empirically, income risk has been volatile and generally increasing since the 1970s in Germany and elsewhere (Bönke et al., 2015; Blundell et al., 2015). We study whether there is a *causal* link between income risk and changes in fertility patterns. Our findings contribute primarily to the academic discussion but are interesting for policy makers, too. Family policies, such as child allowances or parental leave benefits, could potentially be improved, if they address heterogeneity in wage and income risk.

Following Heckman and Walker (1990a,b,c), we study fertility as the time it takes to have a child or to have another child given one or more children. Thus, we model the duration to the next childbirth and estimate discrete time hazard models for the first three fertility transitions, which we augment by the female wage rate, household income and measures of risk inherent in these two sources of income, as well as a rich set of sociodemographic covariates of both partners. Fertility decisions are constrained by biological factors and tastes for the number of children vary considerably; hence, following Heckman and Singer (1984), we include unobserved population heterogeneity in our estimation. The analysis is based on data from the Socio-Economic Panel (SOEP) for the years 1984 to 2014, a representative annual household panel survey in Germany. In order to study changing patterns of fertility, we differentiate our results with respect to two birth cohorts, born between 1960 and 1974 and between 1975 to 1989, and whether women hold a college degree or not. We restrict our analysis to West Germany, since the fertility developments in the former GDR seem to have their own transition dynamics.[3] To the best of our knowledge, we are the first to provide empirical evidence on the relation

[3] In East Germany, the period fertility rate dropped tremendously in the years following unification. Several studies relate this to the adaptation of Western fertility and labor market patterns (Conrad et al., 1996; Witte and Wagner, 1995; Bonin and Euwals, 2001). For 2001–2008, fertility rates seem to have converged between East and West Germany (Goldstein and Kreyenfeld, 2011). Goldstein et al. (2009)

between risk and fertility based on tracking the same women over large parts (up to 32 years) of their fertile life.

Conceptually, we assume that fertility decisions are made by the women alone, hence we have a female chauvinist model of fertility transitions. We include both single and cohabiting women, but assume that cohabitation (and mating) are exogenous. We do not distinguish whether the partner is married or not. Our income and risk measures are based on the permanent income potential rather than the actually observed income and wages. We deliberately do not use actual earnings, since these are related to the presence of children and fertility intentions. Lundborg et al. (2017) show that the effect of having children on earnings is negative because women work less and move to less paid jobs. Adda et al. (2017) show that selection into different careers may be based on the desire of having children, so that some costs of fertility incur even before children are born. Instead of the actual hourly wage rate, we therefore construct the women's parity-specific *potential* hourly wage rate and its risk, which both are unrelated to the actual labor supply decisions.[4] The potential wage rate captures the women's opportunity costs and therefore is an important determinant of the substitution decision between time spent working and time spent on childrearing activities. Of course, the potential wage rate also has an income effect for the fertility decisions; however, in line with previous research, we argue that for the women's earnings the substitution effect should dominate and the income effect becomes negligible. For cohabiting women, we also include the permanent net income of the household, i.e. the households capital income and the labor earnings from the male partner. We assume that it is exogenous for the women and hence should only feature an income effect. We will also assess the effect of the riskiness of this household income on fertility transitions.

For the older cohort we do not find any significant effects of the risk and income measures on the transition to the first child. For the younger cohort however, female wage risk leads to significantly longer time to the first child. For the transition to the second child, women from the older cohort with higher income opportunities tend to have shorter spacing between the first and the second child. We do not find such a significant effect for the younger cohort; however, for these women, the riskiness of the male income significantly reduces the probability of having a second child. We do not find any striking effects for the transition to the third child, which only rarely occurs in Germany. These results are mainly driven by low educated women.

The following section briefly presents existing research on the relationship between income risk, various kinds of uncertainties, and fertility. Section 1.3 presents some *stylized facts* about fertility transitions in West Germany. Section 1.4 describes the methods and data used in this study. In particular, we discuss the specification of our duration

point out that such a pattern was present in other post-communist Eastern European countries and merely reflects a rapid shift to childbearing later in life.

[4] We follow the standard approach and assume that our method to construct these measures unbiasedly reflect the information used for fertility decisions. See Cunha and Heckman (2016) for a discussion of alternative approaches to measure idiosyncratic risk.

model, explain how we constructed the key economic variables, income and wage potential, and how we measure riskiness of these variables. We present and discuss the results in Section 1.5. Section 1.6 concludes.

1.2 Uncertainty, Income Risk and Fertility Decisions

In this study we focus on the impact of wage risk on fertility. Wages are risky if they are subject to uninsurable idiosyncratic shocks, that is if a particular women treats part of her wage as accidental. For instance, unusual sickness, a bad guess about when to buy or sell, and similar factors (Friedman, 1957) may give rise to risk specific to a particular women. A common measure of this risk is the variance of the residual part of log wages. How this kind of risk affects fertility is largely unexplored. Sommer (2016) studies the relationship between uninsurable income risk and fertility using a structural life-cycle optimization model of savings, time allocation and fertility decisions. Young households postpone childbearing when income uncertainty is high, preferring to work and to accumulate more precautionary savings as a self-insurance before founding a family. Sommer also provides evidence that women whose husbands are in high-risk occupations have lower age-specific fertility rates.

Some attempts have been made to measure idiosyncratic risk using survey questions on *perceived insecurity*. Subjective concerns are frequently found to be negatively related to fertility. Various studies exploit the German Socio-economic Panel, which contains a series of question about the perception of the personal employment situation and general economic worries. Bhaumik and Nugent (2011) find that women whose employment situation is insecure (women with a job, but fearing to loose it and women without a job, hoping finding one) are less likely to get children compared to those who are confident to keep their job, or have a high certainty of remaining unemployed. Kreyenfeld (2010) finds that economic worries are related to fertility postponement by educated women which is confirmed by Kind and Kleibrink (2013). Hofmann and Hohmeyer (2013) use a labor market reform as an instrument for the economic concerns and also confirm the negative relationship between the concerns and fertility.

Another strand of the economic literature on fertility studies more predictable and less idiosyncratic forms of economic uncertainty. In Germany and other countries with elaborate labor market institutions it is frequently argued that at least for women with high educational attainment labor market integration precedes family formation (Bernardi et al., 2008). Indeed, occupational uncertainties such as fixed-term contracts, marginal employment, positions below personal educational attainment appear to be related to later transitions to parenthood (Kreyenfeld, 2010; Schmitt, 2012; Sutela, 2012; Auer and Danzer, 2016)[5] and subsequent fertility (Fiori et al., 2013). These results are

[5]For the Netherlands however, de Lange et al. (2014) do not find any significant effect of temporary contracts or the aggregate unemployment rates on the transition to the first child. Gebel and Giesecke (2009) do not find a effect of temporary contracts on fertility in Germany.

in line with a cross-country studies by Adsera (2004, 2005) who find that flexible labor markets such as in the US and in the Nordic countries are positively related to fertility, compared to more regulated labor markets, which tend to produce "protected insiders".

Unemployment is also seen as an adverse economic event which might effect fertility. For Germany, Gebel and Giesecke (2009) as well as Kreyenfeld (2010) report that women with high educational attainment tend to postpone fertility if they experience unemployment. Hofmann et al. (2017) show that the negative effect of job displacement on fertility depends on the business cycle. The adverse effects on the transition to the first child are larger if the job loss occurs during an economic downturn. For the transition to the second child, Wood et al. (2016) report that the aggregate unemployment level and the extent of temporary employment are negatively related in various European countries. On an aggregate level Hondroyiannis (2010) documents a negative impact of the unemployment rate and output volatility on fertility using a cross-country panel. Adsera and Menendez (2011) find similar results for Latin America. For Austria, Del Bono et al. (2012) study the impact of job loss on fertility using plant closures as exogenous variation. Their findings suggest that displaced workers (men and women) delay fertility in the short and the medium run. The effect is the strongest for female white collar employees. In a follow-up paper, Del Bono et al. (2015) argue that the negative fertility effect of unemployment and the associated temporary income loss is minor, compared to the negative effect of job displacement—which can be seen as an idiosyncratic wage shock.

1.3 Timing and Spacing of Births in Germany

In this section we present some stylized facts about fertility developments in Germany for the 1946–1986 annual birth cohorts in order to give an impression of the German fertility transition. Our analysis is based on the German Socio-Economic Panel (SOEP), a representative annual household survey (cf. Wagner et al., 2007). Even though the data from the SOEP is available only for the years 1984-2014, there is retrospective information about the birth histories. Therefore, in this section we cover information from 50 years, i.e. from 1946–1996 assuming that fertile life begins at age 18. Thus, we can trace the fertility developments also for those women which we do not observe during their fertile years (born before 1966). Recall that we restrict our analysis to West Germany. The developments we sketch here are similar to those in other Western- and Southern-European countries, even though there might be some slight differences with respect to the timing of events and the magnitude of the transitions.

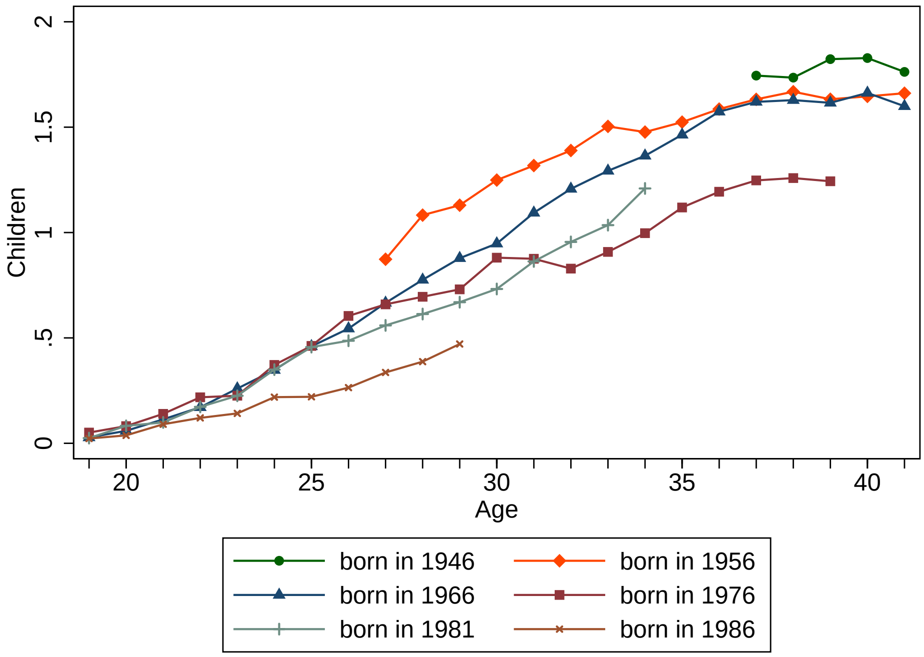
The Intensive and the Extensive Margin of Fertility Figure 1.1a depicts the age-specific fertility rate for six cohorts, namely those born in 1946, 1956, 1966, 1976, 1981 and 1986.[6] As we would expect, the profiles of younger cohorts tend to be below those of older cohorts. However, completed fertility, i.e. the age-specific rate at the end of the fertile cycle, is very close for the first three cohorts (1946-1966). Hence, women born in the 1960s first delayed fertility, but caught up in their thirties to reach fertility levels similar to those in older cohorts. The younger three cohorts however do not seem to feature such a strong catch-up effect and remain at lower levels of completed fertility.

Figure 1.1b follows the same structure as the previous one; however, instead of the fertility rate of all fertile women it displays the age-specific fertility rates only for women who already have a child. Compared to the previous graph, the profiles are much more aligned, especially up to the age of 30. For the years thereafter and for completed fertility, there are notable differences. The patterns suggest that if there was a first birth, the transition to the second child did not change tremendously over cohorts. Higher order fertility transitions occurred at lower rates.

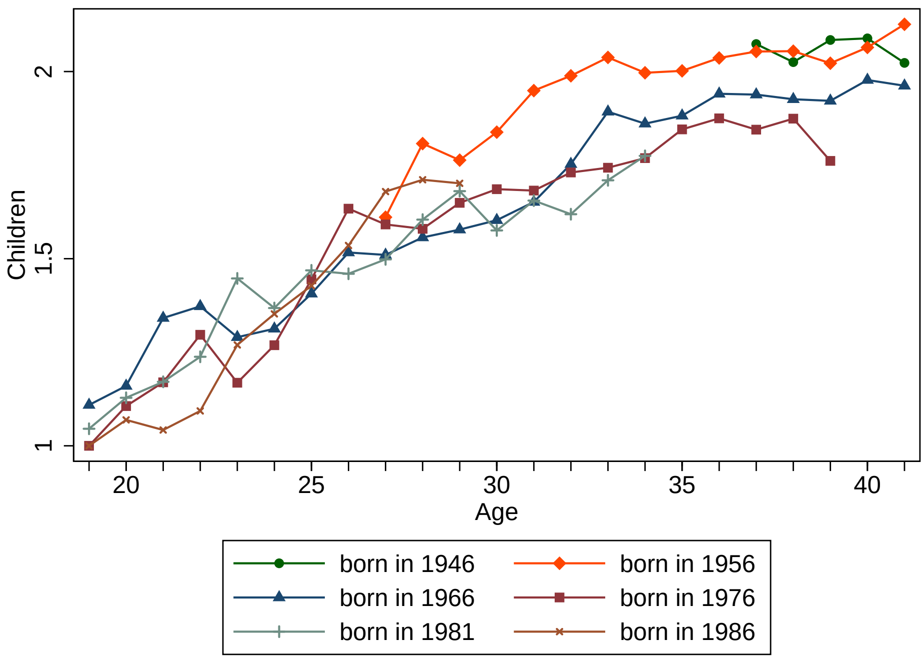
Figure 1.2 again has the same structure as the previous figures; however, in order to scrutinize changes at the *extensive margin* of fertility, we look at the age-specific share of childless women for the six cohorts. The figure provides evidence that starting from the cohort born in 1966 age-specific childlessness increases from cohort to cohort.

The profiles suggest both, a development towards founding a family later in the life-cycle, but also towards permanent childlessness.

[6]Each cohort of women born in the eponymous year or in the year before or after, in order to increase the number of observations. SOEP sample weights are used for all graphs.



(a) Age-Specific Fertility Rates for Different Cohorts



(b) Age-Specific Fertility Rates for Women with at least One Child

Figure 1.1: Development of Age-Specific Fertility Rates across Cohorts

Source: Authors' calculations.

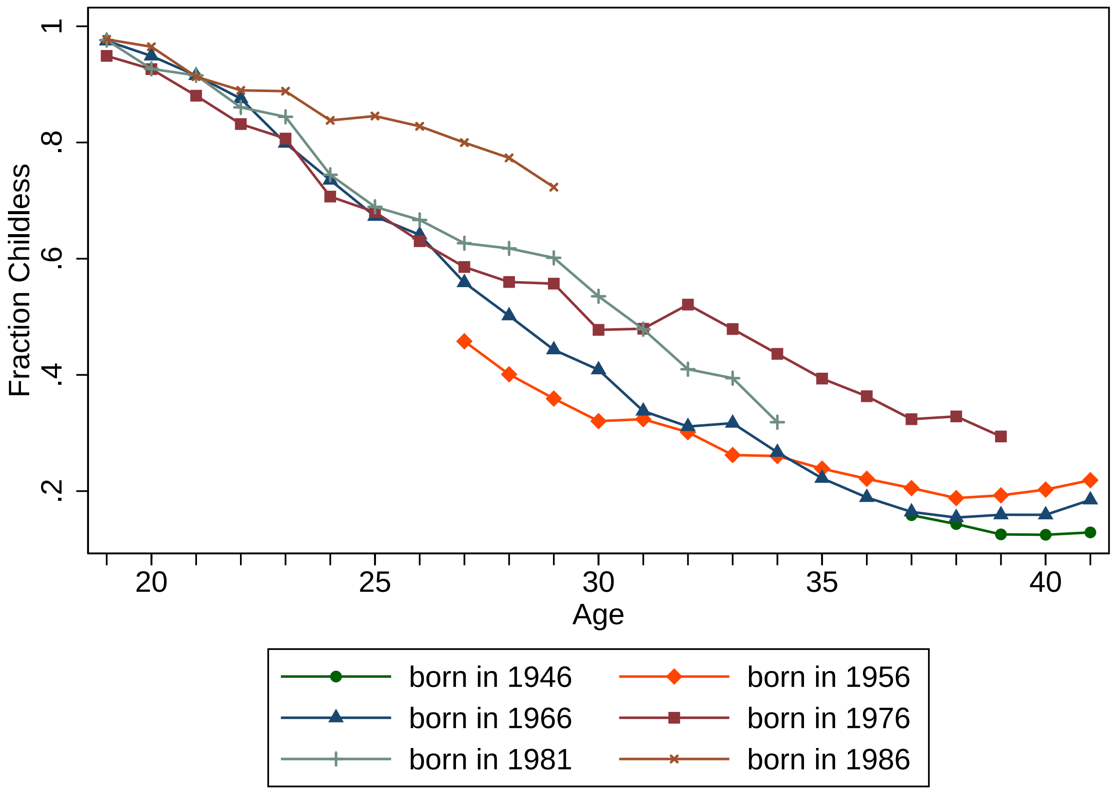


Figure 1.2: Share of Childless Women across Cohorts

Source: Authors' calculations.

Changes in the Age at Birth Another way to look at these adjustments is to plot the mothers' age at birth for different parities, as it is done in Figure 1.3. Since we can only generate meaningful graphs for those with completed fertile cycles, we drop women born after 1975, i.e. older than 40 years in the last year of our sample. For the cohorts from 1946 to 1975, we observe 5 women with 12 children. The particular ages of these women and other mothers with at least three children contribute to the average ages at the first three parities that we present in the figure. The 47,771 women who have two children additionally contribute to the average ages at first and second births but not the third births, 31,614 mothers of one child contribute additionally to the average age at first birth. In order to ease readability we predicted a quadratic trend. All three curves are sloping upwards, indicating that average ages at all parities increased from cohort to cohort. The largest change can be detected for the average age at the first child, which increased from slightly below 24 to 27. The change for the average age at the second child seems to be a parallel shift, going from a little bit above 26 to almost 30. Thus, while this suggests that younger cohorts postponed first births, the average spacing between the first and the second child did not change much. The curve for the average age at the third child, however, is rather flat: Mothers who will have at least three children in total start early, do not to seem to postpone the second child, and therefore must reduce spacing between second and third child.

For the following analysis we are limited to the years 1984 to 2014, hence we only observe the fertile years of the cohorts born in the 60s, 70s, and 80s. The exploratory analysis suggest that there is a structural difference in the fertility patters between those born before and after 1970. Compared to the post war generation, both groups have lower cohort-specific fertility rates. However, only for those born after 1970 we observe a sharp increase in the share of childless women. Hence, we find it informative for the rest of our analysis to split our sample in a cohort 1, born between 1960 and 1969 and a cohort 2, i.e. born between 1970 and 1989.[7]

[7]This definition of cohorts makes sure that we have in both groups full support over the entire fertile cycle, which we define to be between 18 and 49 years (cf. Section 1.4.1).

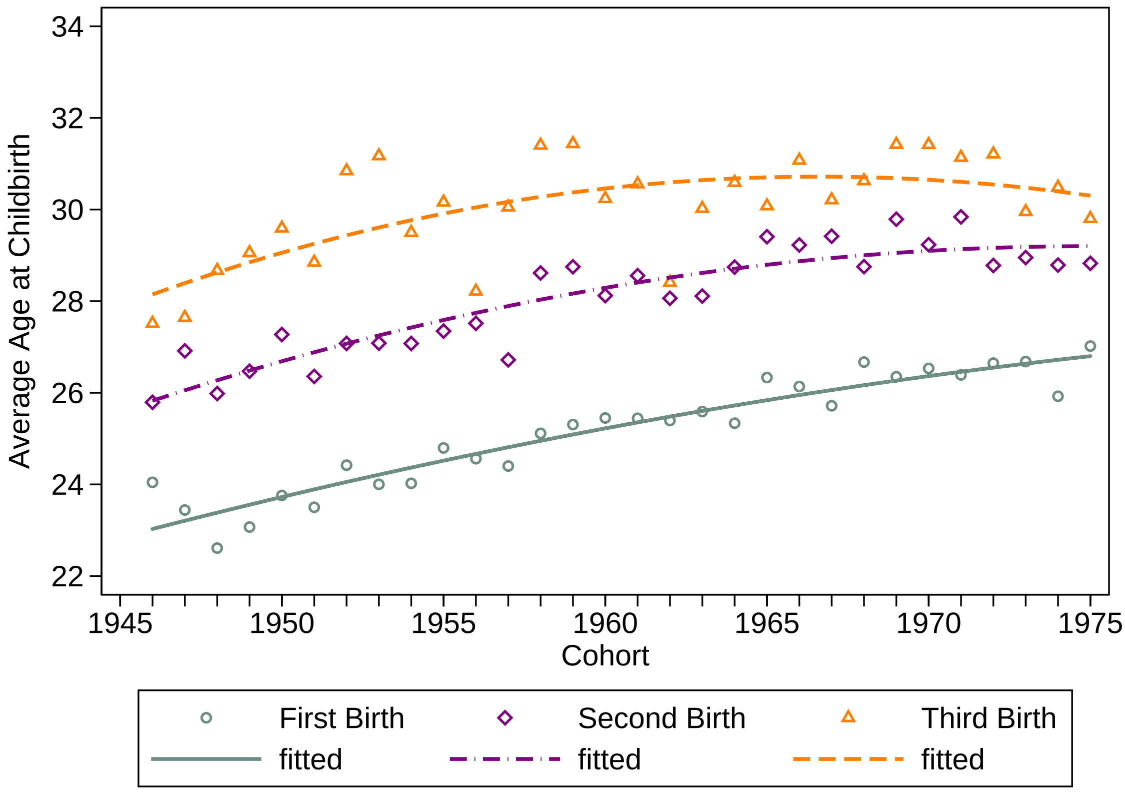


Figure 1.3: Age at 1st, 2nd and 3rd Child across Cohorts

Source: Authors' calculations.

1.4 Method and Data

1.4.1 A Duration Model of Fertility Transitions

We are interested in the effects of risk, female wages and household income on fertility decisions. We model fertility as the waiting time to the next child using a multiple spell discrete time hazard model. We jointly estimate up to three transitions and control for unobserved population heterogeneity. Given that we are interested in fertility *decisions*, we define a transition to occur in the presumable year of the pregnancy and not in the year of the actual childbirth.[8] Hence, modeling is based on the time until the first pregnancy and the duration until the subsequent pregnancy.

The time to the first pregnancy is denoted as the first spell, and the subsequent transitions accordingly second and third spell. Our data, the German Socio Economic Panel has annual frequency. Thus, the duration of women i's j-th fertility spell can be described by a non-negative random variable T, which takes integer values only. This random variable will take the value t if a transition occurs in the t-th year after the beginning of the spell-specific time at risk.[9] The spell-specific hazard function λ_{ij} gives the conditional probability of leaving a parity state by becoming pregnant and can be formally defined for individual i as

$$\lambda_{ij}\left[t|x_{ij}(t), \theta_j \varepsilon_i^m\right] = P[T_{ij} = t|T_{ij} \geq t, x_{ij}(t), \theta_j \varepsilon_i^m]. \tag{1.1}$$

The hazard rate is thus a function of spell-specific unobserved heterogeneity $\theta_j \varepsilon_i^m$ and a spell-specific and possibly time-varying vector $x_{ij}(t)$ which contains the logs of the income and risk variables which we will introduce in the following section, further control variables (cohabitation, years of education) and a specification for the duration dependence, the so-called baseline hazard. Duration dependence means, that the value of the hazard depends on the amount of time that has already elapsed. We model duration dependence using cubic splines of the parity-specific time at risk. Parity-specific cubic splines are a comparatively flexible non-linear way to model duration dependence without relying on too many parameters, in the case of using dummy variables for the time since the last births.[10] For transition 2 and 3 the vector $x_{ij}(t)$ also includes the time spent in the previous spell(s) in order to capture the inter-linkages across spells. Conceptually, we assume that fertility decisions are made by the women alone, hence

[8] We assume that all children born in the fourth quarter of a year also have been conceived in the same year, all other children in the previous year. Of course there can be a considerable lag between the actual fertility decision, e.g. stopping contraceptives, the actual pregnancy. However, given the difficulty in modeling this lag, given the information in the SOEP we decided to abstract from these issues.

[9] Technically, t falls into the interval $[I_{t-1}, I_t]$.

[10] The main results are unaffected by the choice of the bandwidth, as soon as the specification is sufficiently flexible (quadratic, cubic, dummies for certain interval length, etc.). In practice, the cubic splines offered the most reliable specification of the baseline concerning convergence of the models with unobserved heterogeneity.

we have a female chauvinist model of fertility transitions. We include both single and cohabiting women, but assume that cohabitation (and mateing) are exogenous. We do not distinguish whether the partner is married or not. Variables in $x_{ij}(t)$ which refer to the partner (years in education of the partner, permanent household income and its riskiness) are introduced as interactions with the period-specific cohabitation indicator.

As argued among others in Heckman and Walker (1990a,b,c), not controlling for unobserved heterogeneity might severely bias the estimated coefficients. Generally, unobserved heterogeneity refers to any time-invariant factors, not accounted for by the included covariates. In the context of fertility transition, this heterogeneity might either be biological (fecundity), or more importantly for a developed country, captures fertility preferences, which are assumed to be uncorrelated to the included covariates. Following Heckman and Singer (1984), we model the distribution of the unobserved factors in a nonparametric way: The distribution is expressed by a discrete probability distribution with a small number of mass points ε_i^m and their associated probabilities with

$$E(\varepsilon_i) = \sum_{m=1}^{M} P(\varepsilon_i^m)\varepsilon_i^m = 0; \sum_{m=1}^{M} P(\varepsilon_i^m) = 1; E(\varepsilon_i^m x_{ij}(t)) = 0, \forall m \in \{1,2,...M\}.$$

Thus, we assume that the unobserved heterogeneity is time-invariant and uncorrelated with the included covariates. In order to ease identification, we further assume that the heterogeneity is constant across parities. However, we additionally include so-called factor loads θ_j to capture differences in the importance of unobserved heterogeneity across spells. The factor load for the first spell θ_1 is normalized to unity. All factor loads, mass points and their probabilities are jointly estimated.

Given our definition of the hazard function (1.1), the probability of staying in a certain parity state in period t, conditional on not having experiences a(nother) pregnancy up to this point is than given by

$$P[T_j > t | T_j \geq t, x_{ij}(t), \theta_j \varepsilon_i^m] = 1 - \lambda_{ij}\left[t | x_{ij}(t), \theta_j \varepsilon_i^m\right]. \tag{1.2}$$

Likewise, the survivor function, i.e. the unconditional probability of remaining in parity state j until time t can be calculated as the product of Equation (1.2) from the first year at risk up to $t-1$

$$P\left[T_j > t | x_{ij}(t), \theta_j \varepsilon_i^m\right] = S_j\left[t | x_{ij}(t), \theta_j \varepsilon_i^m\right] = \prod_{\tau=1}^{t-1}\left[1 - \lambda_{ij}\left[\tau | x_{ij}(t), \theta_j \varepsilon_i^m\right].\right] \tag{1.3}$$

Putting things together, we can estimate the unconditional probability of transiting in period t as the product of the hazard function (1.1) and the survivor function (1.3)

$$P\left[T_j = t | x_{ij}(t), \theta_j \varepsilon_i^m\right] = \lambda_{ij}\left[t | x_{ij}(t), \theta_j \varepsilon_i^m\right] \times \prod_{\tau=1}^{t-1}\left[1 - \lambda_{ij}\left[\tau | x_{ij}(\tau), \theta_j \varepsilon_i^m\right]\right]. \tag{1.4}$$

In order to obtain the sample likelihood, we need to assume that conditional on all explanatory variables and the individual effects, all observations are independent. We consider up to three fertility transitions, thus $J = 3$. Then the sample likelihood is

$$L = \prod_{i=1}^{n} \sum_{m=1}^{M} P(\varepsilon_i^m) \prod_{j=1}^{3} \left[\lambda_{ij}\left[t_{ij}|x_{ij}(t_{ij}), \theta_j \varepsilon_i^m\right]\right]^{\delta_{ij}} \prod_{\tau=1}^{t_i-1} \left[1 - \lambda_{ij}\left[\tau|x_{ij}(\tau), \theta_j \varepsilon_i^m\right]\right], \tag{1.5}$$

with

$$\delta_{ij} = \begin{cases} 1, & \text{if there is a transition of individual } i \text{ in the } j\text{-th spell} \\ 0, & \text{otherwise.} \end{cases}$$

Finally, we need to specify a functional form for the hazard rate, the so-called link-function. As proposed for instance in Jenkins (1995) we will use the complementary log-log specification. For the application of fertility transitions, the cloglog specification has some appeal over a simple logit specifications, since fertility transitions occur only rarely. For the cloglog link function, the hazard rate λ is than given by

$$\lambda_{ij}\left[t|x_{ij}(t), \theta_j \varepsilon_i^m\right] = 1 - \exp\left(-\exp\left(\beta_j' x_{ij}(t) + \theta_j \varepsilon_i^m\right)\right). \tag{1.6}$$

We will make use of the survivor function in order to illustrate the implication of the estimation results. With the cloglog specification, the survivor function (1.3) becomes

$$S_{ij}(t) = \exp\left\{\sum_{\tau=0}^{t} [\ln(1 - \lambda_{ij}(\tau))]\right\}. \tag{1.7}$$

We obtain the estimation equation by plugging equation (1.6) and (1.7) into the likelihood function (1.5). The resulting (log) likelihood including all mass points, probabilities and factor loads is maximized using maximum likelihood estimation and the GLLAMM package (Rabe-Hesketh et al., 2005).

1.4.2 Data

The model is estimated on the German SOEP data for the years 1984 to 2014. We define women to be at risk for the first transition starting at the age of 18.[11] We track women up to the year they turn 50. Women in our sample are at risk of transiting from parity zero to parity one from the year onwards, in which they turn 18. Women with children born before their 18th birthday are excluded entirely. Additionally, we exclude women in same-sex relationships or with adopted or foster children. We also exclude women from East Germany, because we observe them primarily in the adjustment period after

[11] Many demographic studies, using a similar method follow women from their first menstruation or from a very young age such as 12 or 14. Such a procedure does not seem practical for our purpose, given that we cannot construct meaningful wage rates and income variables.

unification, which has been present in almost all former Socialist countries in the 1990s. The fertility patterns in Eastern Germany are clearly driven by the adjustment to West German models and their inclusions could largely confound our estimation. We also exclude person years (but not entire cases) for women with a male partner younger than 18 or older than 58, women in any form of education (School, university, intern, apprentice, trainees, aspirants), who were already in early retirement or are working in agriculture. We use 31 years of SOEP data from 1984 to 2014. In order to have a representative sample of women mostly in their fertile years, we restrict attention to women born in 1960 or later (i.e. women who are at maximum 25 in the first year of the SOEP) and before 1990 (i.e. women who are at minimum 24 in 2014). Hence we are left with the birth cohorts born in the 1960s, 1970s and 1980s.

Compared to previous studies in the spirit of Heckman and Walker (1990a,b,c), our data has some limitations, since for many cases we only observe a section of the entire time at risk until pregnancy and the data is frequently left and/or right-censored. Since we estimate our model in discrete time, we have by definition interval-censored data. Right-censoring occurs if a woman leaves the study before a pregnancy occurs, and if the study ends before the transition has occurred. Left-censoring occurs if we women joins the study only after the age of 18 or some of the control variables in $x_{ij}(t)$ are only available after a certain age. Since we observe the entire birth history, we can compute the respective process time at risk t starting from the age of 18 for the first spell. For the second and third spell, we start at the year following the last pregnancy. Thus, both types of censoring boil down to a problem of missing observations; however, this limitation is in our opinion unproblematic, as long as we assume *random* sample attrition in the SOEP and have for every sub-group and every transition full support over the relevant time at risk horizon. Given our mild sample restrictions, the high quality of the SOEP, and that we control for time at risk and starting age for the higher parity orders, we are still able to properly estimate the fertility transitions, based on our sample. Our data allows to condition on the complete birth history process, hence we can abstract from initial condition problems and sample selection problems which are discussed in Heckman and Singer (1984).

1.4.3 Construction of Income, Wage, and Risk Measures

We want to model the impact of the different sources of household income and their riskiness on fertility decisions. In line with previous research on the impact of the financial situation on fertility, we split the income in two distinct sources, first, the net annual household income after taxes net of the earnings of the woman, second, the female potential wage rate. Using the women's potential wage rate and not actual labor earnings allows us to abstract from women's actual labor supply decisions, which are likely related to the presence of and preferences for children. Additionally, focusing on the potential wage rate allows us to obtain a measure of the women's permanent income which is usually not estimated in the literature, due to child-related interruptions in employment. The potential wage rate captures the women's substitution decision between time spent working and time spent on childrearing activities. Of course, the potential wage rate also has an income effect for the fertility decisions; however, in line with previous research, we argue that for the women's earnings the substitution effect should dominate. The annual net income of the household without the female labor income (henceforth the household income), is assumed to be exogenous and should only induce an income effect for fertility transitions. All income and wage rate variables are net of taxes and deflated to the year 2010.

Expected Household Income Following the literature on precautionary savings, such as Lusardi (1998); Fossen and Rostam-Afschar (2013); Jessen et al. (2017), we thus decompose both the income and the wage rate into a permanent (or potential) part and an associated measure for its riskiness. We obtain the permanent part of the household income in the spirit of Friedman (1957)[12] as the prediction of a parity-specific[13] regression of the log of the net income on a large set of covariates.[14] These variables include a quadratic polynomial in potential experience of the men, indicators for educational attainment and occupation (single-digit ISCO-88), as well as interactions of them with a quadratic in experience, and a quadratic in tenure in the current job.[15] Additionally, also a quadratic in calender years in order to capture economic growth, regional indicators and cohort dummies are included. The idea is to capture all observable variables, which households could use, when inferring their permanent income potential.

Potential Wage Rate For the female wage rate, the procedure is similar, but requires some modifications in order to deal with participation issues. As for household income, the wage regressions are parity-specific. The left-hand side variable is the log of the net hourly wage rate, the right-hand side contains the same variables as for the household income, but additionally include quadratics in the actual labor market history (years

[12] Friedman defined permanent income as annuity value of total wealth, i.e. of human (present value of lifetime earnings) and financial wealth. Our measure is what Friedman describes as the expected value of a probability distribution reflecting the effect of e.g. the training, ability, personality, occupation, location of the economic activity of the women, and so on.

[13] The regressions are parity-specific in order to allow different returns to factors, depending on the number of children in the household. Additionally, as shown by Lundborg et al. (2017), childbirths seems to have a negative causal effect on the hourly wage rates, hence parity-specific regressions seem more appropriate in order to uncover the true earnings potential.

[14] The resulting predictions are adjusted for the distortion from the exponentiation by multiplying the exponent of the prediction with the coefficient of a simple bivariate regression without constant term of the exponent of the predicted value on the actually observed income as described in footnote 17 in Fossen and Rostam-Afschar (2013).

[15] Of course, e.g., the current occupation reflects a choice and thus would be a bad control. Therefore, occupational choice could be estimated in a joint model together with fertility choice Adda et al. (2017). However, formulating a model for both decision margins increases the likelihood of mis-specifying the estimation equation. Thus we prefer to focus on the piece meal approach.

in part-time, years in unemployment, inactive years), the actual hours worked and a dummy for cohabitation. We apply the selection correction proposed by Heckman (1979), using the presence of young children[16] and a dummy for being pregnant as exclusion restrictions. In order to have a measure for the earnings potential, *unrelated* to the actual decisions regarding fertility, we replace all covariates related to children to hypothetical values. Thus, for the prediction of the potential wage rate, we assume that all women are cohabiting, working close to full-time (32 hours per week) are currently not pregnant, and have no children under three years. Additionally, all labor market history variables are set to age and parity specific averages. As for the household income, we apply the procedure to convert the predicted values to log-scale.

Risk Measures We argue that the residuals from the regression described above capture the stochastic income component. Hence, in order to obtain a risk measure for the household income, we regress the log of the squared residuals on the same set of covariates as before, hence, we fit a heteroskedasticity function. Also these regressions are performed for each parity separately. As in Fossen and Rostam-Afschar (2013), the prediction from these regressions provide us with a measure for the riskiness of the household income.

Also for wage risk, we obtain our risk measure using the heteroskedasticity function approach. For women who are actually working, we calculated the residuals from the regression directly; for the women who are not working, we simulate the residuals, based on the estimated parameters from the Heckman selection correction. We regress the log of the squared residuals on all original covariates using OLS. For the prediction of the risk measure, we set all child and labor supply related covariates to the same previously described hypothetical values. Hence, for both, the wage rate and its riskiness, we control in our regression for child-related factors, but do not let them enter our measures for the wage and risk.

1.4.4 Summary Statistics

Table 1.1 presents the spell and cohort-specific weighted summary statistics.[17] The table reports also the number of cases and the number of pregnancies we observe at the bottom of each panel. Years at risk refers to the time since being exposed to the risk of a first pregnancy. The spell and cohort-specific distribution of times at risk and pregnancies are displayed graphically in Figure 1.4 (described below).

All partner variables (partner's years of education: "Educ male", net income net of female income: "Net Inc male", and risk of net income net of female income: "Risk male"), are reported conditioning on cohabitation. These variables are zero if no partner

[16]A dummy for a child younger than one, and a dummy for a child between 1 and 3.

[17]The summary statistics for the joint sample and separated for educational attainment can be found in the Appendix (Table 1.A1 and 1.A2).

is present. Differently as in the actual estimation sample, all risk and income variables are reported in levels and are not centered. The age at 1st birth is only defined for those with a first birth. At first glance it might be surprising that the age at first birth is much higher for the first cohort than for the second; however, this comes from a composition effect: In cohort 2 are more women in the early years at risk, which one can also see in the summary statistic *years at risk*. Table 1.1 shows that average ages at prior births do not significantly differ between the two cohorts. If we could observe the youngar cohort for 10 more years, i.e. over the same age span as the older cohort, the averages ages of the younger would likely increase. The spacing between births is similar across cohorts as well. While women in transition 2 are about four years older at the second birth, those in transition three are about three years older at second birth and again three years older at the third birth (Again average waiting times could change if we had more observations on cohort 2). The share of cohabiting women seems also rather low, however, also this is a composition effect. Once a women is cohabiting, the chances of having a child are high, hence, these observations leave transition 1.

Figure 1.4 shows that the SOEP data cover the fertile life of the women born in 1960 to 1974 (left-hand figures) and born in 1975 to 1989 (right-hand figures) well. In all figures, the horizontal line starts from the first age for which we observe a woman who is at risk to transit to the next parity and ends at age 50. In Figure 1.4a and 1.4b, where the horizontal runs from age 18 to age 50 or age 40 (the last age observed of those born in 1975 in 2014), respectively, the vertical axis measures the number of women observed who are at risk to make the first transition.[18] In principle, a person enters our data set at the age of 18 and stays inside until she completed her first, second, and third pregnancy. However, we do not observe all women from the first year at risk onwards. Nevertheless, the cohort and subsample definitions are still coarse enough that the entire fertile circle is present, to ensure meaningful estimation.[19]

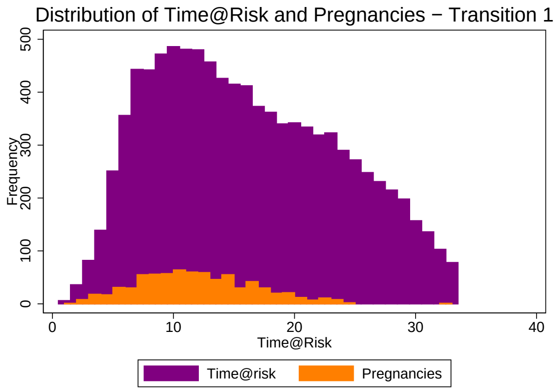
Moreover, the number of pregnancies are shown. The frequency of pregnancies indicates the number of women who left the sample at a specific age before age 50. The figures show that for the first transition, we have fewer observations for women younger than 25 and older women in particular for the younger cohort. For the second and third transition, the number of observations is very high and covers the fertile cycle very well for the older cohort, while for the younger cohort, fewer observations are available for aged women. Note that the number of pregnancies is very small for the third transition, which makes the identification of our estimates difficult.

[18] Expanding or shortening the definition of the fertile cycle does not significantly change the results. Nevertheless we wanted to exclude under-aged women because of the presumably large share of unwanted pregnancies. Extending the fertile horizon beyond 49 does not add many observations. We decided to cut all observation two years after the last observed pregnancy per sub-group.

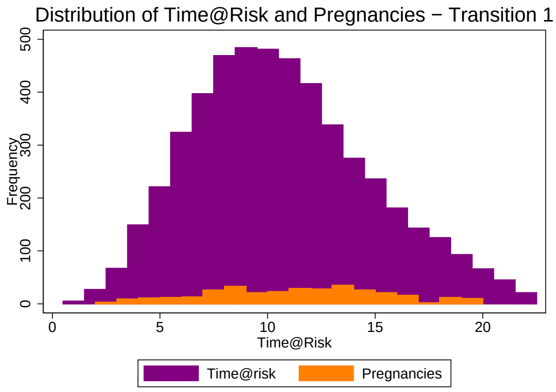
[19] Our sample hence differs to the data in (Heckman and Walker, 1990c) because we can only include those years in which we have information on the variables which we use to construct the predicted measures of income and risk. Hence, we cannot extrapolate based on the birth histories.

Table 1.1: Summary Statistics - By Cohorts

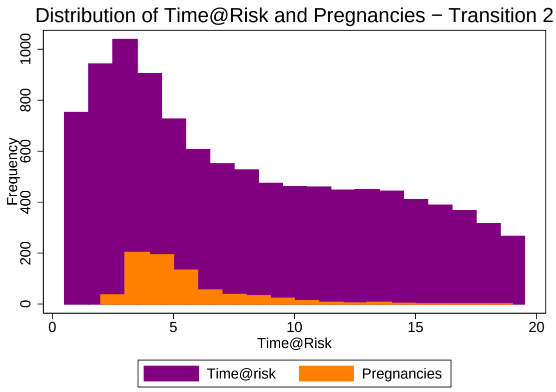
	Cohort 1 - Transition 1				**Cohort 2 - Transition 1**			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Years at risk	12.97	8	1	32	8.85	5.2	1	25
Educ female	12.43	2.6	7	18	12.85	2.64	7	18
Educ male	13.22	2.67	8.5	18	13.47	2.56	7	18
Net Inc male	28964	7315	16851	50102	25693	5539	16855	48579
Net Wage female	8.9	1.19	7.19	12.69	8.65	1.06	7.19	12.69
Risk male	0.041	0.013	0.021	0.094	0.051	0.014	0.027	0.095
Risk female	0.017	0.01	0.008	0.068	0.014	0.006	0.008	0.052
Cohabiting	0.43	0.5	0	1	0.33	0.47	0	1
College	0.29	0.46	0	1	0.36	0.48	0	1
Age	29.97	8	18	49	25.85	5.2	18	42
Age at 1st birth	29.25	5.11	18	49	28.24	4.42	18	39
N	2030 cases, 742 births, 9719 $N \times T$				1535 cases, 329 births, 5008 $N \times T$			
	Cohort 1 - Transition 2				**Cohort 2 - Transition 2**			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Years at risk	7.28	5.21	1	19	4.87	4.03	1	19
Educ female	12.1	2.4	7	18	12.12	2.42	7	18
Educ male	12.78	2.57	7	18	12.86	2.59	7	18
Net Inc male	32235	6993	20297	55568	31420	6310	20297	55384
Net Wage female	9.52	1.26	7.34	14.22	9.5	1.2	7.31	14.2
Risk male	0.031	0.008	0.018	0.076	0.034	0.009	0.019	0.077
Risk female	0.019	0.008	0.008	0.06	0.016	0.007	0.008	0.045
Cohabiting	0.75	0.43	0	1	0.73	0.44	0	1
College	0.24	0.43	0	1	0.25	0.44	0	1
Age	33.73	6.62	18	49	30.06	5.22	18	42
Age at 1st birth	26.45	5.06	17	42	25.18	4.55	17	39
Age at 2nd birth	30.45	4.77	19	43	29.12	4.38	19	40
N	2275 cases, 749 births, 10514 $N \times T$				1197 cases, 363 births, 3558 $N \times T$			
	Cohort 1 - Transition 3				**Cohort 2 - Transition 3**			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Years at risk	8.34	4.87	1	18	5.09	3.59	1	18
Educ female	12.22	2.42	7	18	11.88	2.37	7	18
Educ male	13.05	2.68	7	18	12.6	2.59	7	18
Net Inc male	37418	8700	20716	62493	34611	7380	21126	61977
Net Wage female	10.16	1.43	7.5	15.43	9.79	1.31	7.37	15.32
Risk male	0.033	0.008	0.019	0.084	0.035	0.009	0.021	0.088
Risk female	0.022	0.008	0.007	0.058	0.018	0.006	0.008	0.043
Cohabiting	0.79	0.41	0	1	0.84	0.36	0	1
College	0.25	0.44	0	1	0.21	0.41	0	1
Age	37.12	5.56	20	49	32.38	4.59	20	42
Age at 1st birth	25	4.21	17	40	23.88	4.01	17	36
Age at 2nd birth	28.78	4.22	18	43	27.29	4.06	19	40
Age at 3rd birth	31.9	4.67	21	43	30.89	3.86	22	39
N	2498 cases, 229 births, 14419 $N \times T$				1061 cases, 121 births, 3648 $N \times T$			



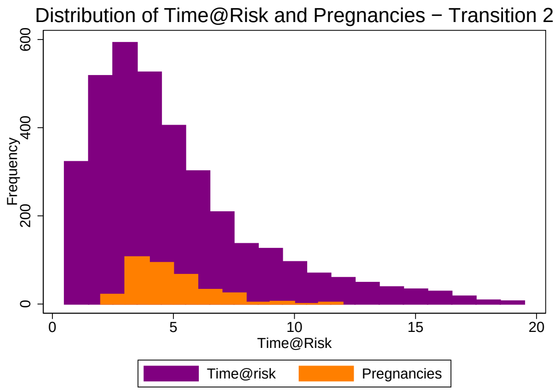
(a) First Parity of Cohorts 1960-1974



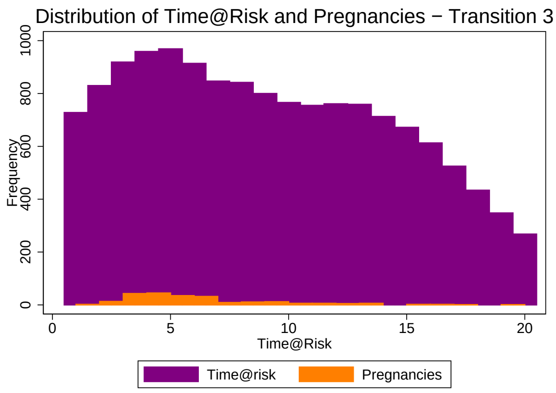
(b) First Parity of Cohorts 1975-1989



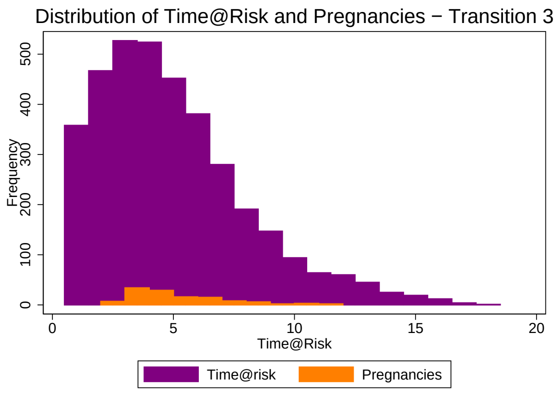
(c) Second Parity of Cohorts 1960-1974



(d) Second Parity of Cohorts 1975-1989



(e) Third Parity of Cohorts 1960-1974



(f) Third Parity of Cohorts 1975-1989

Figure 1.4: Distribution of Time@risk and Pregnancies for all Three Transitions

Source: Authors' calculations.

1.4.5 Income, Wage and Risk over the Life Cycle

Figure 1.5 displays the previously described income measures over the life cycle. We will use the log of the measures for estimation, for the graphs however, we decided to use the levels in order to ease interpretation. Note that all income measures (and hence all risk measures) are deflated to year 2010 values; thus, all values are *real*. The measures are parity and cohort-specific and always refer to the age of the women. The first cohort are women born between 1960 and 1974, the second cohort is born between 1975 and 1989.[20]

Both, the potential wage rate (Figure 1.5a) and the household permanent income (Figure 1.5b) increase over age. The wage rate seems to plateau somewhere in the late thirties, a feature which is less pronounced for household income. The potential wage rate does not show any striking differences across parities, the net household income however differs considerably across parities, which can partially be explained by child allowances.

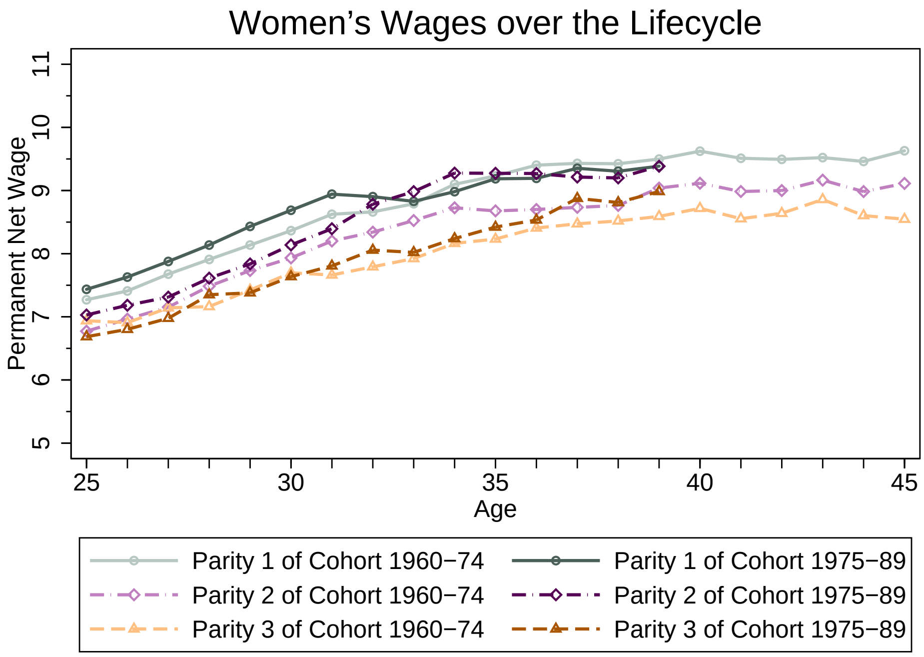
In contrast, there are strong differences in the female wage risk (Figure 1.6a) across parities: The risk for women waiting for the first transition is well below those of higher parities. Wage risk for the first transition increases slightly with age, the risk measures for higher parities on the other hand decrease with the womens age. Differences between cohorts seem to be rather small.

Male household income risk (Figure 1.6b) increases with the women's age. In sharp contrast to wage risk, the average risk is highest for women in the first and lowest for women in the third spell. Income seems to be considerably riskier for the younger cohort in spell 1 and 2, but not in the third spell.[21]

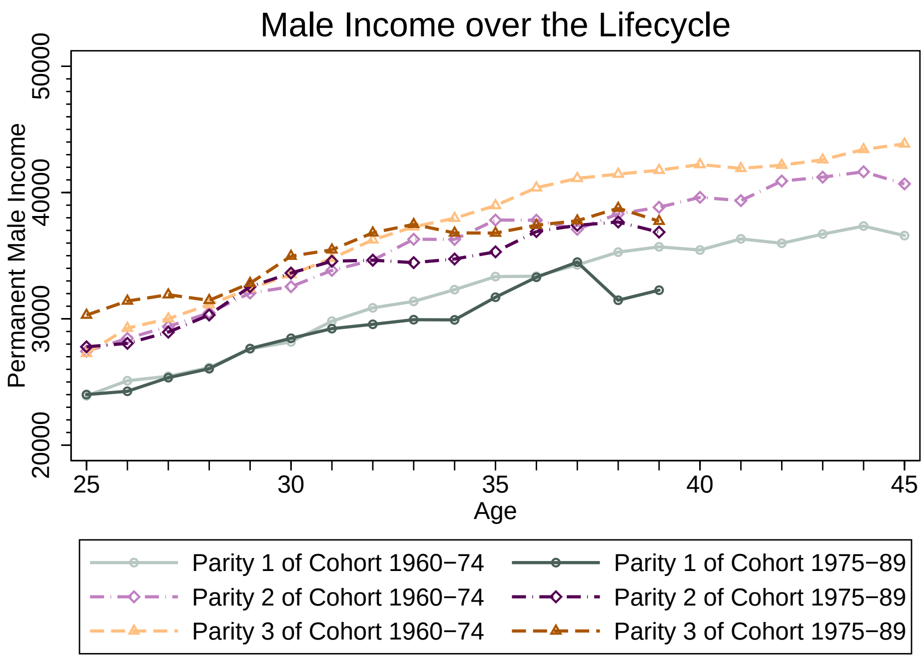
Figure 1.A1 in the Appendix displays the parity-specific development of the four variables of interest over calendar years. Both, the potential wage rate and the male income increase considerably over time. Female wage risk appears rather stable, but male income risk appears to have increased for all parities since the late 1990s.

[20] Graphs without the cohort distinction but separated for educational attainment are shown in the Appendix (Figures 1.A2 and 1.A3).

[21] Note that the jump for male income risk for parity 1 is due to the few observations shortly before age 40.



(a) Women's Potential Wages over the Lifecycle

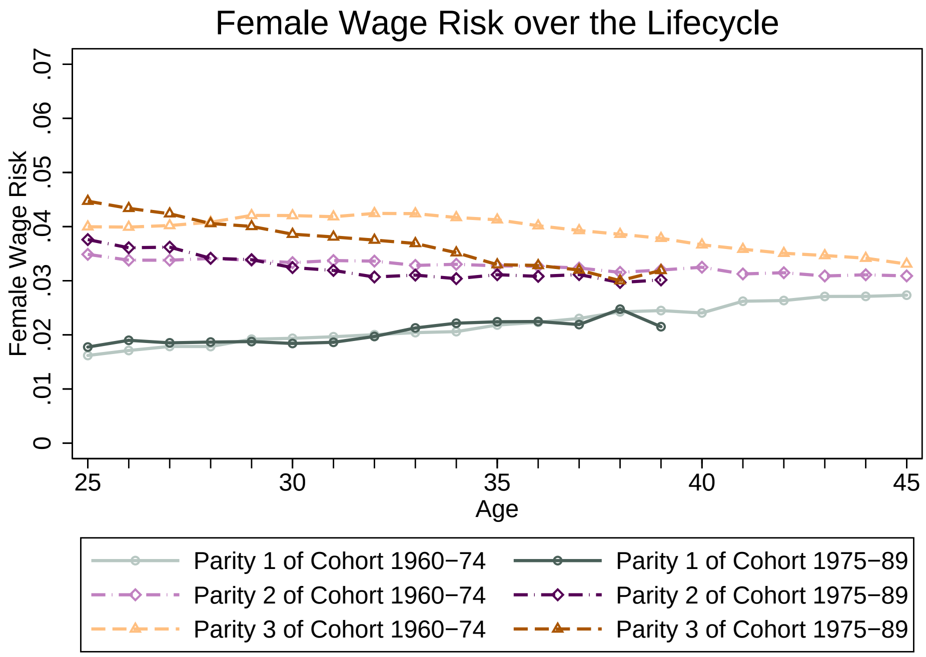


(b) Expected Household Income over the Lifecycle

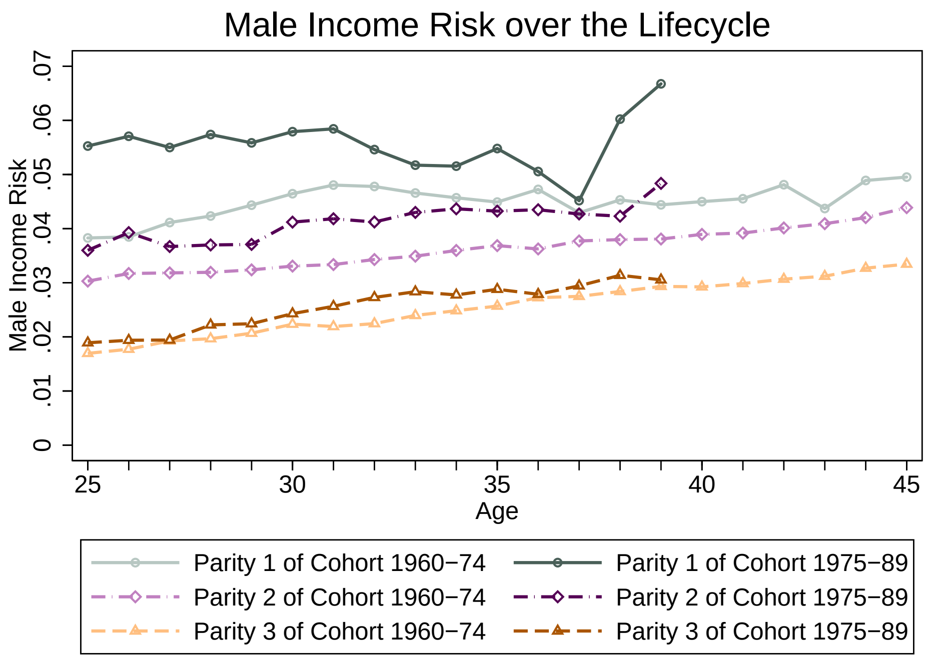
Figure 1.5: Permanent Income Measures Over Age

Note: For both graphs, age refers to the female partner.

Source: Authors' calculations.



(a) Female Wage Risk over the Lifecycle



(b) Household Income Risk over the Lifecycle

Figure 1.6: Income Risk Measures Over Age

Note: For both graphs, age refers to the female partner.

Source: Authors' calculations.

1.5 Results

1.5.1 Pooled Sample

Table 1.2 shows the results from the pooled regression of the first three spells using the entire sample.[22] Coefficients of the cubic spline baseline hazard and further control variables (years of education for both partners, cohabitation indicator, length of the previous spells for transition 2 and 3) are omitted. The upper panel of Table 1.2 displays the estimated coefficients from the link function (1.6) for the first three fertility transitions. The lower panel displays the estimates of the unobserved heterogeneity (mass points, factor loads, probabilities). For all tested specifications, two mass points was the maximum number of mass points, which still led to convergences.

Table 1.2: Results - Entire Sample - Joint Estimation of all three Transitions

	Transition 1	Transition 2	Transition 3
Wage female	.1254	.8289***	-.1856
	(.3212)	(.3061)	(.3810)
Income male	.2004	.0272	-.0059
	(.1865)	(.2013)	(.2899)
Risk female	-.0954	.3603**	.5132
	(.0994)	(.1584)	(.2816)
Risk male	-.0083	-.2837***	-.0945
	(.0862)	(.1054)	(.1499)
Unobserved Heterogeneity			
Factor Load 2	2.3406		
Factor Load 3	0.4689		
Mass Point 1	-0.1178		
Mass Point 2	0.6303		
Pr(M1)	0.8426		
Pr(M2)	0.1574		

$N \times T$ 46871; 8240 distinct cases
Clustered standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For the first and the third transition in the joint model, there are no significant effects of wage, income and the associated household risk. For the transition from the

[22]The associated descriptive statistics can be found in the Appendix in Table 1.A1.

first to the second child however, all, but the household income become statistically significant. Both, female wages and female wage risk appear to have *ceteris paribus* a positive effect on the transition to the second child. We interpret this as evidence that women with a higher earnings potential or with a riskier wage rate actually prefer to have a tighter birth spacing between child number one and two. Since women in Germany most frequently have only two children, finding a job after the second child might be easier, since employers do not fear that women will be soon on another parental leave. The effect of (male) household income however, is negative, thus, couples with a uncertain income seem to transition at lower rates. This corroborates Sommer's finding of a negative relation between the number of births and husband's income risk, even though Sommer used a different risk measure taken from Saks and Shore (2005) and a different estimation method.

1.5.2 Cohort Drift

Table 1.3 shows the estimates from a fully interacted model, which estimates the effects of all covariates separately for the first cohort (women born between 1960 and 1974) and the second cohort (women born between 1975 and 1989). Otherwise, the structure of the table follows the one of Table 1.2. The results for the first cohort are very similar to the ones of the joint sample from Table 1.2: There are no statistically significant effects for the first and the second transition; for the second transition however, female wage rate and its riskiness appear to increase the hazard rate. However, there is no significant effect from the male income risk.

The patterns look different for the second cohort. Female wage risk has a negative effect on the first transition, which corresponds to the finding from the previous literature that labor market integration precedes family formation. For the second transition, female wage rate and its riskiness are no longer statistically significant; however, the point estimates are quite large and in the ballpark of those for the first cohort. The effect from the riskiness of male household income on the second transition however is strongly significant.

In order to ease the interpretation of the estimated coefficients, Figures 1.7 to 1.10 display the survivor functions, which could be termed more appropriately for our context as *no-(further)-child*functions, for the case of a cohabiting woman at the average covariates of all cohabiting women for the two types of of the unobserved heterogeneity. In order to asses the effect of the economic variables, we also show the resulting change of the survivor function due to one additional standard deviation of the respective variable. The advantage of this representation is that it enables us to show how the effect evolves over time and to provide insights whether the effect primarily induces a shift in the timing, or whether there are lasting differences in the probability of having no (further) child at the end of the time at risk. This means that there are two main ways to read the figures: First, a difference between intercepts of the no-(further)-child-function with and without an additional standard deviation of risk with the vertical axis at the last reported

Table 1.3: Results - pooled sample - Split by Cohorts, jointly estimated

	Cohort 1			*Cohort 2*		
	Transit. 1	Transit. 2	Transit. 3	Transit. 1	Transit. 2	Transit. 3
Wage female	.5278	.9416***	-.3110	-.1402	.7499	.1887
	(.3700)	(.3515)	(.4796)	(.6444)	(.5272)	(.6463)
Income male	.2789	-.0470	-.2530	.5514	.3065	.6184
	(.2246)	(.2433)	(.3595)	(.3379)	(.3531)	(.5253)
Risk female	.1050	.3167*	.4128	-.3680**	.5015	.9911*
	(.1183)	(.1874)	(.3667)	(.1795)	(.3116)	(.5808)
Risk male	.1182	-.0545	-.0792	.0951	-.5348***	-.1904
	(.1086)	(.1359)	(.1963)	(.1529)	(.1740)	(.2495)
Unobserved Heterogeneity						
Factor Load 2	2.6716					
Factor Load 3	0.6154					
Mass Point 1	-0.1009					
Mass Point 2	0.5251					
Pr(M1)	0.8388					
Pr(M2)	0.1612					

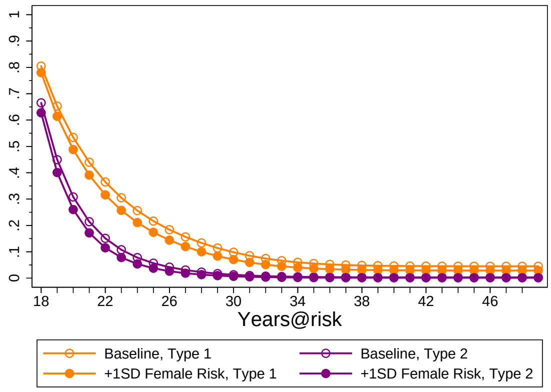
$N \times T$ 46871; 8240 distinct cases
Clustered standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

age, where fertility is completed means that more risk leads to a change in the probability to have a specific number of children. Second, by drawing an imaginary horizontal line starting at 50%, i.e. the median woman, on the vertical axis one can assess that higher risk leads to shorter spacing in this transition if this line crosses the no-further-child function at a younger age with one standard deviation more risk.

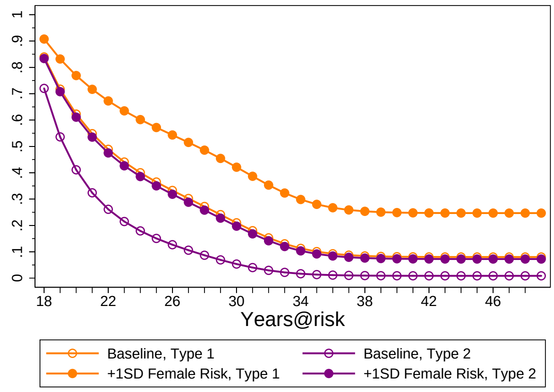
Figures 1.7 to 1.10 compare the four variables of interest for both cohorts for the first two transitions. The results for the third transition are omitted due to the imprecisely estimated point estimates. The corresponding graphs for the joint model from Table 1.2 are shown in the Appendix (Figures 1.A4 and 1.A5).

Figure 1.7a shows that there are two types of women with respect to the pace of transition: A quick group (Type 1) and a slower group (Type 2). The median woman in the quick group makes the transition to parity 1 already after the first year of being at risk. The median woman of the slower group waits until age 21 or 22 to make the transition. The estimates at the bottom of Table 1.3 for Pr(M1) and Pr(M2) indicate that 84% of the women belong to the slower group and 16% to the quicker group.

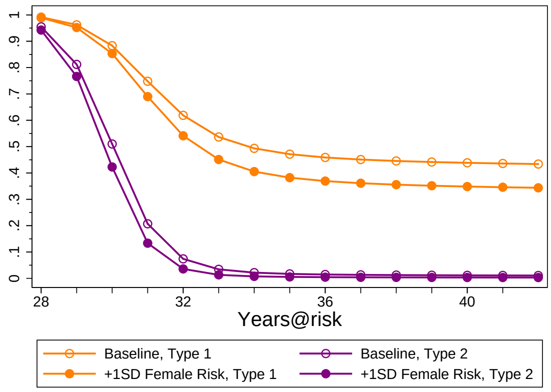
The figure shows for each group two lines, one with filled circles and one with unfilled circles. The line with unfilled circles corresponds to the of a cohabiting woman at the average covariates of all cohabiting women. The filled circles show a new no-(further)-child-function that results after increasing female wage risk by one standard deviation



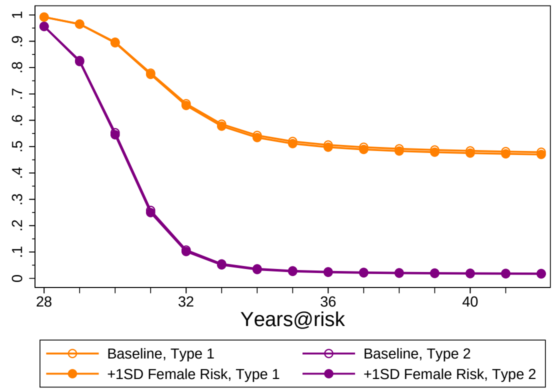
(a) First Parity of Cohorts 1960-1974



(b) First Parity of Cohorts 1975-1989



(c) Second Parity of Cohorts 1960-1974



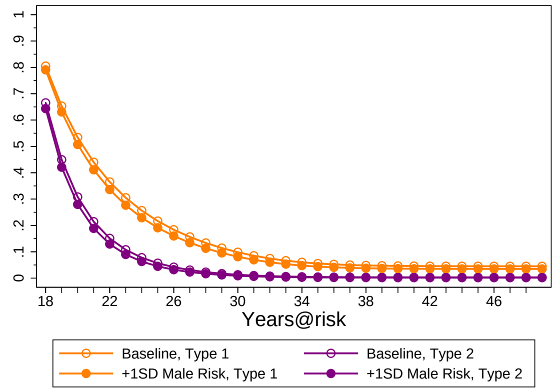
(d) Second Parity of Cohorts 1975-1989

Figure 1.7: Impact of Higher Risk of Women's Wage on Survivor Functions by Cohorts

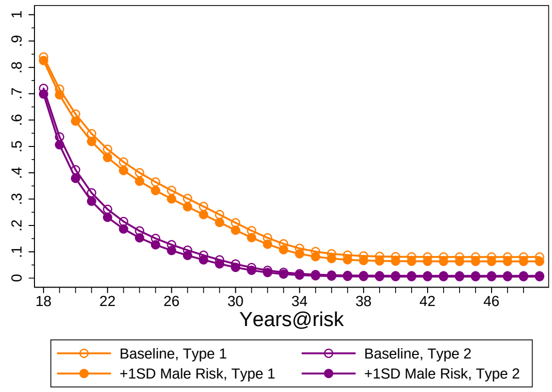
Source: Authors' calculations.

(see Table 1.1) keeping everything else constant. While a substantially riskier wage did not affect the fertility decisions of women the older cohort to make the first transition, such an increase in wage risk causes the median woman of the quick type to defer having the first child by three years in the younger cohort (Figure 1.7b). This delay causes only a small increase in the fraction of childless women among this group. For the majority of slower women, the delay is much more dramatic: the median woman postpones having her first child by six years due to higher own wage risk. The fraction of childless women increases substantially from 8% to 25%.

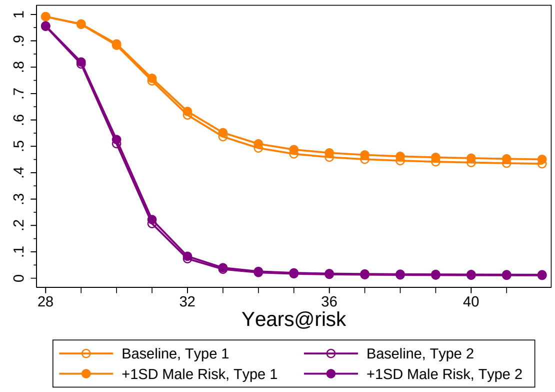
Figures 1.7c and 1.7d show how transitions to the second parity are affected by the same one-standard deviation increase in females wage risk. For the older cohort, the quick type is hardly affected. However, the slower majority of women reduce spacing between the first and second child. The median woman in this group decides to have a second child at age 32-33 instead at age 34. The fraction of women who have on child but do not have a second reduces from 45% to 35% due to higher risk. Note that the



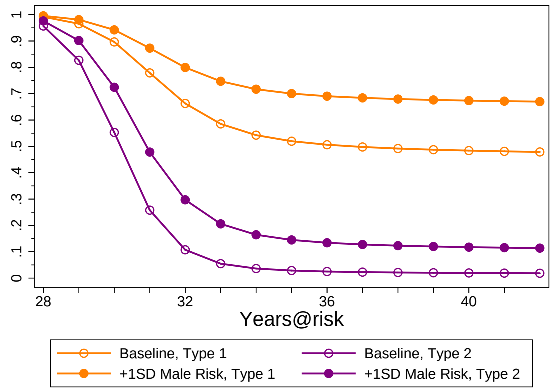
(a) First Parity of Cohorts 1960-1974



(b) First Parity of Cohorts 1975-1989



(c) Second Parity of Cohorts 1960-1974



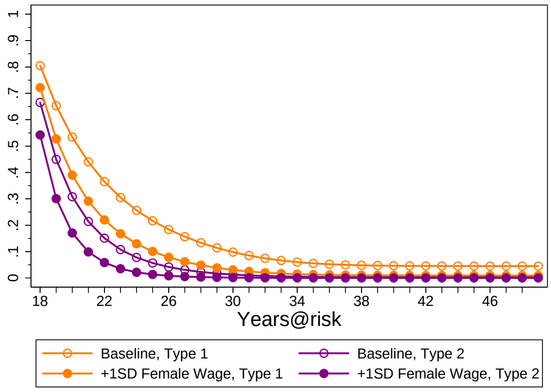
(d) Second Parity of Cohorts 1975-1989

Figure 1.8: Impact of Higher Risk of Household Income on Survivor Functions by Cohorts

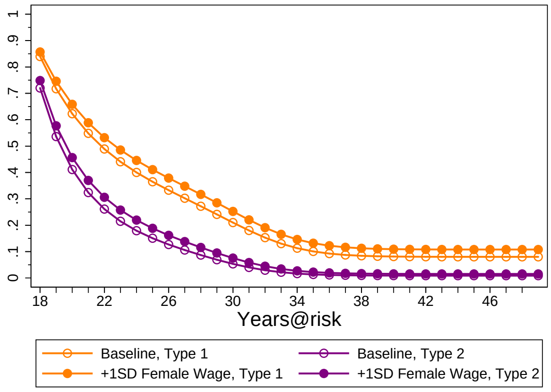
Source: Authors' calculations.

difference in types is much more pronounced in transition 2 as reflected by the fact that the estimate of Factor Load 2 in Table 1.3 is larger than 1.

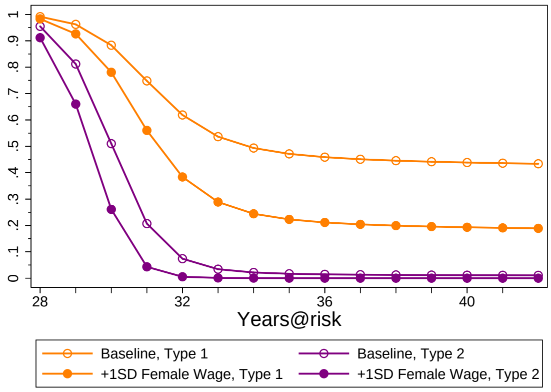
Figure 1.8 presents the for both types the no-(further)-child-function with unfilled circles denoting the line of a cohabiting woman at the average covariates of all cohabiting women and filled circles marking the resulting no-(further)-child-function after an one standard-deviation increase in male income risk (see Table 1.1). This ceteris paribus experiment has virtually no effect for either cohort and both types, except for women of the younger cohort in transition 2. For this group higher income risk leads to a postponement of having a second child and to an increase in the fraction of mothers of one child who will never have a second child. For the median woman of the quick type this postponement amounts to about one year. The fraction of mothers who will not have two children increases to 10%. For the median mother among the 84% of slow women, higher male income risk means that she changes her mind and will never have a



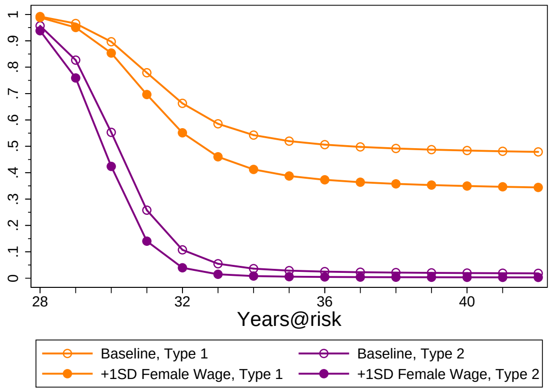
(a) First Parity of Cohorts 1960-1974



(b) First Parity of Cohorts 1975-1989



(c) Second Parity of Cohorts 1960-1974



(d) Second Parity of Cohorts 1975-1989

Figure 1.9: Impact of Higher Wage of Women on Survivor Functions by Cohorts

Source: Authors' calculations.

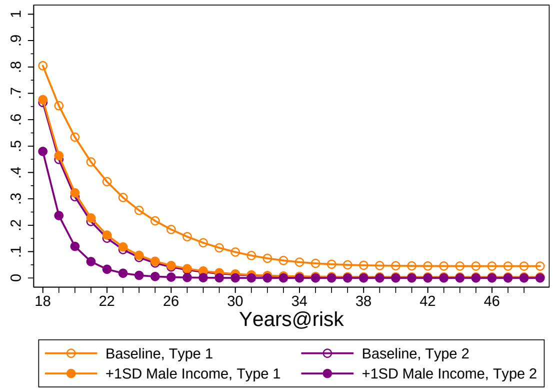
second child. The overall fraction of those who will not have a second child in this group increases from 50% to 65%.

Figure 1.9 illustrates the effects of an increase of female wage potential by one standard deviation. For transition 1 there is no significant postponement effect and also no significant effect on the probability to stay childless in both cohorts.

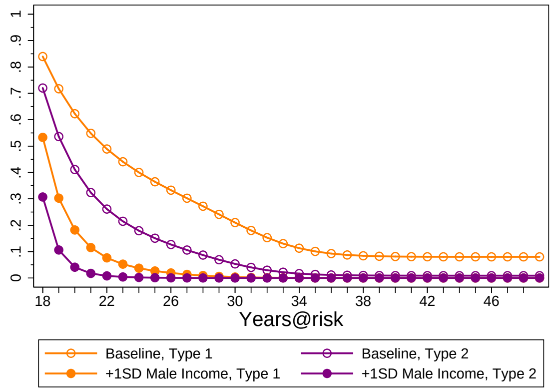
For the second transition, higher opportunity costs of child rearing reduces spacing distance to the firstborn by three (one) years for the median of the slow majority of women (the small group of quick women).

For the younger cohort this preponement is even four years for the median woman of the slow group. The fraction of mothers who will never have a second child reduces from 45% to 20% for the older and to 35% for the younger cohort.

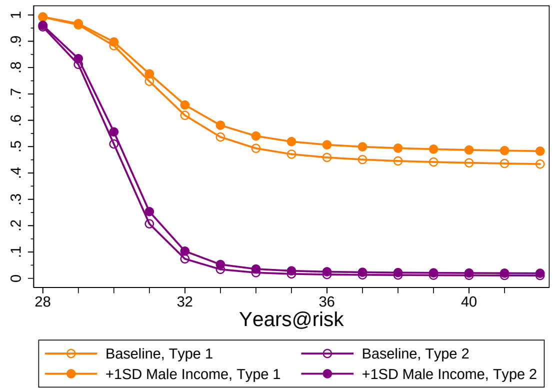
Figure 1.10 shows that the insignificant effects of a substantial increase of the male income (by one standard deviation) is also economically not important for the older cohort. For the younger cohort higher male income leads to a three year preponement of



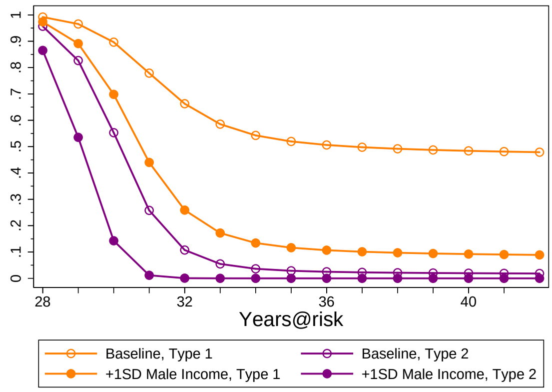
(a) First Parity of Cohorts 1960-1974



(b) First Parity of Cohorts 1975-1989



(c) Second Parity of Cohorts 1960-1974



(d) Second Parity of Cohorts 1975-1989

Figure 1.10: Impact of Higher Household Income on Survivor Functions by Cohorts

Source: Authors' calculations.

the decision to become a mother for the median woman of type 1. Moreover, all women would eventually make this decision.

For the second transition of the younger cohort there is a stronger preponement effect of six years for the median mother of type 1. The probability to not have a second child reduces from 50% to 10%. However, due to the sample size, these results have to be interpreted with caution.

1.5.3 Educational Drift

Alternatively to the split of the sample between cohorts, we also consider a split of the sample by educational attainment. For Germany, it is natural to distiguish women based on whether they have a university degree or not. Low fertility rates among highly educated women have been very prominent in the German public debate. Table 1.4 thus shows the results for such an educational split. The corresponding survivor functions are shown in the Appendix (Figure 1.A6 to 1.A9).

For the college-educated women, the only significant economic variable is the one of female risk for the first transition. This is again in line with previous research about the fertility behavior of well-educated women, for which establishing yourself in the labor market seems to be important before founding a family. The effect of the female potential wage rate is also statistically significant for women without a university degree, however, the effect size is about only half of the size of the one for women with a college degree.

For women without college degree, the potential wage rate is negatively associated for the first, but positively associated for the second transition. The result for the first transition can be explained by specialization to a male-breadwinner household, the positive effect for the second transition again by a tighter spacing between the first and second child argument. The risk of the household income is negative for the second transition. This is again as in Sommer (2016).

Table 1.4: Results - Entire Sample - Split by Education, jointly estimated

	No College			*College*		
	Transit. 1	Transit. 2	Transit. 3	Transit. 1	Transit. 2	Transit. 3
Wage female	-.7230*	1.1246***	.0119	.2528	.4028	-.3097
	(.4191)	(.3980)	(.5088)	(.6332)	(.5090)	(.6943)
Income male	.3637	-.1049	.1128	-.1009	.3371	-.2627
	(.2288)	(.2377)	(.3248)	(.3421)	(.3989)	(.6564)
Risk female	-.2712**	.5345***	.5795*	-.4973**	.0758	.8561
	(.1332)	(.2013)	(.3403)	(.1959)	(.3251)	(.6218)
Risk male	.0648	-.3083**	.0225	-.1698	-.2570	-.4119
	(.1044)	(.1235)	(.1669)	(.1616)	(.2151)	(.3292)
Unobserved Heterogeneity						
Factor Load 2	2.2535					
Factor Load 3	0.3956					
Mass Point 1	-0.1260					
Mass Point 2	0.6701					
Pr(M1)	0.8417					
Pr(M2)	0.1583					

$N \times T$ 46871; 8240 distinct cases
Clustered standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.5.4 Robustness Checks

Table 1.A3 in the Appendix shows the results of the joint model from Table 1.2 for a plain complementary log-log model without controlling for unobserved heterogeneity and for a random effects cloglog model with normally distributed random effects. For both models, standard errors are clustered on the individual level. The results indicate that the specification of the unobserved heterogeneity does not seem to affect the results; for both cases, the estimated coefficients and the pattern of statistically significant results are very similar to those from Table 1.2.

Table 1.A4 shows the results for alternative specifications of the unobserved heterogeneity following Heckman and Singer, first, separately transition by transition, second, in a joint specification of all three transitions, but without factor loads. The third transition estimated alone did not converge and is hence not shown. Nevertheless, also these changes leave the estimation results qualitatively and quantitatively unchanged.

1.6 Conclusion and Discussion

In this paper we study how household income, the female earnings potential and the associated uninsured, idiosyncratic measures of risk affect womens fertility transitions. Using data from the German Socio-Economic Panel for the years 1984 to 2014 we estimated joint duration models for the first three fertility transitions, taking into account unobserved heterogeneity.

We focus on the difference between cohorts, the first one born between 1960 and 1974, the second one between 1975 and 1989. For the older cohort we do not find any significant effects on the transition to the first child. For the younger cohort however, female wage risk leads to significantly longer time to the first child. For the transition to the second child, women from the older cohort with higher income opportunities tend to have shorter spacing between the first and the second child. For the younger cohort, we do not find such a significant effect for the younger cohort; however, for these women, the riskiness of the male income significantly reduces the probability of having a second child. We do not find any striking effects for the transition to the third child, which only rarely occurs in Germany. These results are mainly driven by low to medium educated women. For women with a college degree, only the riskiness of the female wage rate did significantly reduce the hazard rate for the first transition.

1.7 Appendix

Additional Tables

Table 1.A1: Summary Statistics - Overall Sample

	Transition 1				Transition 2				Transition 3			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Years at risk	14.3	7.33	1	33	7.37	4.93	1	19	8.5	5.17	1	20
Net Wage female	8.26	1.54	3.47	13.55	8.1	1.66	4.7	14.63	8.16	1.73	5.27	15.68
Net Inc male	28752	8532	10928	65574	34084	9558	16598	77205	38756	10578	16708	81541
Risk female	0.024	0.034	0.006	1.122	0.035	0.013	0.01	0.159	0.038	0.011	0.011	0.215
Risk male	0.049	0.031	0.006	0.444	0.036	0.018	0.003	0.222	0.027	0.014	0.004	0.194
Educ female	12.82	2.6	7	18	11.99	2.3	7	18	12.06	2.4	7	18
Educ male	12.61	2.75	7	18	12	2.54	7	18	12.2	2.71	7	18
College	0.30	0.46	0	1	0.22	0.41	0	1	0.21	0.41	0	1
Cohabiting	0.49	0.5	0	1	0.77	0.42	0	1	0.88	0.32	0	1
Age	31.3	7.33	18	50	33.04	6.43	19	50	36.47	6.06	19	50
Age at 1st birth	29.72	4.54	18	50	26.15	4.87	18	50				
Age at 2nd birth					30.28	4.56	19	43	24.78	4.02	18	40
Age at 3rd birth					31.5	4.32	20	42	28.45	4.11	18	43
	3565, 1071 births 14727 $N \times T$				3472 cases, 1112 births 14072 $N \times T$				3559 cases, 350 births 18067 $N \times T$			

Table 1.A2: Summary Statistics - By Education

	No college - Transition 1				**College - Transition 1**			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Years at risk	12.97	8	1	32	8.85	5.2	1	25
Educ female	12.43	2.6	7	18	12.85	2.64	7	18
Educ male	13.22	2.67	8.5	18	13.47	2.56	7	18
Net Inc male	28964	7315	16851	50102	25693	5539	16855	48579
Net Wage female	8.9	1.19	7.19	12.69	8.65	1.06	7.19	12.69
Risk male	0.041	0.013	0.021	0.094	0.051	0.014	0.027	0.095
Risk female	0.017	0.01	0.008	0.068	0.014	0.006	0.008	0.052
Cohabiting	0.43	0.5	0	1	0.33	0.47	0	1
College	0.29	0.46	0	1	0.36	0.48	0	1
Age	29.97	8	18	49	25.85	5.2	18	42
Age at 1st birth	29.25	5.11	18	49	28.24	4.42	18	39
N	2626 cases, 769 births, 10163 $N \times T$				1132 cases, 302 births, 4564 $N \times T$			
	No college - Transition 2				**College - Transition 2**			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Years at risk	7.28	5.21	1	19	4.87	4.03	1	19
Educ female	12.1	2.4	7	18	12.12	2.42	7	18
Educ male	12.78	2.57	7	18	12.86	2.59	7	18
Net Inc male	32235	6993	20297	55568	31420	6310	20297	55384
Net Wage female	9.52	1.26	7.34	14.22	9.5	1.2	7.31	14.2
Risk male	0.031	0.008	0.018	0.076	0.034	0.009	0.019	0.077
Risk female	0.019	0.008	0.008	0.06	0.016	0.007	0.008	0.045
Cohabiting	0.75	0.43	0	1	0.73	0.44	0	1
College	0.24	0.43	0	1	0.25	0.44	0	1
Age	33.73	6.62	18	49	30.06	5.22	18	42
Age at 1st birth	26.45	5.06	17	42	25.18	4.55	17	39
Age at 2nd birth	30.45	4.77	19	43	29.12	4.38	19	40
N	2620 cases, 820 births, 10711 $N \times T$				924 cases, 292 births, 3361 $N \times T$			
	No college - Transition 3				**College - Transition 3**			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Years at risk	8.34	4.87	1	18	5.09	3.59	1	18
Educ female	12.22	2.42	7	18	11.88	2.37	7	18
Educ male	13.05	2.68	7	18	12.6	2.59	7	18
Net Inc male	37418	8700	20716	62493	34611	7380	21126	61977
Net Wage female	10.16	1.43	7.5	15.43	9.79	1.31	7.37	15.32
Risk male	0.033	0.008	0.019	0.084	0.035	0.009	0.021	0.088
Risk female	0.022	0.008	0.007	0.058	0.018	0.006	0.008	0.043
Cohabiting	0.79	0.41	0	1	0.84	0.36	0	1
College	0.25	0.44	0	1	0.21	0.41	0	1
Age	37.12	5.56	20	49	32.38	4.59	20	42
Age at 1st birth	25	4.21	17	40	23.88	4.01	17	36
Age at 2nd birth	28.78	4.22	18	43	27.29	4.06	19	40
Age at 3rd birth	31.9	4.67	21	43	30.89	3.86	22	39
N	2728 cases, 265 births, 13792 $N \times T$				911 cases, 85 births, 4275 $N \times T$			

Table 1.A3: Results - pooled sample - Alternative Estimation Methods

	Plain Cloglog			*RE Cloglog*		
	Transit. 1	Transit. 2	Transit. 3	Transit. 1	Transit. 2	Transit. 3
Wage female	.1022	.6658***	-.2230	.1066	.6794***	-.1896
	(.3172)	(.2481)	(.3803)	(.3244)	(.2569)	(.3846)
Income male	.1739	-.0031	-.0027	.2013	.0106	-.0212
	(.1844)	(.1783)	(.2897)	(.1879)	(.1845)	(.2929)
Risk female	-.0949	.2850**	.5191*	-.1002	.3068**	.5148*
	(.0984)	(.1353)	(.2806)	(.1007)	(.1409)	(.2853)
Risk male	-.0095	-.2428***	-.0956	-.0135	-.2487**	-.0861
	(.0851)	(.0932)	(.1491)	(.0871)	(.0970)	(.1517)

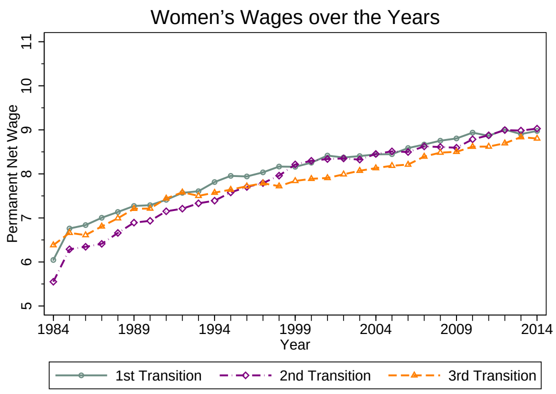
Clustered standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A4: Results - pooled sample - Variations of Heckman and Singer

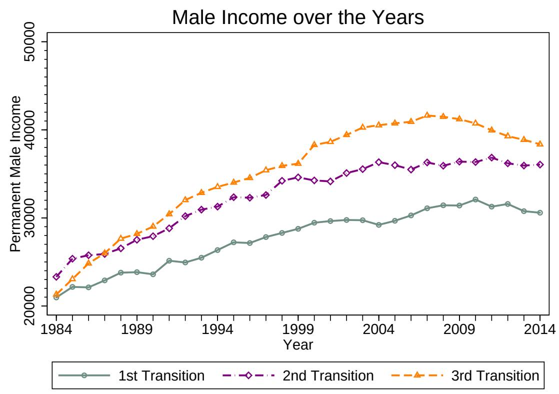
	Transition by Transition		*No Factor Loads*		
	Transition 1	Transition 2	Transition 1	Transition 2	Transition 3
Wage female	.0357	.7567***	.1102	.6817***	-.1903
	(.3309)	(.2799)	(.3255)	(.2588)	(.3847)
Income male	.2189	.0095	.2018	.0100	-.0230
	(.1914)	(.1917)	(.1881)	(.1848)	(.2928)
Risk female	-.1120	.3376**	-.0993	.3073**	.5156*
	(.1032)	(.1518)	(.1006)	(.1411)	(.2854)
Risk male	-.0119	-.2773***	-.0133	-.2516***	-.0848
	(.0877)	(.1009)	(.0871)	(.0973)	(.1517)
	Unobserved Heterogeneity				
Mass Point 1	-13.1413	-0.1959	-0.2952,		
Mass Point 2	1.2142	1.3668	0.4188		
Pr(M1)	0.0846	0.8747	0.5865		
Pr(M2)	0.9154	0.1253	0.4135		
$N \times T$	14728	14073	46871		
# Cases	3567	3472	8240		

Clustered standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Note: For Transition 3 - transition by transition converge not achieved

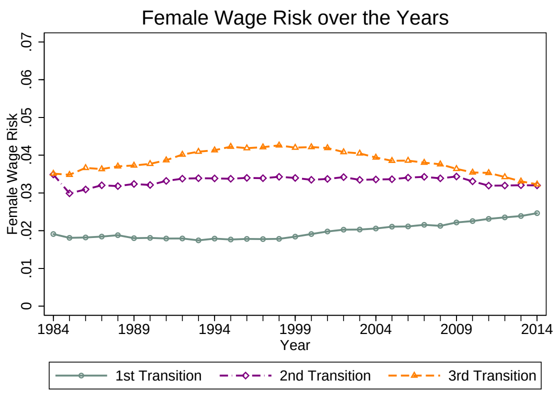
Additional Figures



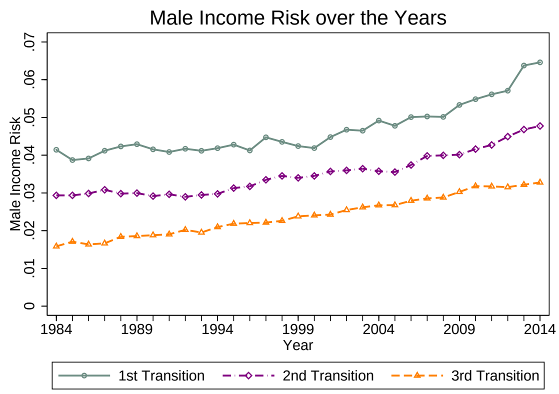
(a) Female Wage over the Years



(b) Male Income over the Years



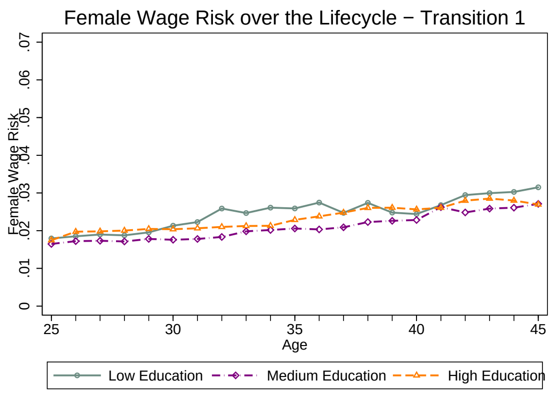
(c) Female Wage Risk over the Years



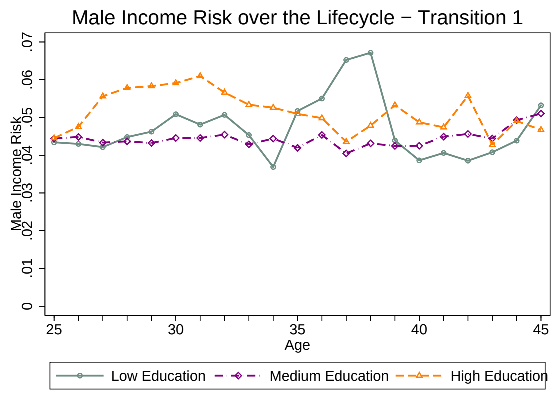
(d) Male Income Risk over the Years

Figure 1.A1: Income, Wages and Risk Measures Over Years

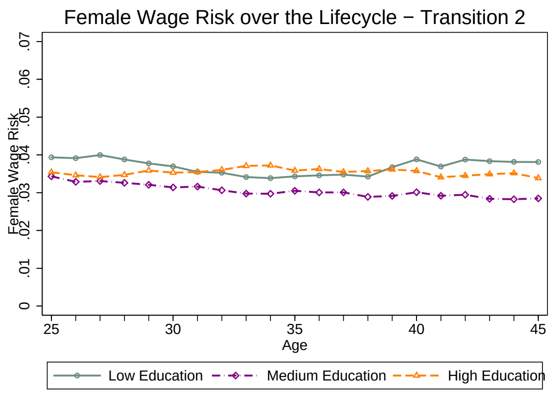
Source: Authors' calculations.



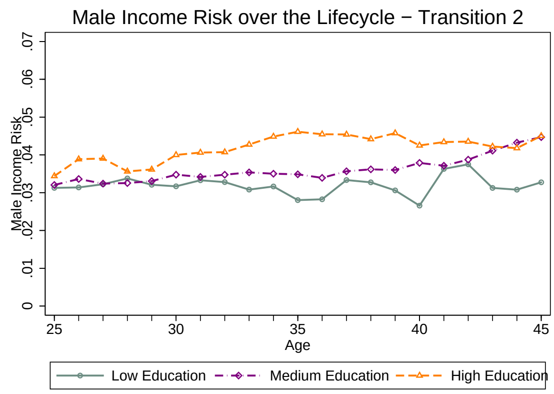
(a) First Parity - Female Wage Risk



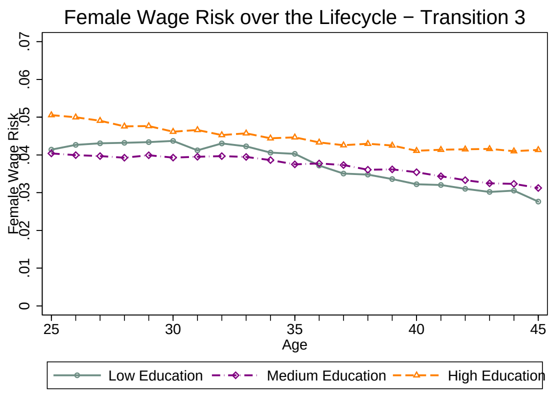
(b) First Parity - Male Income Risk



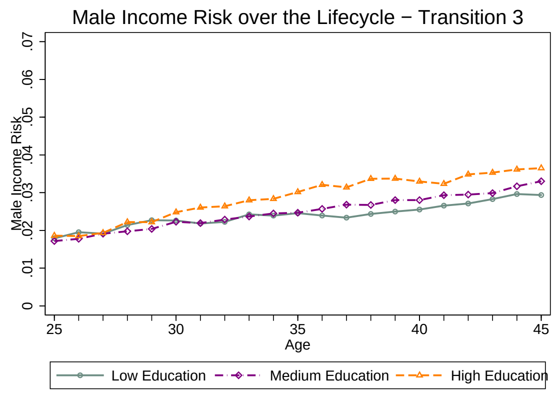
(c) Second Parity - Female Wage Risk



(d) Second Parity - Male Income Risk



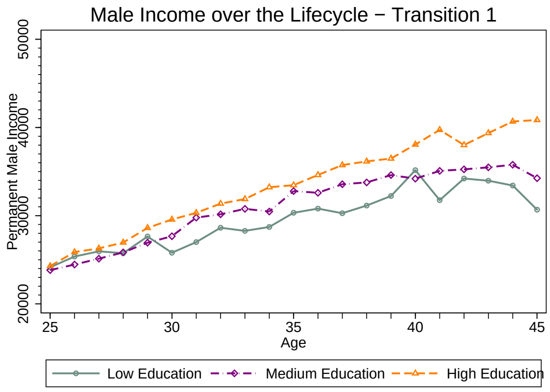
(e) Third Parity - Female Wage Risk



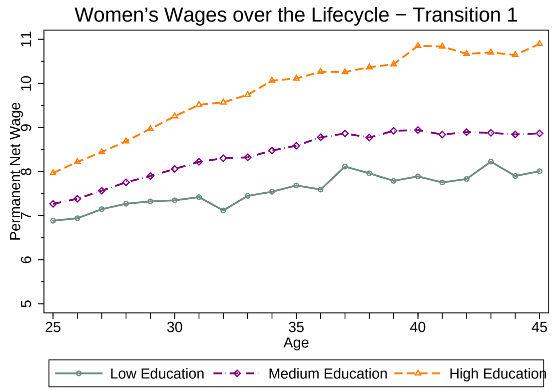
(f) Third Parity - Male Income Risk

Figure 1.A2: Female Wage Risk and Male Income Risk Over Women's Age and Transitions by Education

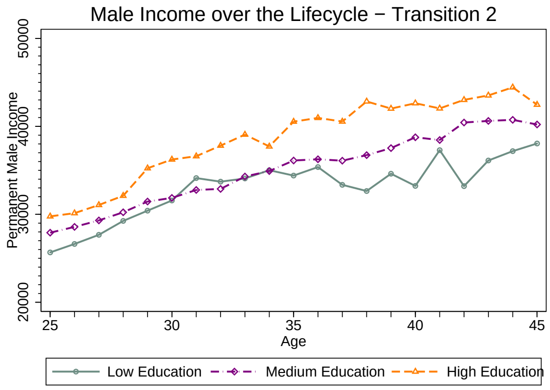
Source: Authors' calculations.



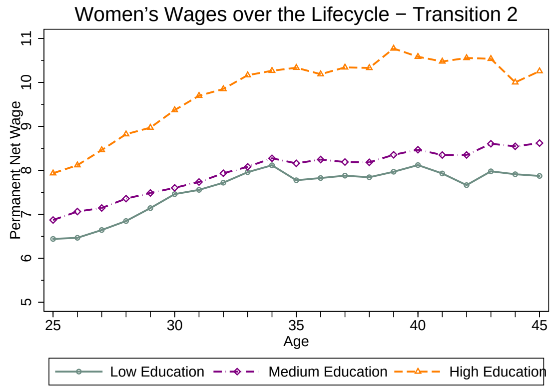
(a) First Parity - Wage



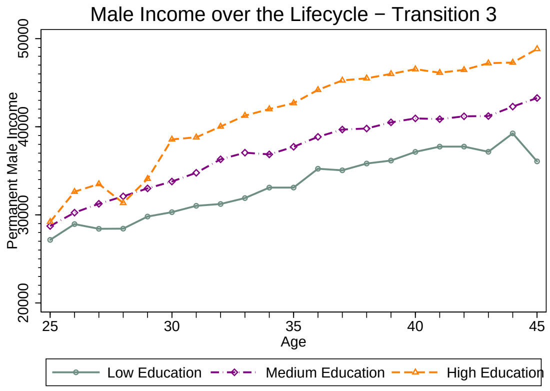
(b) First Parity - Income



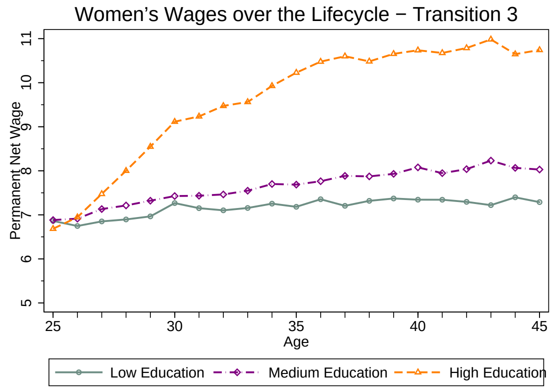
(c) Second Parity - Wage



(d) Second Parity - Income



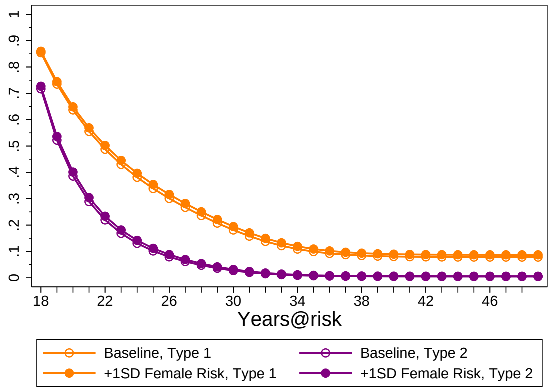
(e) Third Parity - Wage



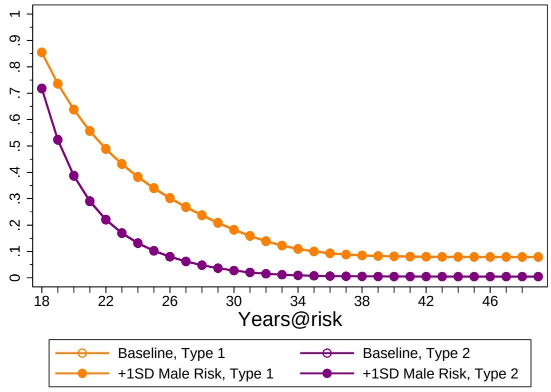
(f) Third Parity - Income

Figure 1.A3: Female Wage Potential and Male Permanent Income Over Age and Transitions by Education

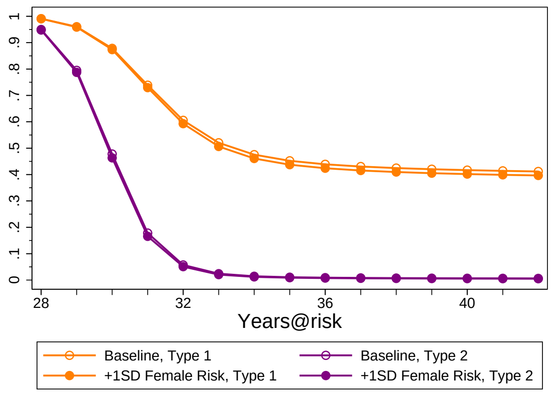
Source: Authors' calculations.



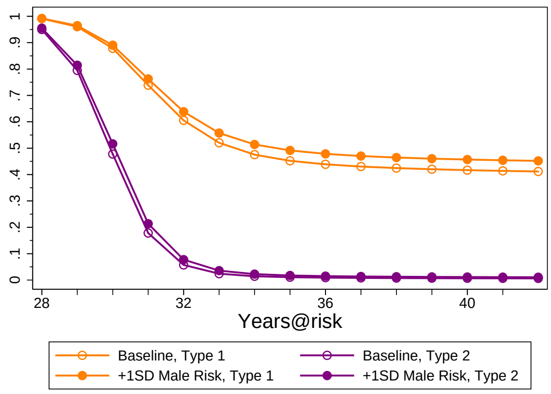
(a) First Parity - Female Wage Risk



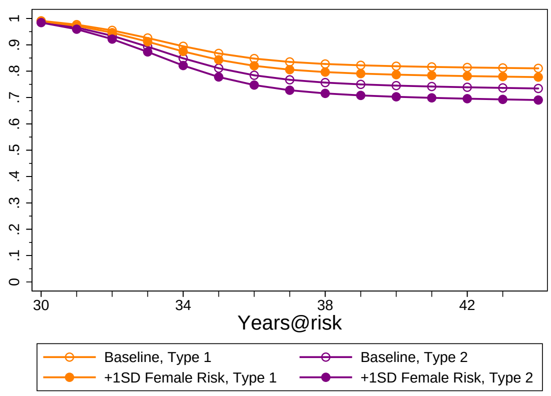
(b) First Parity - Male Income Risk



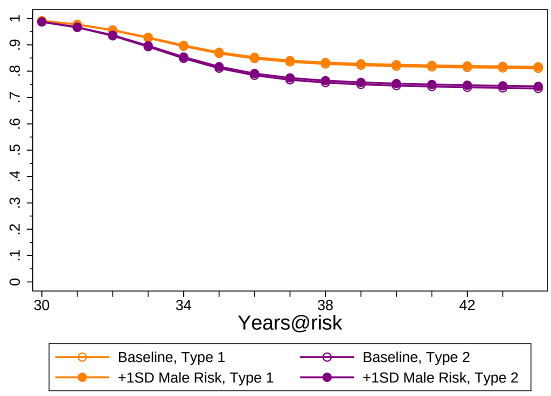
(c) Second Parity - Female Wage Risk



(d) Second Parity - Male Income Risk



(e) Third Parity - Female Wage Risk



(f) Third Parity - Male Income Risk

Figure 1.A4: Impact of Higher Risk on Survivor Functions

Source: Authors' calculations.

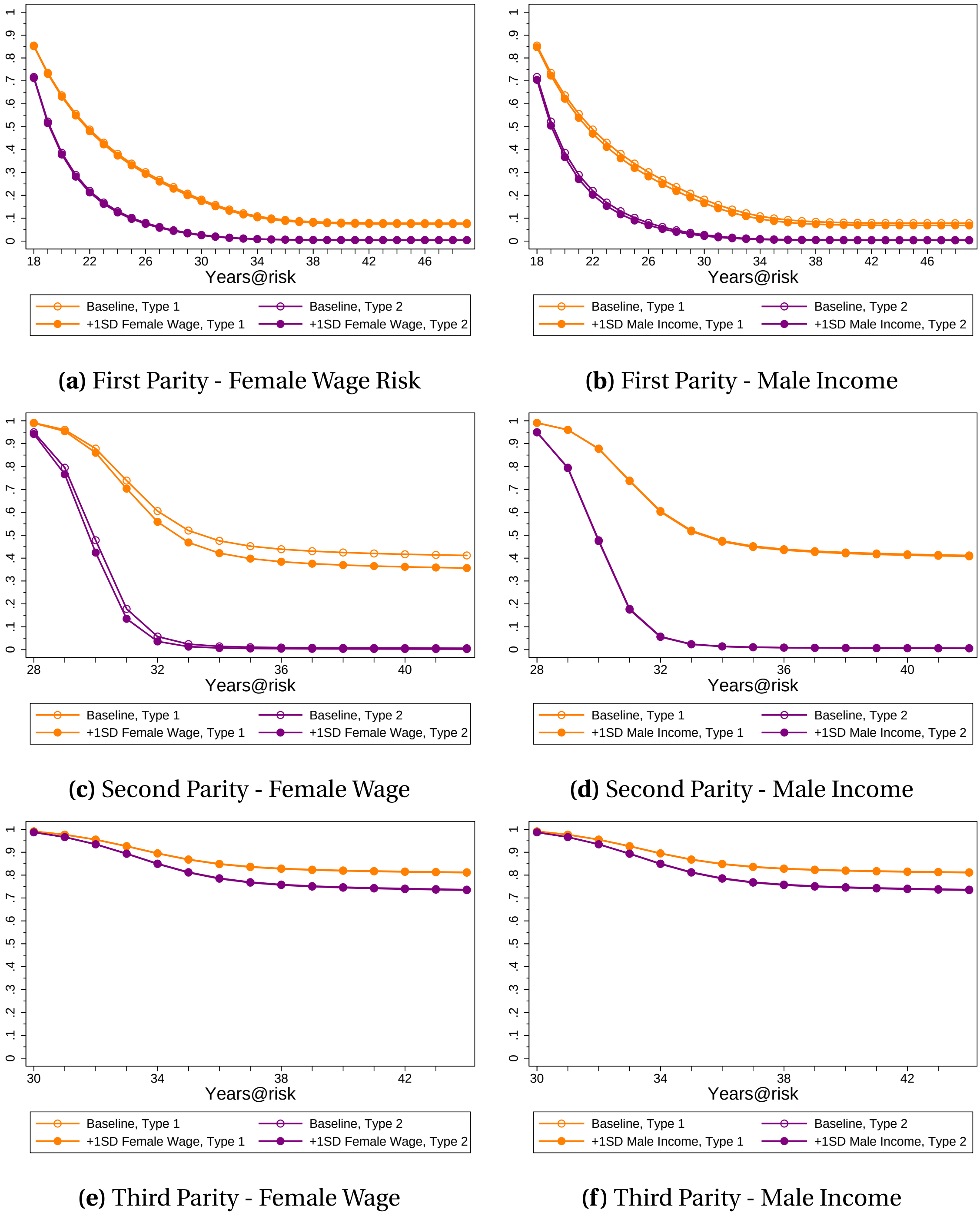
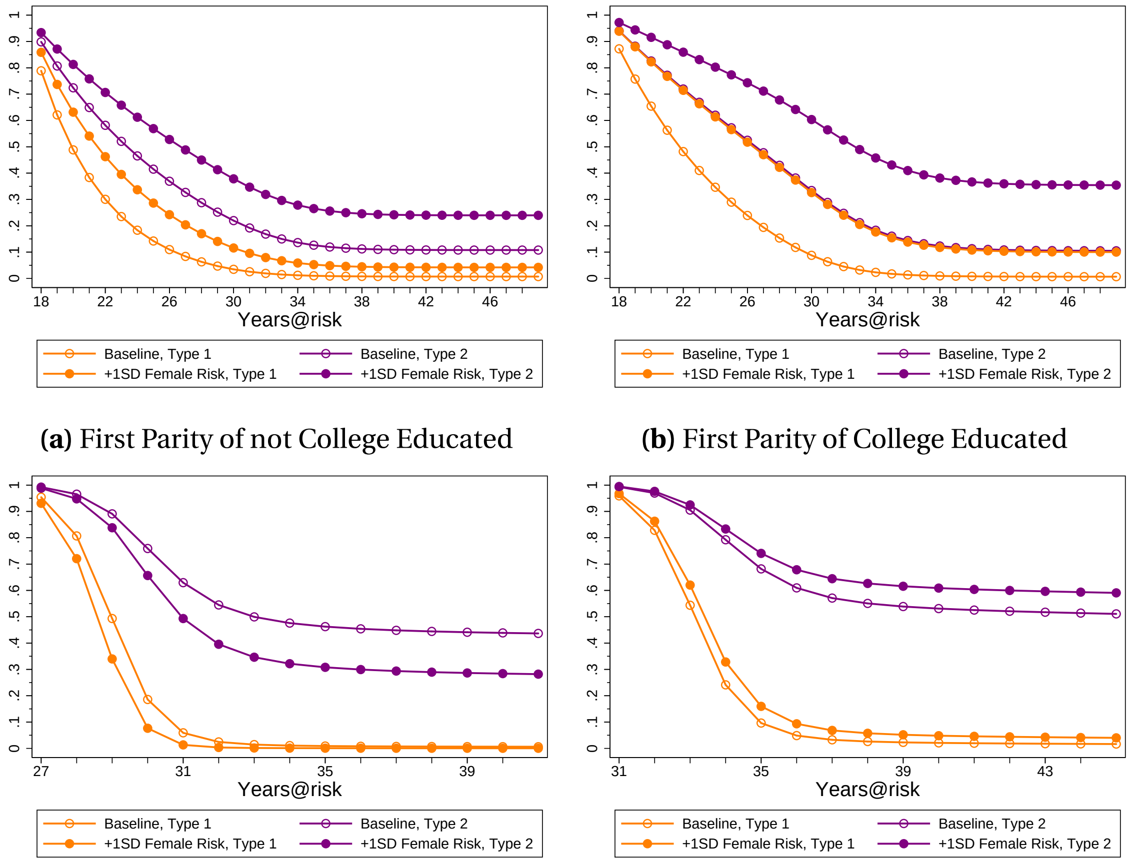


Figure 1.A5: Impact of Higher Wage and Household Income on Survivor Functions

Source: Authors' calculations.



(a) First Parity of not College Educated

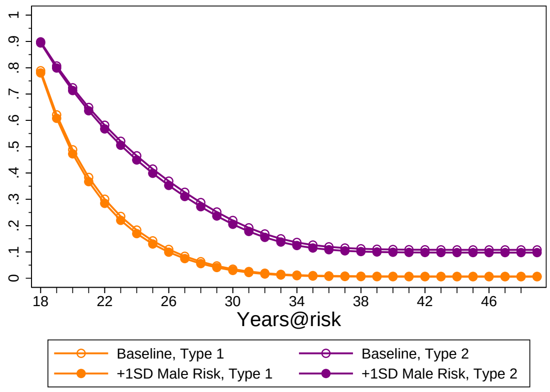
(b) First Parity of College Educated

(c) Second Parity of not College Educated

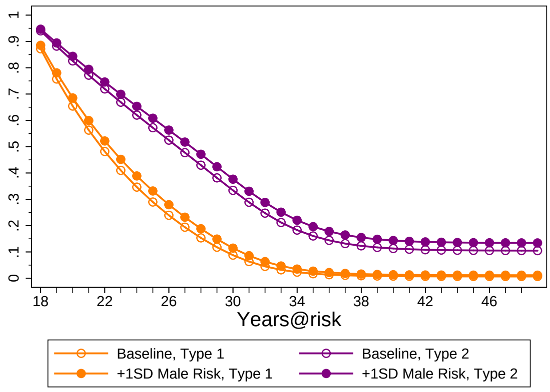
(d) Second Parity of College Educated

Figure 1.A6: Impact of Higher Risk of Women's Wage on Survivor Functions by Education

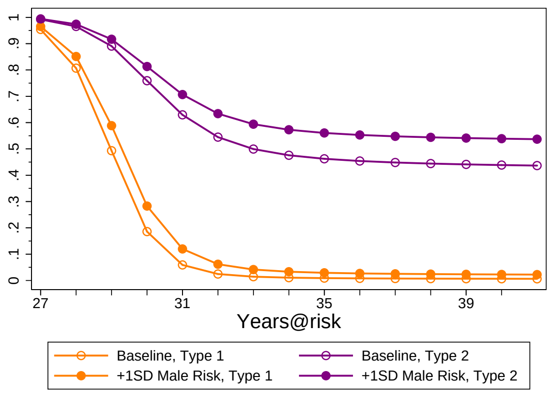
Source: Authors' calculations.



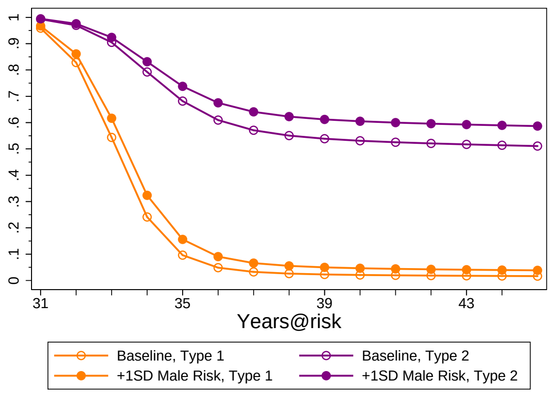
(a) First Parity of not College Educated



(b) First Parity of College Educated



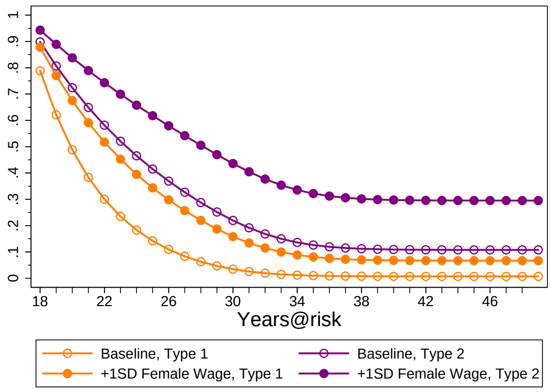
(c) Second Parity of not College Educated



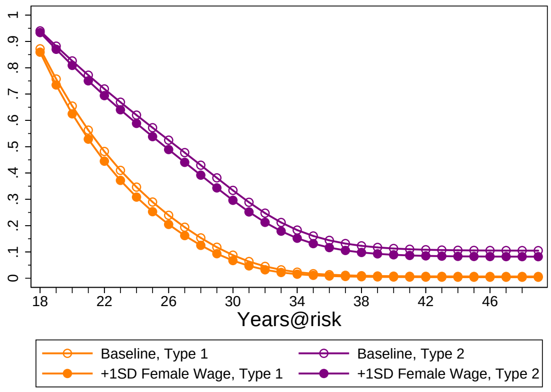
(d) Second Parity of College Educated

Figure 1.A7: Impact of Higher Risk of Household Income on Survivor Functions by Education

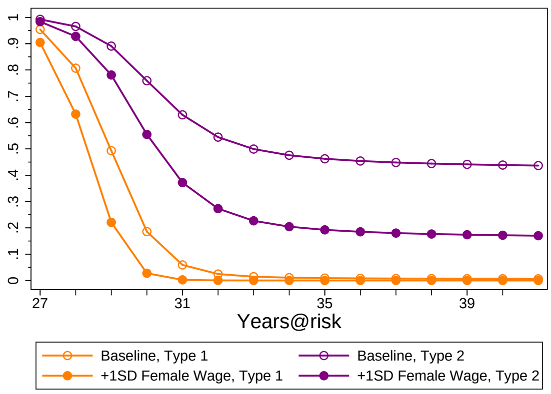
Source: Authors' calculations.



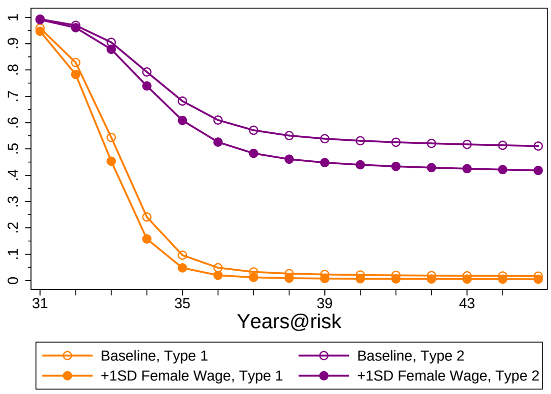
(a) First Parity of not College Educated



(b) First Parity of College Educated



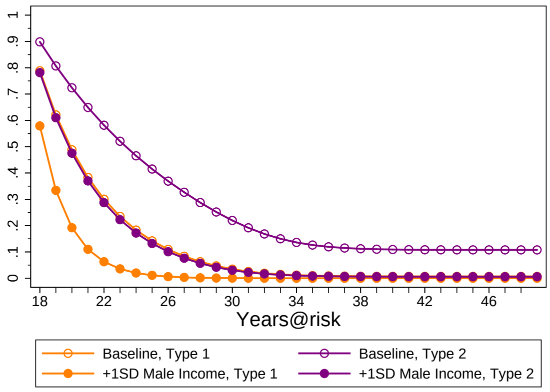
(c) Second Parity of not College Educated



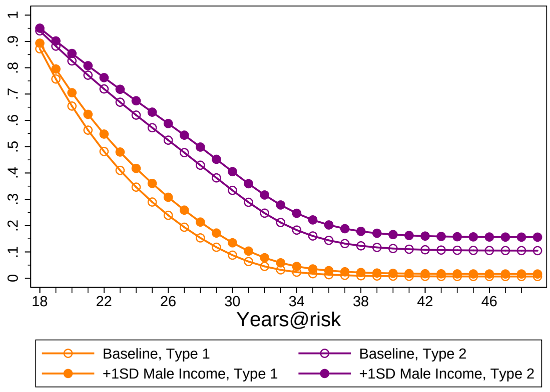
(d) Second Parity of College Educated

Figure 1.A8: Impact of Higher Wage of Women on Survivor Functions by Education

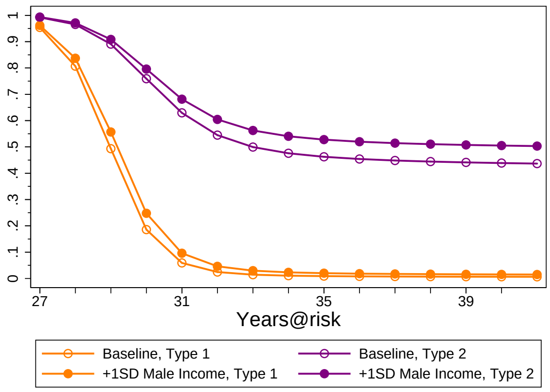
Source: Authors' calculations.



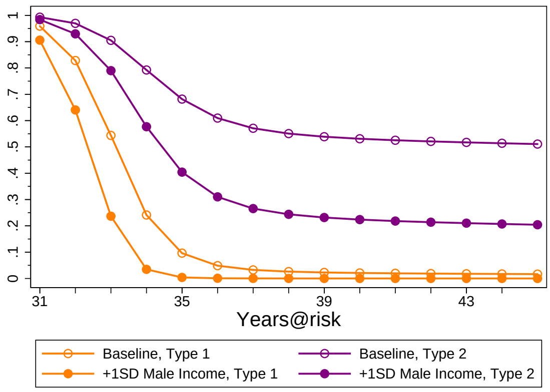
(a) First Parity of not College Educated



(b) First Parity of College Educated



(c) Second Parity of not College Educated



(d) Second Parity of College Educated

Figure 1.A9: Impact of Higher Household Income on Survivor Functions by Education

Source: Authors' calculations.

2 How Important is Precautionary Labor Supply?

2.1 Introduction

This study quantifies the importance of precautionary labor supply, defined as the difference between hours supplied in the presence of risk and hours supplied under perfect foresight. Facing a higher future wage risk, individuals may increase their hours worked in order to insure themselves against bad realizations. Our study provides empirical evidence for this theoretically predicted phenomenon. We examine how strongly labor supply adjusts in response to higher wage risk by focusing on the partial equilibrium case similarly to Carroll and Samwick (1998) or Parker and Preston (2005) for consumption.

A thorough intuition of labor supply incentives over the life cycle is crucial for understanding household behavior and is of primary interest for both labor economics and macroeconomics (Meghir and Pistaferri, 2011). Relevant precautionary labor supply could explain differences in hours worked across occupations or why self-employed work more hours than employees for a given wage. The extent of precautionary labor supply is key for various policy issues, for instance the optimal design of social security programs. Our approach allows us to calculate how labor supply would change in partial equilibrium, if self-employed, blue and white collar workers had the same insurance against wage risk as civil servants, for instance through reforms of the social insurance system.

A number of theoretical contributions have studied precautionary labor supply in models with saving (Flodén, 2006; Low, 2005; Pistaferri, 2003). These studies find that individuals facing higher wage risk work more at the beginning of working life in order to accumulate savings. This behavior is governed by the curvature in consumption, i.e. prudence as defined in Kimball (1990), and in leisure of workers' preferences. When leisure is low, not only the marginal utility of leisure is higher, but also the rate at which the marginal valuation rises when leisure falls. This implicates the precautionary motive because of which individuals save more in anticipation of higher future wage risk. With flexible labor supply they do so by consuming less or by working more. The latter concept is precautionary labor supply. Pijoan-Mas (2006) shows that additional hours of work are a quantitatively important smoothing device in a calibration exercise. In our analysis, we abstract from general equilibrium effects which need to be taken into account to assess whether the effect of uncertainty on aggregate output is positive or negative. Marcet et al. (2007) demonstrate that under reasonable parameter configurations a wealth effect that reduces labor supply may dominate the positive precautionary saving effect on aggregate output documented in Aiyagari (1994) and Huggett (1993). Prior to this, studies like

Block and Heineke (1973); Eaton and Rosen (1980a,b), and Hartwick (2000) predicted theoretically that the relationship between uncertainty and labor supply is positive.[1] Still, the actual importance of precautionary labor supply remains an empirical question.

This paper is one of the few studies that provide empirical evidence on this issue. Pistaferri (2003) finds that the effect of wage risk on labor supply agrees with theoretical predictions, but is economically negligible. This might be due to the fact that Pistaferri (2003) used data collected only every two years for Italy in 1989, 91, and 93. In contrast, we are able to construct growth rates from year to year and to exploit a relatively long time dimension (from 2001 to 2012) of the German Socio-Economic Panel (SOEP).

The relationship between (proxies for) wage risk and hours of work has been documented to be positive for self-employed men in the US (Parker et al., 2005), male employees in the US who work more than 30 hours per week (Kuhn and Lozano, 2008), and for German and US workers (including self-employed) of both sexes (Bell and Freeman, 2001). Benito and Saleheen (2013) show that men and women use hours worked to shield themselves against financial shocks, which the authors define as deviations in the subjective perception of their own financial situation, compared to their expectation from the previous year. We contribute to the literature with several innovations.

First, we specify a dynamic labor supply model that allows for partial adjustment of hours worked. Such a specification reflects constraints in the workers' capacity to adjust immediately to their desired level of labor supply. Our findings reject the immediate adjustment model used in previous work.

Second, we calculate marginal net wages using the tax-transfer-microsimulation model STSM (see Steiner et al., 2012).[2] Therefore, in contrast to the previous studies, we are able to account for partial insurance of wage risk through the tax and transfer system as well as through the social insurance system, which may be an important determinant of precautionary behavior, as argued, e.g., in Fossen and Rostam-Afschar (2013). Bell and Freeman (2001) surmise that "[s]ince we have not taken into account differences in the level of social safety nets or taxation [...] our analysis probably understates the effect of inequality in economic rewards on work time". Our results show that this effect is very small.

Third, the result of Pistaferri (2003) that precautionary labor supply is irrelevant might be due to the fact that he used *subjective information on future income* (see also Mastrogiacomo and Alessie, 2014). We examine several measures for wage risk and do not find relevant precautionary labor supply using subjective risk measures either. If wage risk is—as in our analysis—measured by the standard deviation of past hourly individual net wages, however, precautionary labor supply becomes relevant. Moreover, this does not change if risk from other sources than own wages is included or future wages are used. Most of our measures of wage risk assume—following e.g. Blundell and

[1] See Menezes and Wang (2005) for a study that predicts a negative effect of increased wage uncertainty on labor supply if the substitution effect dominates the income effect.

[2] The Steuer-Transfer-Mikrosimulationsmodell (STSM) is comparable to FORTAX for the UK (Shephard, 2009) or TAXSIM for the US (Feenberg and Coutts, 1993).

Preston (1998), Blundell et al. (2008) or Carroll and Samwick (1998) for income—that information unknown to the econometrician is unpredictable for the worker as well.

Fourth, in addition to wage risk and in contrast to previous studies, we investigate the effect of unemployment probability calculated similarly as in Carroll et al. (2003). We find that unemployment probability also increases labor supply, but is quantitatively less important than wage risk.

Finally, we are the first to quantify precautionary labor supply empirically. Individuals in the main sample choose an additional 2.8% of their hours of work to shield against wage shocks, i.e. about one week per year. Precautionary labor supply is particularly important for the self-employed, a group that faces average wage risks substantially above the sample mean. This group works 6.2% of their hours because of the precautionary motive. If self-employed faced the same wage risk as the median civil servant, their hours of work would reduce by 4.5%.

The next section describes our dataset and construction of the measure of wage risk and probability of unemployment. Section 2.3 presents our empirical specification and the estimation methods. Section 2.4 discusses the main results and occupation specific findings. In Section 2.5 we quantify the importance of precautionary labor supply, Section 2.6 shows that the results are robust, and Section 2.7 concludes.

2.2 Data

Our study uses data from the SOEP (version 30), a representative annual panel survey in Germany. Wagner et al. (2007) provide a detailed description of the data. We use observations from 2001-2012 and focus on men because the extensive margin plays an important role in women's labor supply decisions.The sample is restricted to married men between 25 and 56 years old and working at least 20 hours to allow comparisons with the canonical labor supply literature, for example, Altonji (1986), and MaCurdy (1981).[3] Further, we drop persons who indicated having received social welfare payments because their hours choices are likely driven by institutional constraints rather than precautionary motives. We restrict our sample to individuals working less than 80 hours per week. In total, we observe the main wage risk measure for 10,987 data points from 2,488 persons.[4]

Marginal net wage According to economic theory, individuals' labor supply responds to the *marginal* net wage. The reason is that at the optimum the marginal rate of substitution equals the marginal rate of transformation. The marginal net wage is the price at which leisure is transformed into consumption.

[3]Including workers with less than 20 weekly hours virtually does not affect the results.

[4]Table 2.A1 in the Appendix summarizes the number of observations lost due to each sample selection step.

To construct the marginal net wage, first we calculate the hourly gross wage w_{it}^{gross} by dividing annual gross labor income y_{it} by annual hours of work h_{it}:

$$w_{it}^{\text{gross}} = \frac{y_{it}}{h_{it}}.$$

We calculate net income using the microsimulation model STSM. Jessen et al. (2017) present a comprehensive overview of marginal tax rates for different households (for more information, see Steiner et al., 2012). We obtain marginal net wage rates by scaling the gross wage w_{it}^{gross} with the marginal net-of-tax rate. Define the net-of-tax rate as the net of tax income per Euro of additional pretax income due to an increase in hours of work. Then the marginal hourly net wage is given by:

$$w_{it} = \text{Net-of-tax rate} \times w_{it}^{\text{gross}} = \frac{NetInc(y_{it} + \Delta y_{it}) - NetInc(y_{it})}{\Delta y_{it}} w_{it}^{\text{gross}}. \tag{2.1}$$

$NetInc(y_{it})$ denotes net income given gross income y_{it}. To calculate the net-of-tax rate we increase each person's annual labor income y_{it} marginally.[5] In practice, the relevant concept is the net of tax income per additional time spent on work. We assume that this coincides with the marginal net wage as calculated in equation (2.1). This is true if additional hours of work are fully compensated.

For the calculation of hourly wages we use *paid* hours because an increase in these translates directly into an increase in income. To construct paid hours we follow Euwals (2005), accounting for differences in compensation of overtime hours.[6]

Wage Risk We construct measures for both gross and marginal net wage risk. First, in order to remove variations due to predictable wage growth, we detrend log gross wage growth with a regression on age, its square, education, and interactions of these variables, following, for instance, Hryshko (2012). In a second step, we obtain the sample standard deviation of past detrended log wages for each person similarly to Parker et al. (2005). Hence, our risk measure uses only the variation across time for each individual. Only wage observations from the current occupation are used for the construction of the risk measure such that wage risk is not confounded by occupation choices. Thus at least two (not necessarily consecutive) periods of working in the same occupation are needed to construct the risk measure.

[5]We set $\Delta y_{it} = 2000$ Euro, which implies an increase in labor income of about 40 Euro per week.

[6]The SOEP data provide information on overtime compensation or_{it} in the sense whether overtime was (a) fully paid, (b) fully compensated with time off, (c) partly paid, partly compensated with time off, or (d) not compensated at all. $I(or_{it} = a)$ is an indicator function, in this case indicating that overtime rule (a) applies. We approximate paid hours of work as $h_{it} = hc_{it} + I(or_{it} = a)(ht_{it} - hc_{it}) + 0.5I(or_{it} = c)(ht_{it} - hc_{it})$, where hc_{it} are contracted hours of work and ht_{it} are actual hours of work (Euwals, 2005).

The wage risk measure is given by:

$$\sigma_{w,it} = \sqrt{\frac{1}{\#-1}\sum_{j=t-\#}^{t-1}(\ln \tilde{w}_{ij} - \ln \bar{\tilde{w}}_i)^2}, \tag{2.2}$$

where $\tilde{w}_j$ denotes the detrended (net) wage and # denotes the number of past realizations of wage. The idea behind this measure is that workers use past variations in idiosyncratic wages to form expectations about future risk. As we only use past information, we may treat this measure as exogenous at the moment of the labor supply decision. We denote this measure by $\sigma_{w,it}$. For the estimations, we standardize the risk measure by one standard deviation of the sample used in the regression to facilitate interpretation. We provide robustness tests with different risk measures, such as forward looking, five-year rolling windows, without detrending, using only continuous wage spells, subjective risk measures, other household income risk, and including occupational changes in Section 2.6.

Our measure of wage risk assumes following e.g. Blundell and Preston (1998) or Blundell et al. (2008) that information unknown to the econometrician is unpredictable for the worker as well. Cunha et al. (2005) developed a method that distinguishes information unknown to the econometrician but predictable by the agent from information unknown to both. Applications of this method, see e.g. Cunha and Heckman (2008), Navarro (2011), Cunha and Heckman (2016), Navarro and Zhou (2017), show that equating variability with uncertainty results in overstated risk. To separate the information sets, correlation between choices and future realizations of the stochastic variable may be used.

As in Fossen and Rostam-Afschar (2013), we divide our sample into blue collar workers, white collar workers, civil servants, and self-employed. We are mainly interested in decisions during work life at ages where occupational changes are rare. Nonetheless, we model the selection into occupations as a robustness test in the Appendix.

Figure 2.1 shows how the average net wage risk evolves over the life cycle for each subgroup. We use age groups of three years to obtain a sufficient number of observations for each data point. Only age-occupation combinations with more than 15 observations are displayed, thus the trajectory for self-employed starts at age 35. We find that wage risk decreases slightly over the life cycle for all groups. This is more pronounced for the self-employed. The finding is in line with results in Blundell et al. (2015) who find that income risk decreases over the life cycle in Norway.

As expected, the hourly wages of self-employed workers are more volatile over the entire life cycle than those of employees. At all ages this difference is statistically significant at the 5% significance level.[7] Blue and white collar workers have similar levels of

[7]We use a two-sample t test with unequal variances to obtain the p-values. Test statistics are available from the authors upon request.

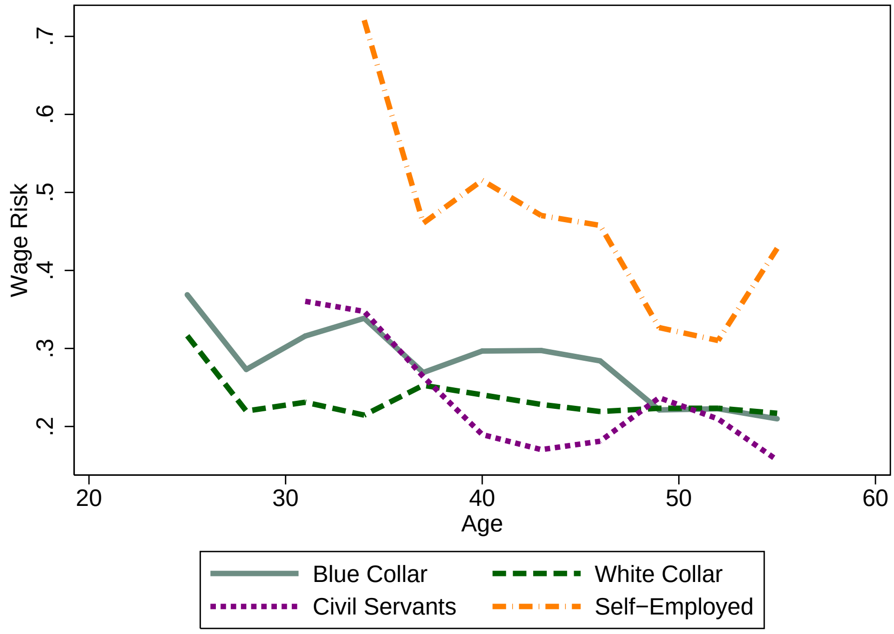


Figure 2.1: Average Net Wage Risk over the Life Cycle
Note: Standard deviations of past marginal net wages for each individual averaged over three years by occupation. We calculate the risk measure for every age for every individual based on past realizations and take the average of this measure over individuals for every age. See equation (2.2). *Source:* Own calculation based on the SOEP

wage risks. Nonetheless, during their 30s and 40s blue collar workers face a statistically significantly higher wage risk than white collar workers. For most age groups, the average net wage risk of civil servants is slightly lower than those of blue collar and white collar workers. This difference is statistically significant at most ages starting in the 40s.

Unemployment Probability The control variable unemployment probability $\Pr_{U,it}$ is the predicted probability to be out of work in the next year. The estimation procedure is similar to the one used by Carroll et al. (2003).[8] Figure 2.2 displays how the average unemployment probability evolves over the life cycle for the four occupational groups.[9] Civil servants have the lowest average unemployment probability, followed by white collar workers. For most parts of the life cycle, blue collar workers face the highest average

[8]We use a heteroskedastic probit model (cf. Harvey, 1976) to estimate the probability of unemployment in the following year conditional on regressors for occupation, industry, region, education, age, age squared, age interacted with occupation as well as with with education, marital status, and unemployment experience. The heteroskedasticity function includes previous unemployment experience and years of education.

[9]As in Figure 2.1, only age-occupation combinations with more than 15 observations are displayed.

unemployment probability. The mean unemployment probabilities of the occupational groups are statistically significantly different at all ages at the 5% level except for the difference between blue collar workers and self-employed at younger ages and white collar workers and self-employed at older ages. As for the wage risk, we standardize the unemployment probability by its standard deviation for the estimations.

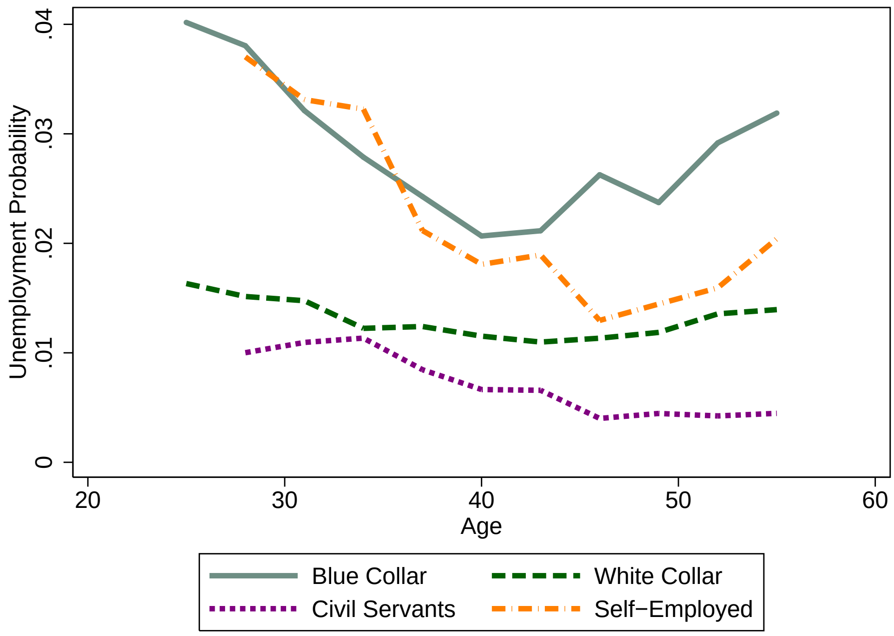


Figure 2.2: Average Unemployment Probability over the Life Cycle
Note: Predicted probability of unemployment next year for currently working married men averaged over three years by occupation. *Source:* Own calculation based on the SOEP

Summary Statistics Table 2.1 provides weighted summary statistics of the most important variables, including wage risk and unemployment probability measures. In the first row we report the average hours worked per week, about 42 in our sample. Hourly wages average 22 Euro, with average marginal net wages of 12 Euro. Hourly wages are constructed by dividing gross monthly labor incomes by paid hours of work. All monetary variables are converted to 2010 prices using the consumer price index provided by the Federal Statistical Office. Labor earnings include wages and salaries from all employment including training, self-employment income, and bonuses, overtime, and profit-sharing.

We use paid hours because an increase in these translates directly into an increase in income.[10] The average gross wage risk in our sample is 0.192, which is similar to

[10]We discuss robustness tests using different measures of hours supplied in Section 2.6.

Table 2.1: Summary Statistics

	Unit	Mean	Std. Dev.	Min	Max	N
Labor Supply						
Weekly Hours Worked	(h)	42.03	7.3	20	80	16,038
Wages and Incomes						
Hourly Gross Wage	(Euro)	21.96	10.22	2.20	98.06	16,038
Hourly Marginal Net Wage	(Euro)	12.42	6.27	1.04	57.67	16,038
Monthly Gross Labor Income	(Euro)	3,764.47	1,997.75	319	27,000	16,038
Monthly Net Labor Income	(Euro)	2,458.91	1,197.49	150	15,000	16,038
Wage and Unemployment Probability						
Gross Wage Risk	(ln Euro)	0.192	0.196	0	3.539	11,040
Marginal Net Wage Risk	(ln Euro)	0.249	0.224	0	3.354	10,987
Unemployment Probability	(%)	1.4	2.2	0	27.4	16,038
BB-Index	(%)	2.7	4.7	-4.9	16.0	16,038
Demographics and Characteristics						
Age	(a)	43.1	7.5	25	55	16,038
Years of Education	(a)	12.8	2.7	7	18	16,038
Work Experience	(a)	21.5	8.5	0.2	41.2	16,038
Children younger than 3 years	(%)	11.6	32.0	0	100	16,038
Children between 3 and 6 years	(%)	14.5	35.2	0	100	16,038
Children between 7 and 18 years	(%)	45.2	49.8	0	100	16,038
East Germany	(%)	14.5	35.2	0	100	16,038
Type of Work						
Self-Employed	(%)	8.0	27.2	0	100	16,038
Blue Collar	(%)	32.5	46.8	0	100	16,038
White Collar	(%)	48.2	50.0	0	100	16,038
Civil Servant	(%)	11.3	31.7	0	100	16,038
One-Digit International Standard Classification of Occupations (ISCO)						
Managers	(%)	10.7	30.9	0	100	16,038
Professionals	(%)	22.0	41.4	0	100	16,038
Technicians	(%)	20.2	40.2	0	100	16,038
Clerks	(%)	7.7	26.6	0	100	16,038
Service and Sales	(%)	4.5	20.7	0	100	16,038
Craftsmen	(%)	20.9	40.7	0	100	16,038
Operatives	(%)	9.7	29.6	0	100	16,038
Unskilled	(%)	4.3	20.4	0	100	16,038

Notes: Data from SOEP (version 30). Sample of married prime-age males; 2001-2012.

the average wage risk of 0.21 reported in Parker et al. (2005). The last three variables in Table 2.1 show that our sample has 8.0% self-employed workers, 32.5% blue collar workers, 48.2% white collar workers, and 11.3% civil servants.

Figure 2.3 shows the evolution of marginal net wages over the life cycle for different occupational groups. Profiles for white collar workers, civil servants, and self-employed are very similar with increasing wages until the age of about 45. In contrast, the wages of blue collar workers are lower and exhibit less wage growth. Figure 2.4 shows the same graph for weekly hours of work. This time, the self-employed are the odd ones out working substantially more than the other groups. For all groups average hours worked are relatively constant over the life cycle.

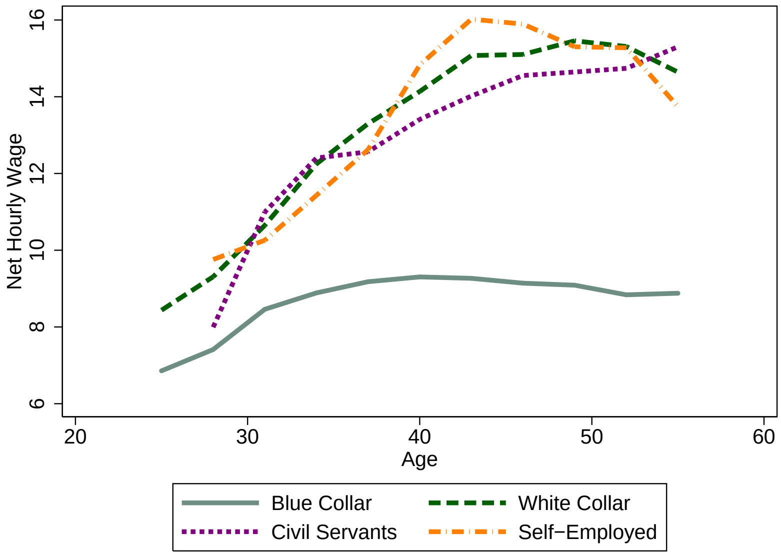


Figure 2.3: Average Marginal Hourly Net Wage over the Life Cycle

Source: Own calculation based on the SOEP

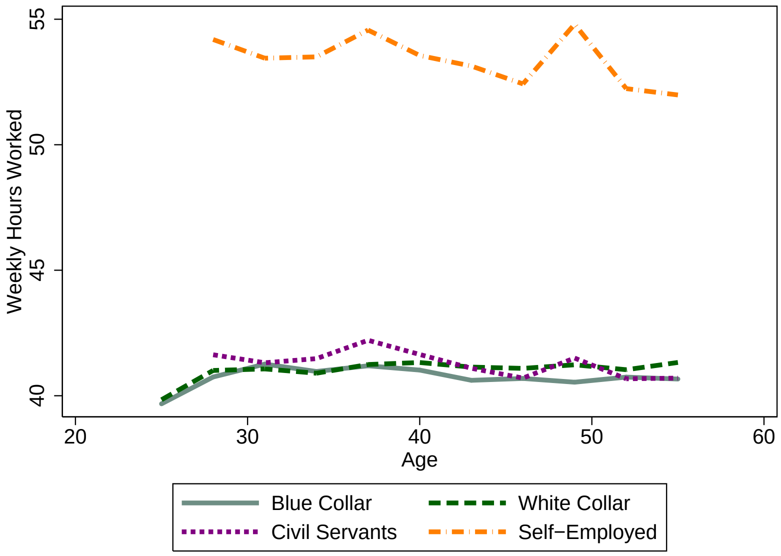


Figure 2.4: Average Weekly Hours Worked over the Life Cycle

Source: Own calculation based on the SOEP

2.3 Empirical Strategy

2.3.1 Constrained Adjustment of Labor Supply

We begin the investigation with the following labor supply equation which is similar to the specification studied in Parker et al. (2005):[11]

$$\ln h^*_{it} = \tilde{\beta}_1 \ln w_{it} + \tilde{\beta}_2 X_{it} + \tilde{\beta}_3 \sigma_{w,it} + \omega_{it}, \tag{2.3}$$

where h^*_{it} denotes desired hours of work, w_{it} denotes the marginal net hourly wage, $\sigma_{w,it}$ is a measure of wage risk, X_{it} contains additional controls, and ω_{it} is the residual.

This specification reflects the view that workers in some occupations, in particular those who are not self-employed, work more or less hours than desired. A reason for this might be contractual rigidities or fixed costs of employment like training or social insurance that make short hours of work unprofitable for firms. For manual workers, Stewart and Swaffield (1997) showed that work hours are significantly higher than the desired level (overemployment) and workers thus "off their labor supply curve". Bryan (2007) uses OLS with correction terms from a fist step random effects ordered probit model that determines the probability of being over-employed, unconstrained or under-employed (but not unemployed). He documents that 45% of manual men were constrained in their choices of hours in a given year in the UK. More recently, Bell and Blanchflower (2013b,a) proposed an index (BB-index) to measure the opposite case, i.e. that workers would like to work more hours (under-employment). They find that under-employment has been substantial in the UK labor market recently. Table 2.1 shows that in Germany as well the average person in the work force is underemployed.[12] Hours constraints might be only temporary e.g. if workers may find another job that matches their preferences better. To reflect constraints in the adjustment of hours worked, we explicitly model the dynamics of actual hours choices h_{it} and specify a partial adjustment mechanism employed by, for example, Robins and West (1980), Euwals (2005), and Baltagi et al. (2005):

$$\ln h_{it} - \ln h_{it-1} = \theta(\ln h^*_{it} - \ln h_{it-1}), \qquad 0 < \theta \leq 1. \tag{2.4}$$

θ may be interpreted as the speed of adjustment. This speed might be determined by costs to immediately adjust the labor supply to desired hours or habit persistence (see, e.g., Brown, 1952). Replace (2.4) in (2.3) to obtain the partial adjustment labor supply specification:

$$\ln h_{it} = \alpha \ln h_{it-1} + \beta_1 \ln w_{it} + \beta_2 X_{it} + \beta_3 \sigma_{w,it} + \varepsilon_{it}. \tag{2.5}$$

[11] Pistaferri (2003) specifies a different labor supply equation, which relies on subjective expectations of future earnings.

[12] Following Bell and Blanchflower (2013b) we constructed a variable that measures the probability of being under- or over-employed and included it in X_{it} along with the probability of unemployment as a robustness test in Table 2.A7 in the Appendix.

This is our empirical labor supply specification. The parameters of (2.3) can be recovered following the estimation of (2.5) with $\alpha = 1-\theta$, $\beta_1 = \theta\tilde{\beta}_1$, $\beta_2 = \theta\tilde{\beta}_2$, $\beta_3 = \theta\tilde{\beta}_3$, and $\varepsilon_{it} = \theta\omega_{it}$ (Baltagi et al., 2005).[13] The partial adjustment model nests the classic labor supply equation with $\theta = 1$ as a special case. The short-run labor supply elasticity is given by $SR_{\eta_w} = \beta_1$, and the short-run labor supply elasticity with respect to risk by $SR_{\eta_{\sigma_w}} = \beta_3$. The corresponding long-run elasticities are $LR_{\eta_w} = \beta_1/(1-\alpha)$ and $LR_{\eta_{\sigma_w}} = \beta_3/(1-\alpha)$.

2.3.2 Instrumentation and Estimation Methods

To estimate our labor supply equation, we need to account for several sources of endogeneity. First, the first difference of the lagged dependent variable is correlated with the first difference of the error term ε_{it}, which includes shocks from $t-1$. We follow Anderson and Hsiao (1981) and instrument the lagged difference in the log of hours with the level $\ln h_{it-2}$ (Anderson-Hsiao estimator). In an alternative specification, we exploit additional moment conditions as suggested by Arellano and Bond (1991) and Holtz-Eakin et al. (1988) and apply the two-step difference GMM estimator (DIFF-GMM) with Windmeijer (2005) finite-sample correction. Blundell and Bond (1998) and Arellano and Bover (1995) show that imposing additional restrictions on the initial values of the data generating process and using lagged levels and lagged differences as instruments improves the efficiency of the estimates. We also present the results from this estimator, called the system GMM (SYS-GMM).

Second, marginal net wage rates may be endogenous for two reasons: First, measurement error in hours leads to downward denominator bias in the coefficient of wage rate since the hourly wage is calculated by dividing labor income by the dependent variable hours of work (cf. Borjas, 1980; Altonji, 1986; Keane, 2011). Second, the marginal net wage depends on the choice of hours because of the nonlinear tax and transfer system. Therefore, we instrument marginal net wages with the first lag of net labor income. This variable is predetermined during the current period labor supply choices and uncorrelated with the measurement error in current period hours.

2.4 Results

2.4.1 Impact of Wage Risk on Weekly Hours of Work

Table 2.2 presents the results of the augmented labor supply equation for different estimators, where the dependent variable is the log of paid hours of work. Standard errors are robust and clustered at the individual level. Columns 1-3 show the results for the immediate adjustment specification, i.e. where the adjustment parameter α in equation (2.5) is restricted to zero. Columns 4–6 show results for the preferred dynamic

[13] Note that ε_{it} might contain an individual time-invariant effect, which is eliminated by first-differencing as in the majority of the estimators used.

specification. The first column displays results for the pooled OLS estimator. The coefficient of marginal net wage is significantly negative. The main coefficient of interest is the one associated with wage risk. The coefficient of 0.028 indicates that an increase in wage risk by one standard deviation would increase labor supply by 2.8%. The coefficient on unemployment probability is very small and not statistically significant.

Column 2 shows results for the pooled 2SLS estimator, where net wage is instrumented with lagged net labor income to overcome the denominator bias.[14] The sign of the coefficient of net wage becomes positive and the coefficient of wage risk remains significantly positive with a point estimate of 0.036. The unemployment probability becomes significant and the point estimate of 0.020 implies that an increase in unemployment probability by one standard deviation translates into 2.0% more hours worked. Column 3 displays the results obtained with the first difference estimator (FD-IV) with the equivalent instrument for net wages. The wage risk coefficient drops slightly but remains significantly positive. The coefficient of marginal net wage is not robust across estimators.

The partial adjustment specification results appear in columns 4–6 with the Anderson-Hsiao estimator displayed in column 4 and the results for the Difference and System GMM estimators displayed in columns 5 and 6, respectively.[15] The immediate adjustment specification is rejected with all three estimators because of statistically and economically significant point estimates of lagged hours of work between 0.14 and 0.2. For all three dynamic estimators, the coefficients of wage risk and unemployment probability are statistically significant. The magnitude of these effects is similar across all dynamic specifications and close to the results of the immediate adjustment specifications.

[14] We estimate it using the ivreg2 package (Baum et al., 2016).

[15] We estimate them using the xtabond2 package (Roodman, 2009).

Table 2.2: Labor Supply Regressions with Alternative Instrumentation Strategies

	OLS	2SLS	FD-IV	Anderson-Hsiao	DIFF-GMM	SYS-GMM
Lag of ln(Hours Worked)				0.155***	0.143***	0.195***
				(0.041)	(0.039)	(0.039)
ln(Net Wage) Risk	0.028***	0.036***	0.010*	0.010*	0.009*	0.024***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Unempl. Prob.	-0.005	0.020***	0.014**	0.015**	0.013*	0.015***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.004)
ln(Marginal Net Wage)	-0.031***	0.183***	-0.073*	-0.060	-0.062*	0.159***
	(0.009)	(0.019)	(0.039)	(0.041)	(0.034)	(0.019)
Controls	✓	✓	✓	✓	✓	✓
Instruments	—	labinc_{it-1}	$\Delta\text{labinc}_{it-1}$	$\ln h_{it-2}$, $\Delta\text{labinc}_{it-1}$	$\ln h_{it-2},\ldots,\ln h_{it-11}$, $\Delta\text{labinc}_{it-1}$	$\ln h_{it-2},\ldots,\ln h_{it-11}$, $\Delta\ln h_{it-2},\ldots,\Delta\ln h_{it-11}$, $\Delta\text{labinc}_{it-1}$
Observations	8,112	8,112	8,112	8,112	8,112	8,112
AR(1) in FD					0.000	0.000
AR(2) in FD					0.954	0.745
Hansen					0.694	0.368

Notes: Columns 1-3: Estimation of an immediate adjustment labor supply equation.
Columns 4-6: Estimation of equation (2.5) *using different estimators.*
We use the sample of the dynamic specifications for all estimations.
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

The coefficient on marginal net wage becomes insignificant in the Anderson-Hsiao and even significantly negative in the difference GMM specification. Blundell and Bond (1998) show that the Difference GMM estimator can be heavily downward biased. Therefore, we prefer System GMM. The wage coefficient is estimated with much higher precision using the system GMM estimator yielding statistical significance at the 1% level. This specification implies a short run labor supply elasticity of $SR_{\eta_w} = 0.16$ and a long run elasticity of $LR_{\eta_w} = 0.20$. For the difference and system GMM estimators, autocorrelation and Hansen tests appear below the estimates. The null hypothesis of no autocorrelation of second order cannot be rejected and the Hansen overidentification test does not indicate any invalidity in the instruments.

Table 2.A2 in the Appendix shows the equivalent of Table 2.2 but using gross wages instead of net wages. This facilitates comparison to the extant literature, e.g., Parker et al. (2005), that does not use microsimulation models, but relies on gross wages. The coefficient of gross wage risk is positive and significant at the 1 percent level in three of the specifications. The preferred system-GMM yields similar coefficients for all variables as the system-GMM for net wages in Table 2.2.

2.4.2 Results by Occupations

As argued by Parker et al. (2005), there should be heterogeneity across occupational groups, especially concerning self-employed. To quantify this heterogeneity, we present the results of our preferred specification across the occupational groups introduced above and the International Standard Classification of Occupations of 1988 (ISCO).

Table 2.3 provides separate results for different occupational groups using the system GMM estimator with the same instruments as in Table 2.2. As before, the risk measures are normalized by one standard deviation; however, this time not by the overall, but the sub-sample specific standard deviation. The point estimate of the wage risk coefficient is positive and statistically significant for self-employed, white collar, and blue collar workers, but not statistically different from zero for civil servants. The point estimate is largest for self-employed workers (0.036) and much smaller for white collar (0.010) and blue collar workers (0.007), suggesting the most important role of precautionary labor supply for the self-employed. Note that the result for self-employed is very similar to the one of Parker et al. (2005) where an additional standard deviation of wage risk implies an increase of annual hours of 3.66%.[16]

The coefficient on the lag of paid hours worked is not statistically significant for the self-employed and civil servants, which makes intuitively sense; these two groups are not as severely constrained in their hours choices as regular employees. Blue collar workers (0.226) are more constrained than white collar workers (0.116). This means that if underemployed blue-collar workers desire to work, say, 40 instead of 30 hours per week

[16]This number is obtained by multiplying the coefficient of risk from Model 2 with the reported standard deviation of the wage risk measure.

in Germany, they need about four years to achieve this, while white collar workers need about two years according to our estimates of the speed of adjustment parameter.

Table 2.3: System GMM Labor Supply Regressions for Occupational Groups

	Self-Employed	White Collar	Blue Collar	Civil Servant
Lag of ln(Hours Worked)	0.109	0.116**	0.226***	0.046
	(0.099)	(0.048)	(0.055)	(0.129)
ln(Net Wage) Risk	0.036***	0.010***	0.007***	-0.007
	(0.012)	(0.003)	(0.003)	(0.007)
Unempl. Prob.	-0.013	0.005	0.009**	-0.001
	(0.014)	(0.004)	(0.004)	(0.005)
ln(Marginal Net Wage)	0.123***	0.133***	0.060***	0.244***
	(0.046)	(0.020)	(0.023)	(0.095)
Controls	✓	✓	✓	✓
Observations	864	5,652	2,987	1,407
AR(1) in FD	0.000	0.000	0.000	0.001
AR(2) in FD	0.688	0.987	0.459	0.286
Hansen	0.213	0.205	0.024	0.298

Notes: Estimation of equation (2.5) *using the SYS-GMM as in column 6, Table 2.2.*
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

The coefficient of marginal net wage is positive and statistically significant for all groups. It is higher for civil servants than for other occupational groups. As in the estimation using the entire sample, we cannot reject the null hypothesis of no autocorrelation of second order. The Hansen test indicates that the instrument may be invalid only for blue collar workers.

Similarly, Table 2.A3 in the Appendix shows results for the four occupations using gross wages instead of marginal net wages. As for marginal net wages, the wage risk coefficient is significantly positive for self-employed, white collar workers and blue collar workers. The coefficients of all other variables are very similar to the main results.

Table 2.A4 in the Appendix shows system GMM estimates of the dynamic labor supply equation for eight professions grouped according to the ISCO. Each one-digit ISCO group is composed of several of the occupational classifications we used above, that is, some managers are self-employed, some not. Only clerks and operatives appear to be constrained in their hours choices. These constraints are quite persistent. The

null hypothesis that wage risk does not affect labor supply is rejected for managers, professionals, technicians, craftsmen, and operatives. An increase in the probability of unemployent corresponds to an increase of hours worked particularly for managers, craftsmen, operatives, and unskilled. The coefficient of marginal net wage is significantly positive for all but clerks, service workers and operatives. Generally, both the coefficients of net wage risk and net wage are of similar magnitude as those obtained in the estimation using the main sample.

2.5 Importance of Precautionary Labor Supply

With our estimates of the wage risk semi-elasticity we can quantify the importance of precautionary labor supply in a ceteris paribus exercise, similarly to Carroll and Samwick (1998) for precautionary savings.[17] We use the estimates from Table 2.2 to simulate the resulting distribution of hours if all individuals faced the same small wage risk. We construct this simulated counterfactual $\hat{h}_{it}$ from the predictions of the dynamic labor supply equation with minimum sample wage risk $\sigma^{\min}_{w,it}$. We use the estimates obtained with the System GMM estimator. We then compare actual hours of work h_{it} observed in the data with their simulated counterfactuals. The difference gives us a measure of the magnitude of precautionary labor supply and, for the short-run, is calculated as

$$\hat{h}_{SR,it} - h_{it} = -\beta_3(\sigma_{w,it} - \sigma^{\min}_{w,it}). \tag{2.6}$$

Figure 2.5 shows three points for each individual in the sample in 2011. The first point (p_i, h_i), denoted by a small circle, indicates the percentile rank p_i of individual i in the actually observed distribution of hours of work (vertical axis) and h_i indicates the actual hours of work (horizontal axis). The second point $(p_i, \hat{h}_{SR,i})$ keeps the percentile ranking p_i from the observed distribution and indicates the simulated *short-run* value of the hours of work $\hat{h}_{SR,i}$ when $\sigma_{w,it}$ is set to $\sigma^{\min}_{w,it}$. The third point $(p_i, \hat{h}_{LR,i})$ shows, as before, p_i from the observed distribution and indicates the simulated *long-run* value of the hours of work $\ln \hat{h}_{LR,i}$ when $\sigma_{w,it}$ is set to $\sigma^{\min}_{w,it}$.

$$\hat{h}_{LR,it} - h_{it} = -\frac{\beta_3}{1-\alpha}(\sigma_{w,it} - \sigma^{\min}_{w,it}). \tag{2.7}$$

The short-run simulated hours lie to the left of the actual hours distribution. The horizontal difference between short-run simulated points and observed points indicates the reduction in the number of hours in the short run if wage risk was reduced to the minimum level. The long-run simulated hours lie to the left of both the actual hours distribution and the short-run simulated points. The horizontal difference between

[17] Precautionary labor supply is likely even more important for singles because spousal labor supply is an additional channel of insurance against risk. However, applying our analysis to singles is difficult because only a small number of individuals in the SOEP are singles over long periods.

Figure 2.5: Reduction in Hours of Work

Notes: Small circles indicate the percentile rank of individual i in the actual observed distribution of hours of work (vertical axis) and the actual hours of work (horizontal axis) in 2011. Plus symbols maintain the percentile ranking from the observed distribution and indicate the simulated short-run *value of the hours of work when $\sigma_{w,it}$ is set to $\sigma^{min}_{w,it}$. Triangles denote the respective long-run hours of work when $\sigma_{w,it}$ is set to $\sigma^{min}_{w,it}$. Source:* Own calculation based on the SOEP

long-run simulated points and observed points indicates the reduction in the number of hours of work in the long-run if wage risk was reduced to the minimum level. The horizontal difference between simulated points in the long- and short-run indicates how much of the adjustment in hours would occur after the immediate reaction to the wage risk reduction.

Table 2.4 reports the labor supply reduction in the short run (columns 1 and 2) and the long-run (columns 3 and 4) if wage risk was reduced to the sample minimum (columns 1 and 3) or the median wage risk of civil servants (columns 2 and 4). In the pooled sample, hours of work would reduce by 2.77% in the long run if wage risk were reduced to the sample minimum. Keep in mind that this is a ceteris paribus exercise neglecting general equilibrium effects. Defining precautionary labor supply as the difference between hours worked in the status quo and in the absence of wage risk and given the average of 42 weekly paid hours of work in our sample, precautionary labor supply amounts to 1.16 hours per week on average.

If wage risk was reduced instead to the median wage risk of civil servants, labor supply would decrease on average by 1.03% in the long run. The wage risk of civil servants is

Table 2.4: Percentage Reduction for Different Occupations

	Short-Run		Long-Run	
	Perfect Foresight	Civil Servants	Perfect Foresight	Civil Servants
Self-Employed	5.01	3.65	6.17	4.49
Blue Collar	2.17	0.76	2.68	0.94
White Collar	2.03	0.62	2.51	0.77
Civil Servants	2.00	0.60	2.48	0.74
All	2.24	0.84	2.77	1.03

Notes: Simulated percentage reduction in hours of work when reducing wage risk to the sample minimum (perfect foresight) or the median risk faced by civil servants.
Source: Own calculation based on the SOEP

below average, therefore this group may be regarded as an important benchmark with particularly low uncertainty. For the self-employed, the long-run labor supply reduction would amount to 4.49%. If the wage risk of all civil servants was reduced to its median, civil servants' labor supply would decrease by 0.74%.[18]

2.6 Robustness

This section discusses the results from various robustness tests. If not indicated otherwise, the results are estimated using the preferred estimator (System GMM). The tables are delegated to the Appendix.

Table 2.A5 shows the main results for four alternative dependent variables. *Annual hours* (column 1) refers to the SOEP-imputed annual hours of work. *Weekly hours*, another variable imputed by the SOEP, is the basis for our main hours worked definition but without adjusting for paid overtime. Respondents are asked directly about *Contracted hours* and *Desired hours*. From a theoretical point of view, desired hours should not be constrained by a partial adjustment mechanism (cf. Euwals, 2005); hence, we use an immediate adjustment model for this specification. Annual hours, weekly hours and desired hours increase with increasing wage risk, while the coefficient for contracted hours is insignificant. The likely reason is that contracted hours cannot be as easily adjusted as actual hours. While still significant and economically important, the coefficient of wage risk in the desired hours specification (0.007) is smaller than in the main specification. This is not surprising because respondents might understand the question in different ways. Therefore, this measure could be affected by measurement errors, which biases the coefficient towards zero.

[18]This effect would equal zero if the distribution of wage risk were symmetric for civil servants.

Table 2.A6 shows results for eight alternative risk specifications. Column 1 shows the case with a forward looking risk measure, i.e., the standard deviation of future detrended log wages. This is similar to the approach in Feigenbaum and Li (2015). Column 2 uses a five year rolling window for the construction of the wage risk measure. Column 3 shows results obtained using the risk measure constructed using undetrended wages. This measure corresponds to the one used by Parker et al. (2005). Column 4 uses only observations with continuous employment spells, i.e., we drop observations of individuals whose employment is interrupted by periods of unemployment or changes between occupations. Columns 5 and 6 include indicators of subjective risk perceptions (Some Worries, Big Worries), column 7 includes the risk of additional household income as an additional control. This is constructed like our main risk measure, but using net household income minus net labor income of the husband instead of the husband's wage. The coefficient of this risk measure is significant and positive, so this source of risk also leads to precautionary labor supply. In column 8 we construct the wage risk measure using all past wages including those from different occupations than the current one. This increases the number of observations and the coefficient of wage risk substantially. This risk measure includes not only wage risk but also occupational risk and implies that these additional risks cause even more important precautionary behavior. The coefficients of the other regressors change only slightly. The wage risk coefficient is similar as in the main specification and remains statistically significant in all other columns.

It is possible that selection into job types could be driven by risk attitudes and the desire for hard work. If these variables are correlated with risk, this would lead to omitted variable bias. Fuchs-Schündeln and Schündeln (2005) exploit the natural experiment of the German reunification to find that risk-averse individuals self-select into low-risk occupations. Not accounting for this selection mechanism might lead to omitted variable bias. To make sure that our results are robust to such concerns, we employ two strategies, including additional controls and estimating a selection correction model. Fortunately, the SOEP elicits information on both risk preferences and the attitude towards hard work. Therefore, our first strategy is to include these additional control variables in the main model. The results are reported in Table 2.A7. In column 1 we add a variable reporting to what degree respondents agree with the assertion "Success takes hard work" on Likert scale from 1 to 7. As expected, this variable has a positive and significant impact on hours. An increase of 1 on the the Likert scale leads to an increase of 1 percent in hours of work. All other coefficients remain virtually the same. In column 2 we include a control that measures the stated willingness to take risk on a scale from 0 to 10, but do not include the preference for hard work variable. A one unit increase in this variable increases hours of work by 0.3 percent. In column 3 we include both additional control variables. Their coefficients are identical to those reported in the previous columns. The main results are very robust to this variation. In column 4 we report results, where we add a variable that captures the stated willingness to take risks in financial matters on a scale from 0 to 10 in addition to the variable capturing attitudes towards hard work. In column 5 we control for the hard-work variable and a variable capturing stated attitudes towards risks in occupational matters. An increase in the variable capturing attitudes towards occupation risk by one unit leads to an increase in hours of work by 0.4 percent, while the variable for risk attitudes in financial matters is insignificant. Again, the main results do not change.

While we explicitly model hours constraints on the occupational level in our dynamic specification, differences in hours constraints between individuals might still bias our results. Therefore we follow Bell and Blanchflower (2013b,a) and construct a region-specific indicator for under- or overemployment. The Bell-Blanchflower underemployment index (BB-index) is defined as

$$u_{BB} = \frac{U\overline{h} + \sum_k h_k^U - \sum_j h_j^O}{U\bar{h} + \sum_i h_i},$$

where U is the number of unemployed, $\overline{h}$ average hours worked by employed, h^U is preferred additional hours, which are aggregated over all workers k who desire to work more, while h^O is the preferred reduction in hours, which are aggregated over all workers j who desire to work less. $\sum_i h_i$ is the sum of actual hours of work over all workers. We use a variable for desired hours of work in the SOEP to calculate over- and underemployment. In the case that all currently employed workers are satisfied with their hours of work, the BB-index simplifies to the unemployment rate. The higher the value of this index, the more likely it is that workers are underemployed, i.e., wish to work more. Negative values indicate overemployment, i.e., people in the labor force on average wish to work less hours. As shown in Table 2.1 the value of the index is 2.7 percent on average for our sample. Column 6 of Table 2.A7 shows that an increase in the BB-index by 1%-point leads to a decrease in hours of work by 0.001 percent. The sign of the coefficient is in line with theoretical predictions. People who are more likely to be underemployed on average work slightly less, although they potentially want to work more. However, the magnitude is economically not relevant. In Column 7 we include both the BB-index and the general risk preferences variable. The BB-index becomes statistically insignificant, although the reported standard error and coefficient are identical. The reason is that the forth digit after the decimal point differs between the columns. The main results are virtually unchanged. This shows that our main results are highly robust to inclusion and exclusion of these additional control variables.

In addition to these controls, there might be selection into occupations on unobservables. We account for this possibility by estimating a Heckman (1979) selection correction model for each of the four occupations. Indicator variables for the occupation and education of both parents, and spatial planning regions are included only in the selection equation. The results are reported in Table 2.A8. The coefficient of the marginal net wage is biased downwards because we do not instrument it. Moreover, the model omits the dynamic structure of our main estimation. The focus is on the coefficients of wage risk and unemployment risk. Wage risk is positive and statistically significant at the 1 percent level and of the same order of magnitude as in Table 2.3 for the first three occupations. As before, the effect is strongest for the self-employed. The coefficient for civil servants remains insignificant. The effect of the unemployment probability remains the same except for the self-employed, where it indicates that an increase in the probability of unemployment leads to a 3.5%-decrease in hours of work. An explanation for this is that the unemployment probability for the self-employed is also a measure for the deterioration of the business and a decreasing number of orders. In the case of self-employed this is directly related to the number of hours worked. Overall, the results suggest that the main result that increases in wage risk lead to increases in hours of work is not confounded by selection bias.

Given that we do not observe many young self-employed and civil servants in our sample because these occupations are typically chosen by older individuals, we repeat the analysis by occupations including only individuals aged at least 35. The results are reported in Table 2.A9. This makes sure that the comparison is based on common support regarding the life cycle. The results are very similar to those reported in Table 2.3. This shows that the differences between occupations are not driven by differences in age.

We also show results obtained for the main sample, but including transfer recipients in Table 2.A9. This group is dropped from the main analysis because institutional insurance through the transfer system is likely to play a much larger role than precautionary behavior and even constrains precautionary behavior (Hubbard et al., 1995; Cullen and Gruber, 2000; Engen and Gruber, 2001). On the other hand, this group might be subject to more gross wage risk and therefore have stronger precautionary motives. The obtained coefficients of wage risk are virtually unchanged, when this group is included in the estimation sample.

Finally, we reestimate the main specification by occupations including interactions between year indicators and the wage risk measure (Table 2.A10). Overall, the estimates of the impact of wage risk are less precise due to less observations for a given year. Nonetheless, the coefficient is economically and statistically significant for many years except for civil servants, as in the main results. When looking at the crisis known as the Great Recession and its aftermath, i.e., 2008-2010, the effect is particularly strong for the self-employed and white collar workers. A similar pattern is not observable for blue collar workers, which does not surprise, since German manufactures made excessive use of short-time work allowance to cushion the effects of the crisis (Burda and Hunt, 2011).

2.7 Conclusion

We quantify the importance of wage risk to explain the hours of work of married men. The analysis is based on the German Socio-Economic Panel data for 2001 to 2012. We find that workers choose slightly more than an hour per week to shield against wage shocks. These effects are statistically significant for various occupations, but not for civil servants, which is in line with previous studies. We observe the largest effects of wage risk for the self-employed who have typically less coverage by institutional insurance like short term unemployment benefits. Our result for this group is quantitatively similar to previous results by Parker et al. (2005).

Precautionary labor supply is economically important. Considering a person who works 42 hours per week, precautionary labor supply amounts to about one week per year or in monetary terms, about 710 Euro per year, with a typical net wage rate of 13 Euro. If all workers faced the same risk as the median civil servant, hours worked would decrease on average by 1% in the long run. Precautionary labor supply is particularly important for the self-employed, a group that faces average wage risk substantially above the sample mean. This group works 6.2% of their hours because of the precautionary motive. Our findings suggest that unemployment probability also plays a statistically significant role, but is quantitatively less important than wage risk.

2.8 Appendix

Table 2.A1: Sample Restrictions for the Main Sample

Full sample: 416,241 person years	*Eliminated*	*Remaining*
Incomplete interviews	9,829	406,412
Drop if female	207,407	199,005
Drop if not married	55,457	143,548
Drop if younger than 26 or older than 55 in each year	86,223	57,325
Drop if in military or agriculture	2,155	55,170
Drop if transfer recipients	6,806	48,364
Drop if very low hours worked	495	47,869
Drop if unrealistic hours changes	115	47,754
Drop if unrealistic wage changes	670	47,084
Drop if without net wage or risk	36,097	10,987
After first differencing, drop if no available IVs	2,875	8,112

Source: Own calculation based on the SOEP

Table 2.A2: Comparison of Specifications, Gross Wages

	OLS	2SLS	FD-IV	FD-IV	DIFF-GMM	SYS-GMM
Lag of ln(Hours Worked)				0.173***	0.153***	0.189***
				(0.039)	(0.037)	(0.033)
ln(Gross Wage) Risk	0.044***	0.051***	0.002	0.002	0.002	0.036***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)
Unempl. Prob.	-0.003	0.013***	0.005	0.005	0.005	0.008**
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.003)
ln(Marginal Gross Wage)	-0.081***	0.130***	0.000	0.012	-0.003	0.112***
	(0.010)	(0.015)	(0.023)	(0.026)	(0.025)	(0.016)
Controls	✓	✓	✓	✓	✓	✓
Instruments	—	labinc_{it-1}	$\Delta\text{labinc}_{it-1}$	$\ln h_{it-2}$, $\Delta\text{labinc}_{it-1}$	$\ln h_{it-2}, \ldots, \ln h_{it-11}$, $\Delta\text{labinc}_{it-1}$	$\ln h_{it-2}, \ldots, \ln h_{it-11}$, $\Delta \ln h_{it-2}, \ldots, \Delta \ln h_{it-11}$, $\Delta\text{labinc}_{it-1}$
Observations	11,276	11,276	11,276	11,276	11,276	11,276
AR(1) in FD					0.000	0.000
AR(2) in FD					0.193	0.100
Hansen					0.708	0.238

Notes: Columns 1-3: Estimation of an immediate adjustment labor supply equation.
Columns 4-6: Estimation of equation (2.5) *using different estimators.*
We use the sample of the dynamic specifications for all estimations.
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

Table 2.A3: Occupational Groups, System GMM, Gross wages

	Self-Employed	White Collar	Blue Collar	Civil Servant
Lag of ln(Hours Worked)	0.132**	0.161***	0.197***	0.015
	(0.064)	(0.048)	(0.040)	(0.127)
ln(Gross Wage) Risk	0.019**	0.013***	0.010***	-0.005
	(0.009)	(0.003)	(0.003)	(0.007)
Unempl. Prob.	-0.019	0.007*	0.011***	0.002
	(0.014)	(0.004)	(0.003)	(0.005)
ln(Marginal Gross Wage)	0.082**	0.115***	0.055***	0.226**
	(0.034)	(0.018)	(0.021)	(0.093)
Controls	✓	✓	✓	✓
Observations	1,328	6,755	5,414	1,512
AR(1) in FD	0.000	0.000	0.000	0.001
AR(2) in FD	0.244	0.159	0.953	0.302
Hansen	0.916	0.146	0.052	0.582

Notes: Estimation of equation (2.5) *using the SYS-GMM as in column 6, Table 2.2.*
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

Table 2.A4: System GMM Labor Supply Regressions for ISCO Groups

	Managers	Professionals	Technicians	Clerks	Service and Sales	Craftsmen	Operatives	Unskilled
Lag of ln(Hours Worked)	0.135	0.111	-0.054	0.429***	0.016	0.046	0.323***	0.327
	(0.093)	(0.076)	(0.105)	(0.142)	(0.125)	(0.068)	(0.090)	(0.262)
ln(Net Wage) Risk	0.025***	0.027***	0.021***	0.005	0.012	0.022***	0.034***	0.016
	(0.008)	(0.007)	(0.008)	(0.003)	(0.010)	(0.006)	(0.013)	(0.019)
Unempl. Prob.	0.019**	0.007	0.007	-0.008*	0.000	0.019***	0.012*	0.015*
	(0.009)	(0.006)	(0.007)	(0.004)	(0.010)	(0.007)	(0.006)	(0.008)
ln(Marginal Net Wage)	0.187***	0.299***	0.174***	0.043	0.057	0.191***	0.092	0.162*
	(0.059)	(0.051)	(0.041)	(0.027)	(0.059)	(0.044)	(0.066)	(0.085)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1314	3007	2197	797	398	1985	880	332
AR(1) in FD	0.000	0.000	0.000	0.000	0.084	0.000	0.001	0.017
AR(2) in FD	0.496	0.259	0.712	0.720	0.451	0.351	0.107	0.765
Hansen	0.703	0.042	0.366	0.466	0.526	0.303	0.062	0.393

Notes: Estimation of equation (2.5) *using the SYS-GMM as in column 6, Table 2.2.*
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

Table 2.A5: Alternative Hours Definitions

	Annual Hours	Weekly Hours	Contracted Hours	Desired Hours
Lag of ln(Hours Worked)	0.114	0.110	0.205**	
	(0.075)	(0.070)	(0.081)	
ln(Net Wage) Risk	0.024***	0.020***	-0.001	0.007**
	(0.004)	(0.004)	(0.001)	(0.003)
Unempl. Prob.	0.012**	0.018***	0.001	0.015***
	(0.006)	(0.005)	(0.003)	(0.004)
ln(Marginal Net Wage)	0.218***	0.215***	0.032***	0.144***
	(0.024)	(0.023)	(0.008)	(0.018)
Controls	✓	✓	✓	✓
Observations	11,034	10,845	8,739	10,768
AR(1) in FD	0.000	0.000	0.000	0.000
AR(2) in FD	0.475	0.139	0.726	0.929
Hansen	0.514	0.547	0.810	

Notes: Estimation of equation (2.5) *using the SYS-GMM as in column 6, Table 2.2.*
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

Table 2.A6: Alternative Risk Definitions

	Forward	Five years	Undetrended	Cont. Spells	Subj. Risk	Subj. & Wage	Household Risk	With Occ. Changes
Lag of ln(Hours Worked)	0.223***	0.200***	0.192***	0.226***	0.187***	0.195***	0.171***	0.157***
	(0.049)	(0.039)	(0.039)	(0.044)	(0.041)	(0.041)	(0.042)	(0.034)
ln(Net Wage) Risk	0.020***	0.019***	0.023***	0.013***		0.021***	0.013**	0.088***
	(0.003)	(0.003)	(0.004)	(0.003)		(0.004)	(0.005)	(0.013)
Unempl. Prob	0.010***	0.009***	0.009***	0.008***	0.012***	0.011***	0.007**	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.002)
ln(Marginal Net Wage)	0.154***	0.156***	0.160***	0.158***	0.158***	0.164***	0.107***	0.164***
	(0.022)	(0.019)	(0.019)	(0.020)	(0.021)	(0.020)	(0.030)	(0.015)
Some Worries					0.016	0.055		
					(0.042)	(0.043)		
Big Worries					-0.086	-0.044		
					(0.076)	(0.075)		
ln(Net Household Inc.) Risk							0.061**	
							(0.031)	
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	5,675	8,089	8,112	6,614	8,101	8,101	8,014	15,544
AR(1) in FD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) in FD	0.577	0.835	0.800	0.776	0.425	0.318	0.870	0.498
Hansen	0.233	0.111	0.614	0.014	0.408	0.614	0.521	0.366

Notes: Estimation of equation (2.5) *using the SYS-GMM as in column 6, Table 2.2.*
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

Table 2.A7: Additional Control Variables

	I	II	III	IV	V	VI	VII
Lag of ln(Hours Worked)	0.196***	0.198***	0.199***	0.200***	0.203***	0.195***	0.198***
	(0.040)	(0.039)	(0.040)	(0.041)	(0.041)	(0.039)	(0.040)
ln(Net Wage) Risk	0.021***	0.021***	0.020***	0.021***	0.020***	0.021***	0.020***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Unempl. Prob.	0.009***	0.009***	0.010***	0.009***	0.009***	0.010***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
ln(Marginal Net Wage)	0.154***	0.156***	0.149***	0.151***	0.147***	0.158***	0.151***
	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)
Success Takes Hard Work	0.010***		0.010***	0.010***	0.010***		0.010***
	(0.002)		(0.002)	(0.002)	(0.002)		(0.002)
General Risk Preference		0.003**	0.003***				0.003**
		(0.001)	(0.001)				(0.001)
Financial Risk Preference				-0.001			
				(0.001)			
Occupational Risk Preference					0.004***		
					(0.001)		
BB-Index						-0.001*	-0.001
						(0.001)	(0.001)
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	7,862	8,109	7,859	7,686	7,653	8,112	7,859
AR(1) in FD	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) in FD	0.884	0.604	0.709	0.770	0.807	0.764	0.725
Hansen	0.280	0.312	0.149	0.324	0.204	0.297	0.252

Notes: Estimation of equation (2.5) *using the SYS-GMM as in column 6, Table 2.2.*
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

Table 2.A8: Two-step Heckman Selection Correction Model

	Self-Employed	White Collar	Blue Collar	Civil Servant
ln(Net Wage) Risk	0.033***	0.016***	0.006*	-0.010
	(0.011)	(0.003)	(0.004)	(0.006)
Unempl. Prob.	-0.035***	0.006	0.009**	0.001
	(0.010)	(0.004)	(0.004)	(0.008)
ln(Marginal Net Wage)	-0.100***	-0.024***	-0.050***	-0.296***
	(0.017)	(0.008)	(0.010)	(0.022)
Inverse Mills Ratio	-0.004	-0.003	0.012	0.026*
	(0.024)	(0.012)	(0.010)	(0.015)
Observations	4,758	4,758	4,758	4,758

Notes: Estimation of the immediate adjustment labor supply equation using the two-step Heckman selection model. Exclusion restrictions are: Indicator variables for the occupation and education of both parents, and spatial planning regions. Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Own calculation based on the SOEP

Table 2.A9: Variations of the Sample I

	All, age> 34	SE, age> 34	WC, age> 34	BC, age> 34	CS, age> 34	Incl. TR
Lag of ln(Hours Worked)	0.200***	0.105	0.129***	0.210***	0.018	0.201***
	(0.040)	(0.102)	(0.050)	(0.065)	(0.137)	(0.038)
ln(Net Wage) Risk	0.023***	0.036***	0.010***	0.009***	-0.004	0.023***
	(0.004)	(0.012)	(0.003)	(0.003)	(0.008)	(0.004)
Unempl. Prob.	0.010***	-0.015	0.005	0.008**	-0.001	0.015***
	(0.003)	(0.015)	(0.005)	(0.004)	(0.005)	(0.004)
ln(Marginal Net Wage)	0.162***	0.125***	0.135***	0.069***	0.257***	0.156***
	(0.019)	(0.048)	(0.021)	(0.025)	(0.096)	(0.018)
Controls	✓	✓	✓	✓	✓	✓
Observations	7,547	830	5,216	2,539	1,337	8,660
AR(1) in FD	0.000	0.000	0.000	0.000	0.001	0.000
AR(2) in FD	0.627	0.667	0.890	0.434	0.244	0.854
Hansen	0.255	0.204	0.345	0.057	0.299	0.248

Notes: Estimation of equation (2.5) *using the SYS-GMM as in column 6, Table 2.2. SE: Self-employed; WC: White collar, BC: Blue collar, CS: Civil servants; TR: Transfer recipients. Robust standard errors clustered at the individual level in parentheses.*

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Own calculation based on the SOEP

Table 2.A10: Year-Specific Effects

	Self-Employed	White Collar	Blue Collar	Civil Servant
Lag of ln(Hours Worked)	0.103	0.117**	0.229***	0.058
	(0.097)	(0.048)	(0.056)	(0.123)
ln(Net Wage) Risk × year				
2003	0.041**	0.007	0.000	0.012
	(0.018)	(0.009)	(0.009)	(0.026)
2004	0.011	0.011	0.013	-0.046
	(0.022)	(0.011)	(0.010)	(0.039)
2005	0.032	0.041***	0.047***	-0.020
	(0.026)	(0.013)	(0.015)	(0.037)
2006	0.044**	0.026	0.004	-0.013
	(0.020)	(0.016)	(0.011)	(0.032)
2007	0.063***	0.020*	0.035***	-0.032
	(0.022)	(0.011)	(0.011)	(0.038)
2008	0.060*	0.026**	0.027**	-0.013
	(0.031)	(0.012)	(0.013)	(0.017)
2009	0.076**	0.031**	0.017	-0.001
	(0.030)	(0.012)	(0.014)	(0.022)
2010	0.120***	0.043***	0.020	-0.022
	(0.028)	(0.016)	(0.020)	(0.048)
2011	0.040	0.040***	0.025	0.030
	(0.034)	(0.012)	(0.017)	(0.040)
Unempl. Prob.	-0.007	0.003	0.002**	-0.000
	(0.006)	(0.002)	(0.001)	(0.005)
ln(Marginal Net Wage)	0.119***	0.133***	0.061***	0.243***
	(0.041)	(0.020)	(0.023)	(0.092)
Controls	✓	✓	✓	✓
Observations	864	5,652	2,987	1,407
AR(1) in FD	0.000	0.000	0.000	0.001
AR(2) in FD	0.666	0.954	0.390	0.331
Hansen	0.229	0.227	0.027	0.312

Notes: Estimation of equation (2.5) *using the SYS-GMM as in column 6, Table 2.2.*
Robust standard errors clustered at the individual level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own calculation based on the SOEP

3 The Effects of Germany's New Minimum Wage on Employment and Welfare Dependency

3.1 Motivation

On the 1st of January 2015, Germany introduced a federal, statutory minimum wage of € 8.50 per hour. There were some fears about negative employment effects of this new policy. However, proponents of the minimum wage, such as the German Social Democrats (SPD, 2013), argued that apprehensions regarding job losses were ungrounded. In fact, a minimum wage would be necessary to supplement earnings in the growing low-wage sector and to cushion the large-scale labor market reforms of the early 2000s, the so-called Hartz reforms. A key target group of the minimum wage are households that receive supplementary welfare benefits (unemployment benefit II, *UBII*) while working, the so-called *Aufstocker*. The proponents of the minimum wage argued that increasing the labor earnings via the minimum wage would reduce welfare spending and help households to end their welfare dependency.

At first glance, the labor market outcomes since the reform seem rather comforting. The Federal Ministry of Labor and Social Affairs (BMAS, 2015) reports that after the introduction of the minimum wage unemployment actually fell and that regular employment is at an all time high, although there was some loss of marginal employment.[1] The ministry estimates that 3.7 million employees profited from higher labor earnings. The number of the *Aufstocker* dropped by 50,000 and related welfare expenditures on UBII were expected to decrease by € 0.9bn. The ministry concludes that the "minimum wage works" (BMAS, 2015).

Without a doubt, Germany introduced the minimum wage in times of a healthy labor market and solid economic circumstances. However, it is impossible to determine the *causal* effects of the minimum wage just based on the aggregate employment and social security statistics. Accordingly, this study evaluates the effects of the statutory minimum wage on employment and welfare dependency using county-level administrative data. In order to construct a credible counter-factual, i.e., the development of the labor market under the absence of the minimum wage, identification is based on variations in the relative depth of the intervention on regional level.

[1] Regular employment (sozialversicherungspflichtige Beschäftigung) refers to jobs subject to social security contributions, i.e. with an average monthly income of more than € 450. Jobs below or just paying € 450 are exempt from these contributions and constitute marginal employment (geringfügige Beschäftigung.).

The analysis for employment effects will focus on the impact on both, regular and marginal employment, where the latter refers to so-called *Minijobs*, which are fully social-security exempt (up to € 450 per month). Regarding welfare dependency, there will be a distinction between those who are capable of working, but do not work while receiving UBII payments and those who are working, but top-up their labor earnings with UBII (*Aufstocker*). Also the composition of the *Aufstocker* will be considered: I separate the analysis according to labor income brackets[2] and employment status (self-employed or dependently employed). Furthermore I will evaluate the effects of the minimum wage on the regional wage distribution.

The results suggest that the minimum wage had a considerable negative effect on marginal employment. A back-off-the-envelope calculation indicates that in 2015 150,000-200,000 marginal jobs have been lost, due to the minimum wage. Concerning regular employment, the results indicate a rather small (short-run) negative effect of the minimum wage. Concerning welfare dependency, the minimum wage reduced the number of working welfare recipients, with some indication that about one half of them left welfare receipt due to the minimum wage. The effect on welfare reduction in absolute terms is rather small.

The remainder of this paper is structured as follows: Section 3.2 introduces the institutional background and briefly summarizes the previous studies on employment and welfare effects. Section 3.3 describes the identification strategy, the different outcome variables, and presents the data. Section 3.4 discusses the results. Section 3.5 concludes.

3.2 Minimum Wages in Germany

Background and Institutional Factors Germany introduced a statutory, federal minimum wage as a response to a large variety of economic and political trends. Traditionally, the majority of wages in West Germany have been determined by collective wage bargaining. Hence, trade unions and employers associations alike opposed minimum wages as interference with their autonomous wage- setting. However, with declining coverage of collective agreements starting in the 1990s, and increasing dispersion of gross labor income and equivalized net income[3], trade unions began to favor broadly applicable, legal wage floors.

Germany started to introduce sectoral minimum wages from 1997 onwards, using the legislation on the posting of workers (*Arbeitnehmer-Entsendegesetz*). The majority of

[2]There are three different categories, based on the rules regarding social security contributions: First, fully social-security exempt (*Minijobs*, up to € 450 p.m); second, the phase-in zone for social security contributions (*Midijob*, between € 450 and € 850 p.m.) and third, regular employment (more than € 850 p.m., fully subject to payroll taxes)

[3]See among others Gernandt and Pfeiffer (2007); Antonczyk et al. (2010, 2011); Biewen and Juhasz (2012); Card et al. (2013). This upward trend in inequality was most pronounced from German unification in the early 1990s till the mid 2000s. Income inequality stabilized in more recent years (c.f. Grabka and Goebel, 2014; Möller, 2016).

the covered branches are in the crafts and construction sector and were introduced in order to shield domestic firms against foreign competitors who are also subject to the minimum wages if they operate in Germany.

Minimum wages became also politically more desirable as a means to supplement and cushion the large-scale labor market reforms of the early 2000s, the so-called Hartz reforms.[4] The reforms attempted to lower reservation wages and introduced new forms of marginal employment, exempt from social security contributions (so-called *Minijobs*). It is often argued that the reform package stimulated the expansion of the low-wage sector.[5] The Hartz reforms also encouraged the use of welfare payments as an implicit combination wage. Households with low labor earnings can supplement their income with the *unemployment benefit II* (UBII) to reach subsistence level. People making use of this provision are commonly referred to as *Aufstocker*, literally "those who top-up". Some researchers argue that these reforms are (at least partially) responsible for the success of the German labor market and the German economy in the last decade (Carlin and Soskice, 2009; Boysen-Hogrefe and Groll, 2010; Gartner and Klinger, 2010; Burda and Hunt, 2011). Yet, there is disagreement of its relative importance compared to other factors, such as wage moderation (Akyol et al., 2013; Dustmann et al., 2014).

Proponents of the minimum wage, such as Rürup and Heilmann (2012) or the Social Democrats (SPD, 2013) argue that the introduction of a general minimum wage is necessary to counter the negative effects of the new low-pay sector and will increase the efficiency of the combination wage scheme, since the *Aufstocker* will obtain a larger share of their income from work and not from transfers. Assuming the absence of detrimental employment effects,[6] a modestly set minimum wage would be beneficial for the public budget, due to the reduction in supplementary welfare payments for the *Aufstocker* and increases in payroll and income taxes. The Social Democrats, the driving force in introducing the minimum wage, argued that the minimum wage would generate a fiscal surplus of € 7bn per year (SPD, 2013, pg. 69).[7]

After the federal election in Autumn 2013 and the change of government, the political climate shifted in favor of a statutory minimum wage.[8] In January 2015 a minimum wage

[4]For an overview, see Ochel (2005).

[5]The low-wage sector is usually defined as wages below 2/3 of the median wage. The share of jobs considered to be in the low-wage sector increased between the 1997 and 2007 from ca 16 to 22% and remained constant afterwards. However, most of that increase took place already *before* the Hartz reforms (Schäfer and Schmidt, 2012; Brenke, 2012).

[6]Standard economic theory about minimum wages predicts unambiguous negative effects and involuntary unemployment. However, as for instance argued in Manning (2003); Garloff (2010), a minimum wage does not necessarily need to reduce employment because of some monopsony power of the employers, for instance due to search frictions (Card and Krueger, 2015).

[7]This claim is based on Ehrentraut et al. (2011) who calculate the fiscal effects of a minimum wage, assuming the absence of negative emplyoment effects.

[8]Chancellor Merkel (Christian Democrats) remained in office, but the Social Democrats replaced the Liberals as the coalition partner.

of € 8.50 per hour was intoduced. The minimum wage passed into law in the summer of 2014[9], only few exemptions apply.[10] Sectoral minimum wages remain unaffected.

The following literature review will present the evidence on the economic effects of minimum wages in Germany.[11] I will focus on the empirical insights from ex-ante and ex-post studies concerning employment effects and welfare dependency.[12]

Employment Effects There are various ex-ante studies on the effects of a statutory minimum wage.[13] These simulations generally point to rather substantial employment losses, but the variation of potential effects is large (for a comparisson, see Müller, 2009). Concerning the sectoral minimum wages, there is also a large body of ex-post studies.[14] Not surprisingly, these studies indicate that the bite, i.e. the share of directly affected workers matters a lot for the effects of minimum wages. In West Germany, the sectoral minimum wages were usually comparatively low. Hence, employment effects have been very small or not statistically significant. Moreover, spillover effects might be an issues, i.e. minimum wages can affect the wage distribution above the level of the wage floor.

There are also several studies that analyze descriptively or with ex-post evaluations the employment effects of minimum wage. Based on aggregate employment statistics Groll (2015) shows descriptively that there is a striking reduction of marginal employment in the beginning of 2015. Vom Berge et al. (2016) study data on individual transitions and show that there is no large flow into unemployment, but indeed transitions from marginal to regular employment. Based on the IAB Establishment Panel, Bossler (2016) and Bossler and Gerner (2016) exploit the self-declared affectedness of establishments by the minimum wage and detect a small negative effect on employers' employment expectations (before the reform) and estimate a reduction in employment growth of 60,000 jobs due to the minimum wage. Garloff (2016) uses regional data of the Federal Employment Agency and does not find any evidence for a decrease of employment growth or an increase in unemployment growth. His results suggest that there was a transformation from marginal into regular employment. Knabe et al. (2016) argue that a simple East/West comparison can already detect effects of the minimum wages, since East Germany is much more exposed to the minimum wage than most of the West. They consider aggregate employment statistics and argue that the minimum wage had an effect on labor market dynamics. In 2015, overall employment grew by 300,000 jobs; however,

---

[9] The German *Bundestag* voted for the minimum wage on 3 July 2014, the second chamber *Bundesrat* confirmed the law on 11 July 2014. The law became effective on 16 August 2014.

[10] Apprentices, compulsory internships, long-term unemployed for the first six month.

[11] For a general overview about the empirics of minimum wages see Brown (1999); Neumark and Wascher (2008), for a European focus Dolado et al. (1996).

[12] For a discussion about the theoretical arguments concerning minimum wages in the German context see Fitzenberger (2009).

[13] Among others Bachmann et al. (2008); Bauer et al. (2009); Knabe and Schöb (2009); Müller and Steiner (2011); Knabe et al. (2014); Arni et al. (2014); Henzel and Engelhardt (2014).

[14] Among others Möller and König (2008); Müller (2010); Boockmann et al. (2013); Frings (2013); Aretz et al. (2012, 2013); Gregory (2014); Rattenhuber (2014).

only by 0.2% in the East compared tp 0.9% in the West. Additionally, they conclude that even if there was some transformation of marginal into regular employment, not all marginal jobs have been upgraded.

Effects on Welfare Dependency Proponents of the minimum wage argue that it could lift poor households, such as the *Aufstocker* out of welfare. Some of the ex-ante studies (e.g. Bachmann et al., 2008; Müller and Steiner, 2009; Bauer et al., 2009; Knabe and Schöb, 2009; Knabe et al., 2014; Arni et al., 2014) do not only analyze the employment, but also the fiscal effects of a minimum wage. These studies commonly find that the effects on the *Aufstocker* and welfare dependency are very small or negligible and usually offset by negative employment effects, which are especially severe for this group. The vast majority of the *Aufstocker* is found to remain in welfare receipt, either because of the household context (single parents, many children), hours constraints (disabilities, child care), or both of it. Additionally, Müller and Steiner (2009) estimate that only 25% of the gross income increase due to a minimum wage sticks with households and further argue that minimum wages are not well-targeted for poverty reduction, since they also affect secondary earners in households above the poverty line. In a more recent account, Bruckmeier and Wiemers (2014) argue along similar lines. However, even if the disposable income of the households would not change much due to high transfer withdrawal rates, reduced welfare stigma could improve well-being considerably (c.f. Hetschko et al., 2016).

Bruckmeier and Wiemers (2016) describe the developments for the *Aufstocker* after the introduction of the minimum wage and report that their numbers decreased from December 2014 to January 2015 by 2% (-23,000). This reductions is larger than at previous turns of the years, mostly driven by former marginally employed *Aufstocker*. They provide evidence that in the following months more *Aufstocker* than before managed to leave welfare dependency.

Summing Up So far, no evidence for substantial employment losses due to the new minimum wage exists. There is a loss of marginal employment that seems to be (partially) offset by transformations into regular employment; however, there might be some reductions in employment dynamics. Since the introduction of the minimum wage there was a reduction in the number of the *Aufstocker*, especially those with a marginal employment. There is tentative evidence that more *Aufstocker* than before left welfare receipt due to higher labor earnings.

3.3 Method

3.3.1 Identification Strategy

In order to identify the *causal* effects of the minimum wage on employment and welfare dependency, this study will exploit regional differences (county level, N=402) in the *bite* of the minimum wage as the source of exogenous variation. I will define the bite of the minimum wage as the county-specific share of workers with wages less than € 8.50 per hour before the introduction of the minimum wage.[15] Unlike in the United States, where states can set their own wage floor above the federal minimum wage, in Germany this kind of regional variation does not exist. However, a uniformly set minimum wage of € 8.50 per hour has rather different repercussions across the country. In prosperous economic regions, such as Munich or Frankfurt, the vast majority of workers already receives a wage rate well above € 8.50. On the other hand, € 8.50 is a relatively high wage rate in most parts of East Germany and also in rural, economically struggling regions in the West. In that sense, even though the minimum wage is nominally the same in all regions, the effective strength of the treatment differs considerably. Card (1992) uses this type of variation in order to study the effects of minimum wages on teenage employment, Garloff (2016) uses the same data set as this study, but considers employment outcomes only.

The estimation will make use of observations before and after the policy change, hence a difference-in-differences (DID) estimator is appropriate. The effect of the minimum wage is recovered as the difference between strongly and only mildly "treated" regions before and after the turn of the year 2014/15. Instead of a binary treatment, the regional bite of the minimum wage functions as an indicator for the strength of the treatment. Given that we have a monthly panel running from January 2012 to December 2015 (T=48) of the 402 counties, the difference-in-differences estimator in log-levels can be implemented as

$$\log(y_{it}) = bite_i \cdot D_t^{\text{MW}} \cdot \beta_L + \sum_t D_t^{\text{month}} \cdot \gamma_t + \theta_i + \varepsilon_{it} \tag{3.1}$$

where y_{it} is one of the outcomes of interest[16], measured in period t in county i, the $bite_i$ is the county-specific (but not time-varying) depth of the intervention, interacted with an indicator variable D_t^{MW} which is equal to one for all periods after the introduction of the minimum wage, i.e. the entire year 2015. Furthermore, Equation 3.1 features time fixed effects γ_t and county fixed effects θ_i. Standard errors are clustered on county level as advocated by Bertrand et al. (2004). Thus, β_L is the parameter of interest. In order to ease interpretation of β_L, the bite will be normalized by one standard deviation

[15] The details will be presented in the following Section 3.3.2.

[16] Regular and marginal employment, working and non-working recipients of UBII. Information on wages is only available on an annual basis. The specification for the impact on the wage structure will be presented in Section 3.4.

and divided by 100. Hence, the estimate of β_L from Equation 3.1 corresponds to the *percentage change* of the outcome variable, due to one additional standard deviation of the county-specific bite.

Alternatively, one could specify a model of growth rates instead of levels, where only the left-hand side of Equation 3.1 is modified, yielding

$$\triangle_{12}\log(y_{it}) = \log(y_{it}) - \log(y_{i,t-12}) = bite_i \cdot D_t^{\mathrm{MW}} \cdot \beta + \sum_t D_t^{\mathrm{month}} \cdot \gamma_t + \theta_i + \varepsilon_{it} \qquad (3.2)$$

Note that Equation 3.1 and 3.2 are two distinct models and the estimated coefficients of interest, β and β_L have entirely different interpretations. For the specification in growth rates (Equation 3.2), β provides an estimate for the *percentage point change* of the annual month-specific growth rate of the outcome variable, due to one additional standard deviation of the county-specific bite.

Since the difference-in-differences identification is scale dependent, at best only one of the two specifications is valid. I will argue graphically in Section 3.3.4 that the developments of the outcome variables follow a process modeled more appropriately by growth rates as in Equation 3.2 than in log-levels, as specified in Equation 3.1. The advantage of the specification in growth rates is that stochastic patterns of seasonality are accounted for by having the month t-specific seasonally adjusted annual growth rate of y_{it} on the left-hand side of the equation.

As in any difference-in-differences estimator, identification rests on the validity of the comparison between the treatment and control groups. When the treatment is continuous, all regions are affected by the policy change, but the intensity of the treatment differs. Regions with a high bite were affected more than those with a low bite. The standard DID framework with an unambiguously defined binary treatment requires that the comparison group is unaffected by the treatment. When the treatment is continuous, the requirement is that regions with a lower bite are *proportionally* less affected by the policy. Given this setting, identification rests on two canonical difference-in-differences assumptions, the common trend and the stable unit treatment value assumption.

The common trend assumption (CTA) states that in the absence of the policy, the development of the outcomes of interest should have been *parallel* in regions which are highly affected and regions which are only mildly affected by the new policy. For the specification in log-levels, this implies that percentage changes of the outcome over time should be unrelated to bite if there would be no minimum wage. For the alternative specification from Equation 3.2, the CTA implies that changes in growth rates would not be systematically related to the bite, if there was no minimum wage.

It is conceivable that regions with a high bite exhibit different trend (or growth rate) behavior for the outcomes of interest than those with a low bite. This issue will be addressed with an alternative specification, in which $bite_i \cdot t \cdot \delta$ is introduced as an additional regressor. The term t is a linear time trend and δ a bite-specific linear trend differential. If the differences in the pre-treatment trend behavior can be adequately captured with a deterministic linear trend proportional to the bite, β will provide a valid estimator for the ATT. The causal effects are then recovered as deviations from a pre-treatment trend differential due to the larger impact of the minimum wage. Section 3.3.4 compares the trend behavior of the treated and the control groups graphically, in order to decide whether this alternative specification is appropriate or not.

The stable unit treatment value assumption (SUTVA) stipulates that more or less treated observations do not interfere, so that for example a stronger treatment in one region does not affect the outcome in another region. It is very likely that the SUTVA does not hold in the medium to long run, for instance through firms investment decisions. A firm with two factories, one in a high wage and one in a low wage region might shift its investments to the high wage region, due to the change in relative wages between the two regions. However, these investments and relocation decisions take time and I assume that they are negligible in the narrow time frame of this study. Also labor mobility and migration could invalidate the SUTVA; however, as for the decisions of the firms, I assume that the effects of the minimum wage on labor mobility and migration are negligible in the first year of the new policy.

Having multiple time periods does not only allow to argue graphically for the validity of the CTA, but also to study how the effect evolves over time. This is done by specifying multiple pre- and post-treatment periods, as for instance in Autor (2003). One obtains a more flexible specification for the annual growth rate in which D_t^{MW} is replaced by a series of indicators:

$$\triangle_{12}\log(y_{it}) = bite_i \cdot \sum_{\tau} D_{\tau}^{\text{month}} \cdot \beta_{\tau} + \sum_{t} D_t^{\text{month}} \cdot \gamma_t + \theta_i + \varepsilon_{it} \tag{3.3}$$

where τ is an indicator for the periods of interest. Of course, also the (log-)levels specification from Equation 3.1 could be respecified in the fashion of Equation 3.3. In the analysis, τ will take seven post-treatment values for lagged adjustment[17] and six pre-treatment values (July - December 2014) for anticipation effects after the minimum wage law was passed. Also Equation 3.3 can be supplemented with $bite_i \cdot t \cdot \delta$, if the alternative CTA (deviations from trend differential) is more appropriate.

Figure 3.1 provides evidence that shortly already after the passing of the minimum wage bill, some anticipation effects arose. The figure plots the weekly relative search intensity for the Google search query *"Mindestlohn"* originating from Germany for the years 2013 to 2015 (Google Trends, 2017). The search intensity index is set to 100 for the week with the highest number of queries relative to all search queries. The graph shows a first large spike in the first week of July 2014, when the minimum wage was voted for (light blue vertical line) and a second large spike at the turn of the year when the minimum wage took effect (dark blue vertical line) Hence, it is likely that the majority of

[17] Six month adjustment after introduction (January 2015, February 2015, ..., June 2015) and a joint medium-run effect, i.e. the time between July and December 2015.

people was already well-informed about the new minimum wage, half a year before the official start and that anticipation effects are possible.

Note: Measured between January 2013 and December 2015, Own Graph, Data: Google Trends (2017)

Figure 3.1: Google Trends for Search Query *"Mindestlohn"* over Time

3.3.2 Measurement of the Bite and Treatments

The bite is calculated based on the wage statistics of the Federal Employment Agency, an administrative dataset, aggregated on county level, containing the distribution of gross labor earnings. The statistic is based on social security notifications and refers to regular, full-time employment (Statistik der Bundesagentur für Arbeit, 2016c). As in Garloff (2016), all earnings up to € 1400 per month are supposed to be subject to the minimum wage.[18] The bite is calculated as the county specific share of these earnings in relation to all recorded full-time employees. The wage statistics are available only once per year, namely in December. In order to avoid anticipation effects, the bite is calculated based on data from December 2013, hence one year before the introduction of the new minimum wage.

This measure of the bite has some important caveats. First, there is no information on hours worked and the calculation is only based on the social security records of full-time employees. If this lack of information results in a classic symmetric measurement error, the estimated coefficients and the ATT $\hat{\beta}$ would suffer from a downward attenuation bias. The problem would be exacerbated if there are systematic differences across counties, related to the size of the bite, for instance due to differences in the prevalence of part-time employment. Second, recipients of minimum wages are frequently working in marginal employment, i.e. jobs not subject to social security contributions and hence not covered in the wage statistics and in the calculation of the bite. This feature might actually be desirable for the effects on regular employment, but could potentially be misleading for marginal employment and for *Aufstocker* and welfare dependency. Given that there are neither alternative wage statistics, nor credible instrumental variables available, I will abstract from these issues and

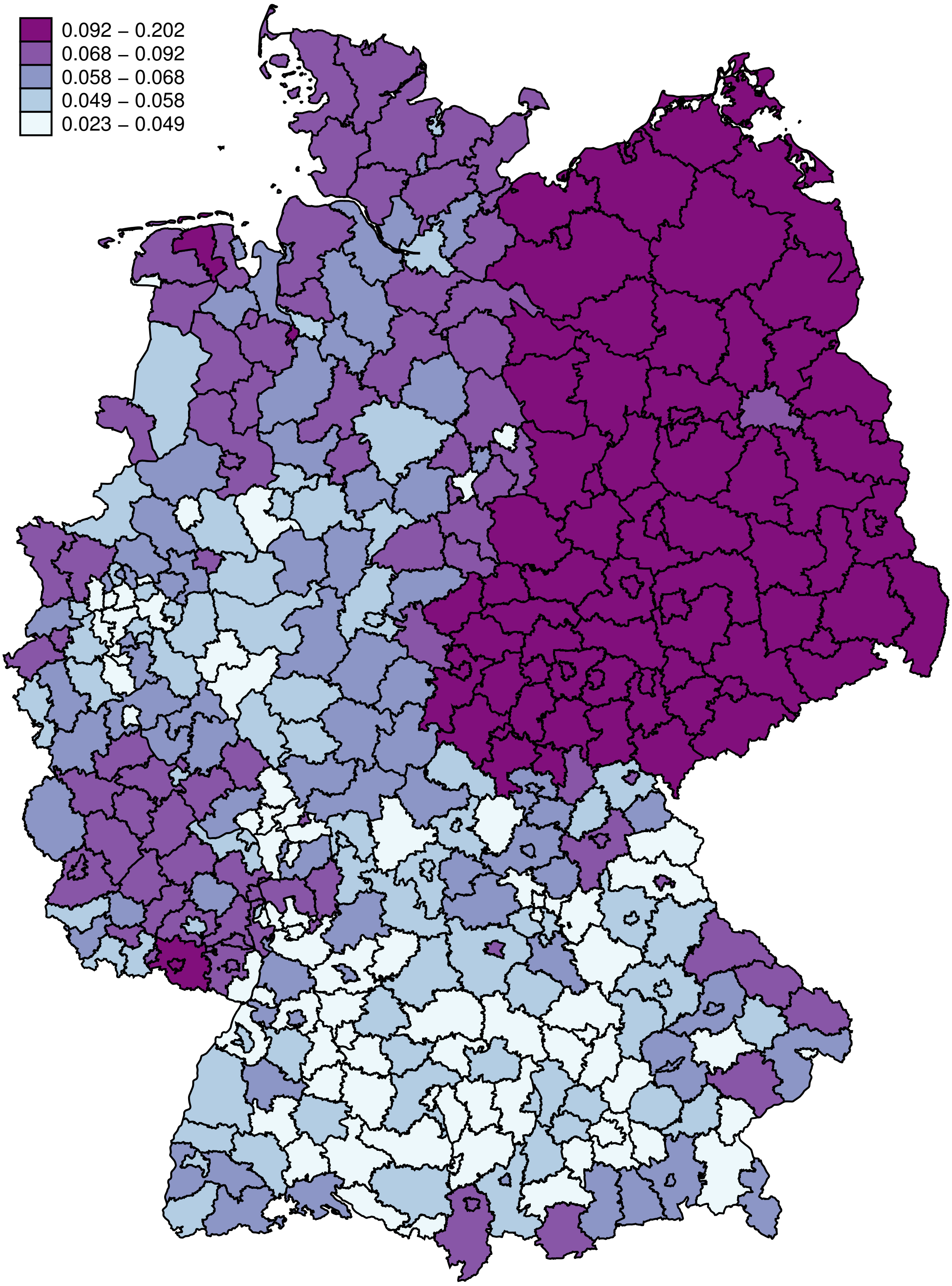
Figure 3.2 maps the distribution of the bite across counties. It shows five different quintiles of the bite in ascending darkness (from light blue to dark purple). The bite ranges from 2.3% to more than 20%; so variation across regions is substantial. The most striking pattern is that the entire former GDR - except for Berlin - is in the highest quintile. This observation matches expectations and echoes the simple East West comparison used by Knabe et al. (2016), however, as displayed in Figure 3.A1 there is also considerable variation within East German regions. The variation within West Germany confirms intuitions: The prosperous regions in Bavaria and Baden-Württemberg in the South of Germany and other metropolitan regions (e.g. Frankfurt Rhine-Main are, Düsseldorf Rhine-Ruhr, Hamburg) are less affected than more rural and less prosperous areas such as the south of Rhineland-Palatinate or East Frisia.

Accordingly, I generate two alternative binary treatments. First, by splitting the counties into those with a bite above and those with a bite below the median bite. This

[18]Assuming 4.35 workweeks per month and reasonable 38 hours work week implies a gross hourly wage rate of € 8.47. As a robustness check, the main results are also replicated with bite measures based on income thresholds of € 1500 and 2000 per months. The bite measures are highly correlated, hence the results remain largely unaffected.

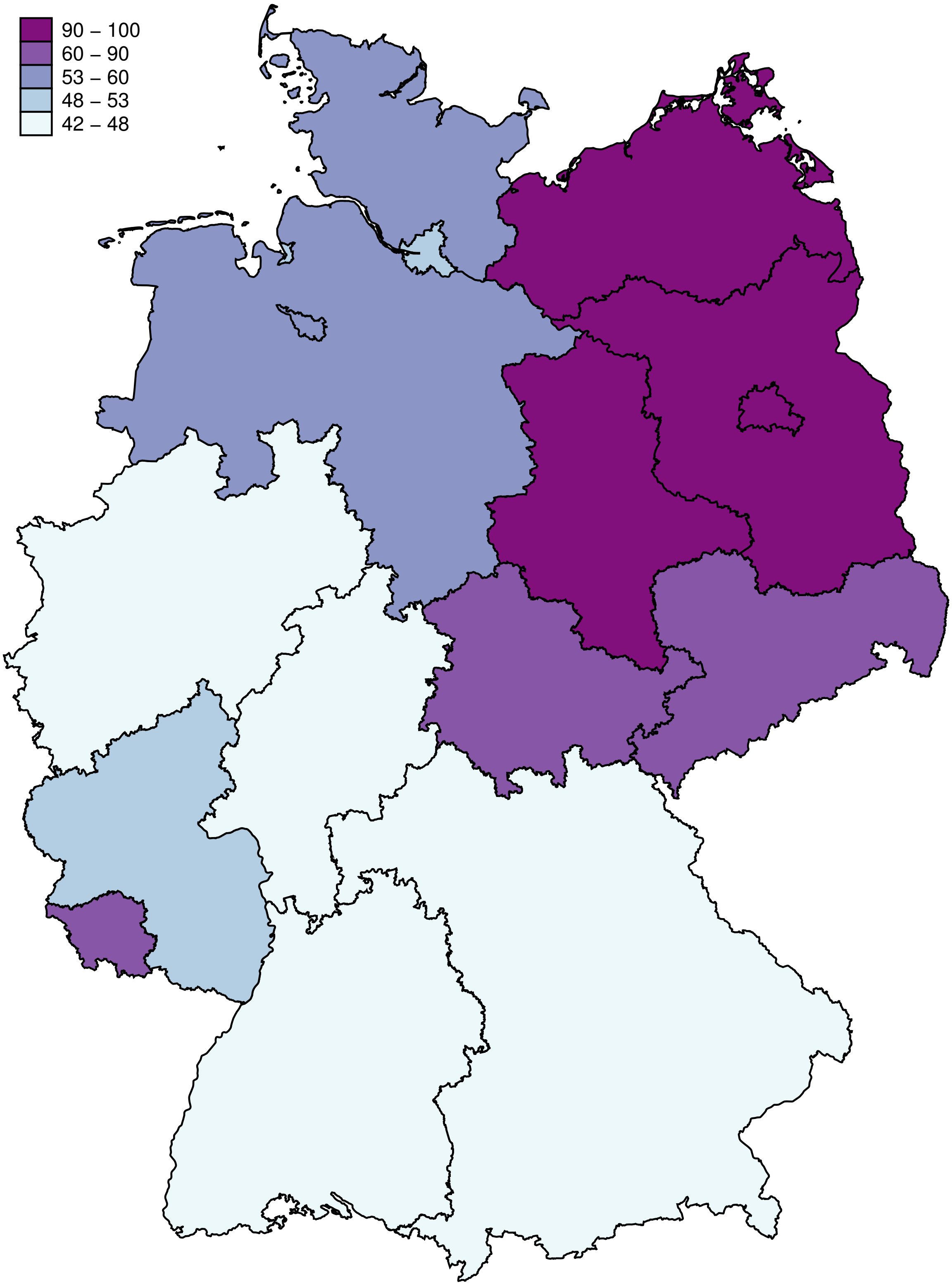
alternative treatment assumes that only the counties with a large bite -above the median- are treated, while the other half is not affected. The resulting map is shown in the Appendix (Figure 3.A2 - left panel). The second alternative binary treatment ignores counties close to the median on either side, since they are very similar in their exposure to the new policy. In this *robust* binary treatment, counties are *treated* if their bite lies above the 60th percentile of the bite distribution, counties below the 40th percentile belong to the control group, and counties close to the median (above 40th and below 6oth percentile) are excluded. The resulting map is also shown in the Appendix (Figure 3.A2, right panel).

One can exploit the regional variation in Google search queries to validate the bite measure. Figure 3.3 shows the relative search intensity for the term *"Mindestlohn"* across federal states in the years 2013 to 2015. The search intensity index is set to 100 in the state in which it was most popular relative to all search queries (Mecklenburg-Vorpommern). A value of 50 implies that the term was only half as popular than in the reference region. The lowest value (42) is observed for Baden-Württemberg (Google Trends, 2017). The map also uses quintiles of the search intensity distribution. Although the data is only available at the coarser state level, the resulting map shows that the bite is highly correlated with interest in the new policy. Regions with a higher bite also have a higher relative search intensity for the minimum wage. As argued by Askitas and Zimmermann (2009), Google search queries can be a powerful predictor for the analysis of labor market outcomes. Hence, the bite measure appears to be a reasonable indicator for the strength of the treatment.



Note: Measured in December 2013

Figure 3.2: Average Bite of the Minimum Wage across Counties ,



Note: Measured between January 2013 and December 2015, Own Graph, Data: Google Trends (2017)

Figure 3.3: Relative Search Intensity *"Mindestlohn"* across Federal States

3.3.3 Data

The impact of the minimum wage will be evaluated using the results from four main outcomes: Concerning the employment effects, regular and marginal employment are considered. Concerning welfare dependency, the stock of recipients of unemployment benefit II, who are deemed to be able to work, will be decomposed into those working (*Aufstocker*) and those not working (*NW UBII*).[19] In order to get an in-depth look at the effects for the *Aufstocker*, I will decompose them based on monthly income (Minijobs, Midijobs and Maxijobs) and employment status (self-employed or dependently employed). The county-specific, monthly time series for all four outcomes are provided by the Federal Employment Agency (Statistik der Bundesagentur für Arbeit, 2016a,b) for the years 2012 to 2015.

Table 3.1: Totals for December 2013

Main Outcomes		*Composition of Aufstocker*	
Regular Employment	29.88m	Self-employed	118,584
Marginal Employment	7.44m	Employees	1,189,417
Able to Work - UBII	4.31m	Minijobs	639,942
Aufstocker	1.30m	Midijobs	233,757
Non-working UBII	3.01m	Maxijobs	315,718

Sum over all 402 counties

Table 3.1 displays the aggregate values of the outcomes of interest for December 2013. The left column contains the employment and welfare outcomes, the right column the composition of the *Aufstocker*. The data contains just short of 30 million regular jobs and 7.5 million marginal jobs. Out of the 4.3 million recipients of UBII who are deemed to be able to work, about 30% (1.3m) top-up their labor income with welfare payments. In the public debate, this group was considered a core target group of the minimum wage, even though economic research (Müller and Steiner, 2009; Bruckmeier and Wiemers, 2014) dampened expectations about the effectiveness of minimum wages. Figure 3.A3 in the Appendix illustrates the composition of the *Aufstocker* graphically. Circa 9% of the *Aufstocker* are self-employed, but do not earn enough. Almost 50% of the *Aufstocker* only have a marginal job, paying up to € 450 per month. Table 3.1 highlights that the vast majority of marginal jobs are not held by people depending on welfare benefits, but by people from households that do not receive welfare payments. Marginal jobs are often held by secondary earners, due to the favorable tax treatment.

[19] There are also recipients of UBII that are not deemed to be able to work, namely the unfit household members of those who are able to work and receive UBII. Thus, the overall number of UBII recipients in Germany is larger.

Table 3.2 shows the most important variables for December 2013, i.e. the reference period for the bite measure. The average bite from the wage statistics is 6.8% and ranges from 2 to 20%. All outcomes display considerable variations across counties, which is not surprising, given that these administrative units are very heterogeneous. Recall that county fixed effects will be included in the estimation and identification is based on inter-temporal differences. The table does not only feature the outcomes of interest, but also the average labor earnings in 2013 and 2015 from the wage statistics (Statistik der Bundesagentur für Arbeit, 2016c).[20]

Table 3.2: Summary Statistics for December 2013

	Mean	**Std. Dev.**	**Min.**	**Max.**
Bite	0.068	0.034	0.023	0.202
Regular Employment	211,742	292,078	12,023	1,250,649
Marginal Employment	44,379	50,972	2835	210,496
Aufstocker	11,185	24,996	215	127,939
Non-working UBII	26,183	55,993	550	283,446
Avg Income 2013	3556	628	2220	5082
Avg Income 2015	3723	649	2361	5386
Composition of Aufstocker				
Employees	9868	21,018	205	107,192
Self-employed	1420	4297	8	22,128
Minijobs	4951	9955	105	50,849
Midijobs	2174	5004	40	25,309
Maxijobs	2744	6086	42	31,034

N=402 counties, weighted with county-specific employment.

As outlined in Section 3.3.1, the DID framework will make use of inter-temporal variation in order to identify the effects of the minimum wage. Figure 3.4 shows the development of the four outcomes (in logs) over time. The graphs run from January 2012 to December 2015 showing time series normalized by the value of January 2012. Hence, all lines start at zero and are growth rates with respect to January 2012. A light blue vertical line indicates the passing of the minimum wage bill in July 2014, the dark blue vertical line the turn of the year 2014/15, the introduction of the minimum wage.

[20] Recall from Section 3.3.2 that the wage statistics only feature full-time regular employment. The data is available in € 50 brackets. For each bracket, the mean value is assumed. The data is top-coded. Average income in the highest bracket is imputed, using a Pareto distribution with $\alpha = 2.6$. The imputation for the top income bracket does not affect the bite, given that the bite is defined only as the share of monthly labor income below 1400 over all full-time employees.

start of the new policy. Note that the series are not yet separated by treatment or intensity of treatment, which will be delegated to the following subsection.

In the upper graph, the two employment outcomes are shown. First, the upward trend for both employment outcomes indicates that Germany introduced the minimum wage into a very robust and expanding labor market. Not surprisingly, the series exhibits a stochastic pattern of seasonality. Strikingly, there is a pronounced reduction in marginal employment exactly after the turn of the year 2014/15. This observation matches the descriptive evidence concerning marginal employment reported for instance by Groll (2015). There does not seem to be a comparable movement in the series of regular employment; however, the relatively small seasonal decrease in the winter 2014/15 could indicate that at least some of the marginal jobs have been converted into regular employment. The stock of regular employment is about four times as large as the stock of marginal employment; hence, in a graph displaying growths rates, such a transformation is certainly difficult to spot.

The lower graph shows the development for the working and non-working recipients of unemployment benefit II. As in the upper graph, there are strong seasonal patterns; however, contrarily to the employment outcomes, the patterns appear to be shifted. This observation indicates that people frequently shift from one category to the other. The non-working series seems to be pretty stable in its average level prior to the introduction of the minimum wage. There seems to be a small upward trend in the first months of 2015. Concerning the working UBII recipients (*Aufstocker*), there seems to be a slight downward trend before the introduction of the policy; however, this downward trend is amplified after the passing of the minimum wage bill in July 2014 with a very pronounced drop in the first months of 2015.

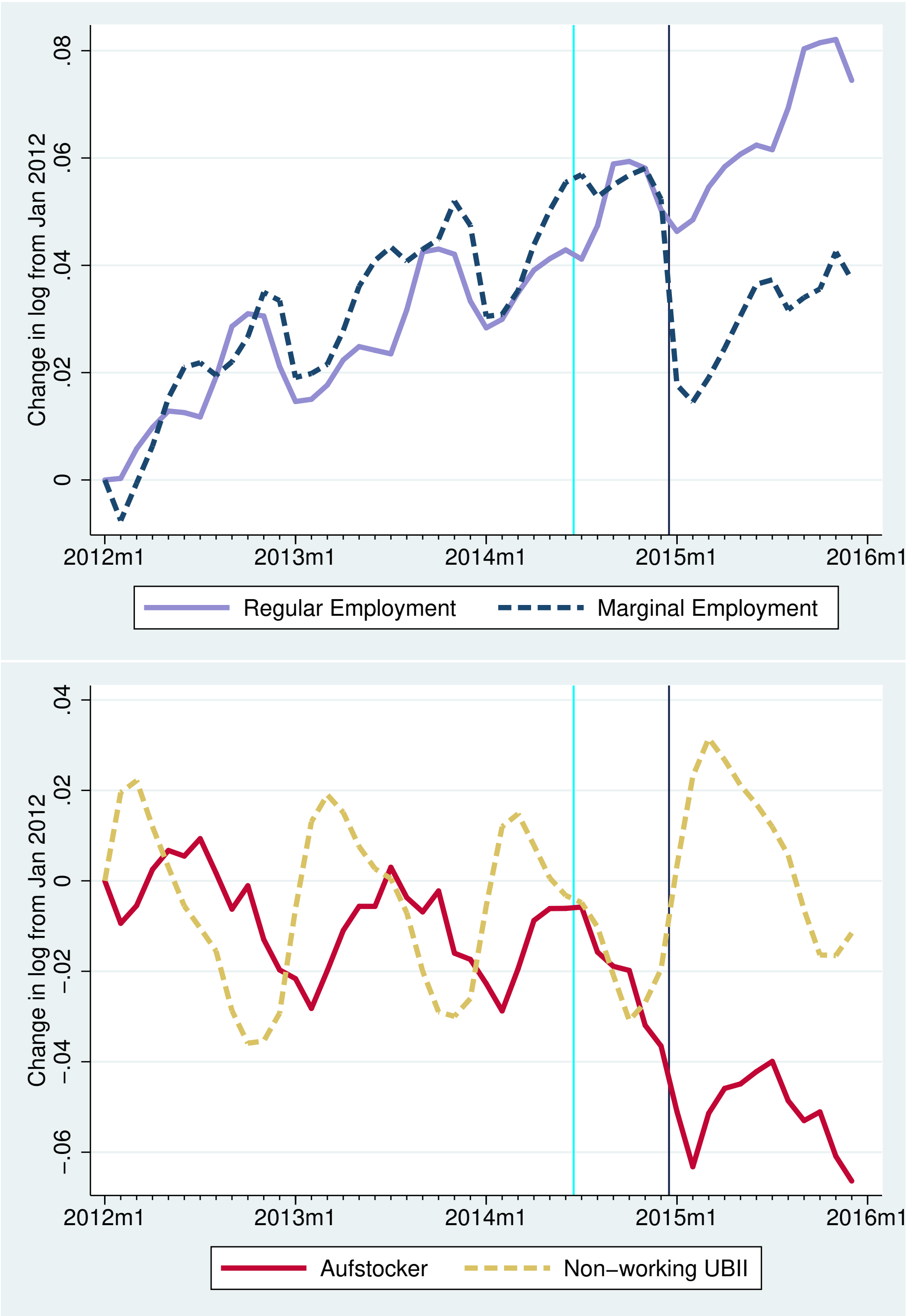


Figure 3.4: Employment and Welfare Outcomes over Time

3.3.4 Graphical Evidence & Trend Assumptions

The graphs in Figure 3.4 provide a descriptive account of the events; nevertheless, they suggest that the minimum wage might had a negative effect at least on marginal employment and on *Aufstocker*. However, these graphs do not account for the differences in the strength of the treatment and possible pre-trend differences in deterministic trends. Additionally, the series are noisy, due to the presence of seasonality. In this section, these concerns will be addressed.

Another goal of this section is to answer two important questions concerning the regression specification. First, whether the left-hand side of the equation should be specified in log-levels (Equation 3.1) or in growth rates (Equation 3.2), and second, whether to include a pre-treatment trend differential or not. As outlined in Section 3.3.1, identification hinges on the so-called common trend assumption (CTA), which is not directly testable. However, if for a certain specification more and less heavily treated counties move parallel *before* the policy change, one would be more confident hat under the absence of the new policy, the parallel movement would have continued also *after* the policy change. Thus, this section compares graphically the different specifications.[21]

In order to ease the graphical exposition, treatment won't be based on the continuous bite, but on the binary treatment indicator.[22] The resulting two time series are normalized by the average value in January 2012 (for levels specification) or 2013 (for the growth rate specification) respectively.[23] As in Figure 3.4, a vertical dark blue line indicates the official start of the minimum wage in January 2015; a light-blue vertical line six months before indicates the the passing of the law in July 2014.

Concerning the first question, the specification of the left-hand side, the time series for the levels specification is based on the residuals of a regression on time and county dummies. For the growth rate specification, the time series is differenced. As it is evident from Figure 3.4, all four outcomes exhibit strong seasonal patterns. The levels approach treats these patterns deterministically with a set of time-specific fixed effects, while the growth rate approach removes the seasonality *stochastically* by differencing.

For the second question, the pros and cons for the inclusion of a pre-treatment trend differential, the resulting time series can also be adjusted by the interaction of a trend

[21] Besides this graphical exposition, I will report in Section 3.4 the estimated coefficients of the pre-treatment trend differentials. Note that a test of their statistical significance cannot directly test the CTA, since the assumption refers to the hypothetical behavior *after* the policy change.

[22] Conveniently, one can compare the movements of two distinct groups, rather than 402 different counties. Treatment and control group are averages of their respective counties, weighted by the county-specific regular employment in December 2013. Recall that in the binary treatment counties are *treated* if the bite is above the median bite; otherwise, counties belong to the control group. The resulting binary treatment is displayed in a map (Figure 3.A2, left panel) in the Appendix.

[23] Hence, values can be interpreted as percentage changes relative to January 2013 net of seasonal effects. The structure of the graph is similar to Figure 1 in Angrist and Krueger (1999) referring to the Mariel Boatlift study by Card (1990).

term (linear, quadratic, or in logs) with the region-specific bite.[24] Due to this adjustment, the values for this graph can no longer be read as a seasonally adjusted growth rate relative to the first period. The resulting series are the ones to be compared for the plausibility of the CTA, conditional on a deterministic pre-treatment trend differential.[25]

Concerning the first question, the graphical analysis suggests that the log-levels specification is not appropriate to distinguish seasonality and trend behavior. The graphs without trend differential and a linear trend differential for the four outcome variables are shown in the Appendix in Tables 3.A4 to 3.A7. The seasonal pattern is not entirely removed and there remain important differences between counties above and below the median. Hence, I will not consider the log-levels specification except for a robustness check and focus on the specification in growth rates (Equation 3.2). Thus, all remaining graphs in this section are seasonally adjusted by taking the 12-month difference. Additionally, the graphs are centered relative to the value for January 2013.

Figure 3.5 shows the two relevant graphs for the growth rates of regular employment. The upper panel shows seasonally adjusted data, the lower panel additionally corrects for a deterministic trend differential. In the upper panel, the two lines move parallel almost everywhere, except a short period in Spring 2014. A common trend appears plausible, and there seems to be no striking effect after the introduction of the minimum wage. In the graphs controlling for a trend differential, the movements before treatment are similar but less congruent. Thus, the graphical analysis speaks in favor of a specification *without* an additional trend differential.

For marginal employment, the two corresponding graphs are displayed in Figure 3.6. The upper and the lower panel appear to be very similar. In both cases, treatment and control group move parallel and almost horizontal until the minimum wage comes into effect. Both series experience a drop at the turn of the year; however, the one in the treatment group is much more pronounced. Recall that the smaller drop in the control regions is likely also related to the minimum wage, given that also the control group is partially treated, but only to a lesser extent.

Figure 3.7 displays the development of the growth rates of the *Aufstocker*. Contrary to the employment outcomes, there seems to be a clear discrepancy between the treatment and the control group before the introduction of the minimum wage. If one does not control for a trend differential (upper panel), the CTA appears implausible. Fortunately, the picture changes, once one controls for a simple linear trend differential. The movements of the control and the treatment are parallel until the introduction of the minimum wage. In 2015 a large discrepancy appears between the two groups.

[24] Technically, the seasonally adjusted values are regressed on the bite-trend interaction for the sample from January 2013 till December 2014. The remainder of the procedure described above is performed on the predicted (out of sample) residuals from this regression.

[25] I will focus on a linear trend specification. The resulting graphs for the growth rate specification with a quadratic polynomial or a logarithmic trend differential are shown in the Appendix (Figures 3.A8 to 3.A11)

The last of the four outcomes in growth rates, the non-working UBII recipients, are displayed in Figure 3.8. There are apparent similarities to the graphs for the *Aufstocker* in Figure 3.7. Without any trend correction, the CTA seems implausible. However, as soon as one controls for a deterministic linear trend differential, the movements align. Unlike the *Aufstocker*, there is no strong indication for a diverging movement of the treatment group after the policy change.

Summing Up The graphical inspection of the binary treatment and control groups indicates that the growth rates specification is more appropriate than the specification in log-levels. Additionally, the CTA should be satisfied for the two employment outcomes without any inclusion of pre-treatment trend differentials. For the two outcomes studying welfare dependency however, a specification with a deterministic trend differential seems more plausible. The graphical analysis suggests the presence of effects for marginal employment and the *Aufstocker*. For regular employment and non-working UBII the graphs do not reveal any striking impact.

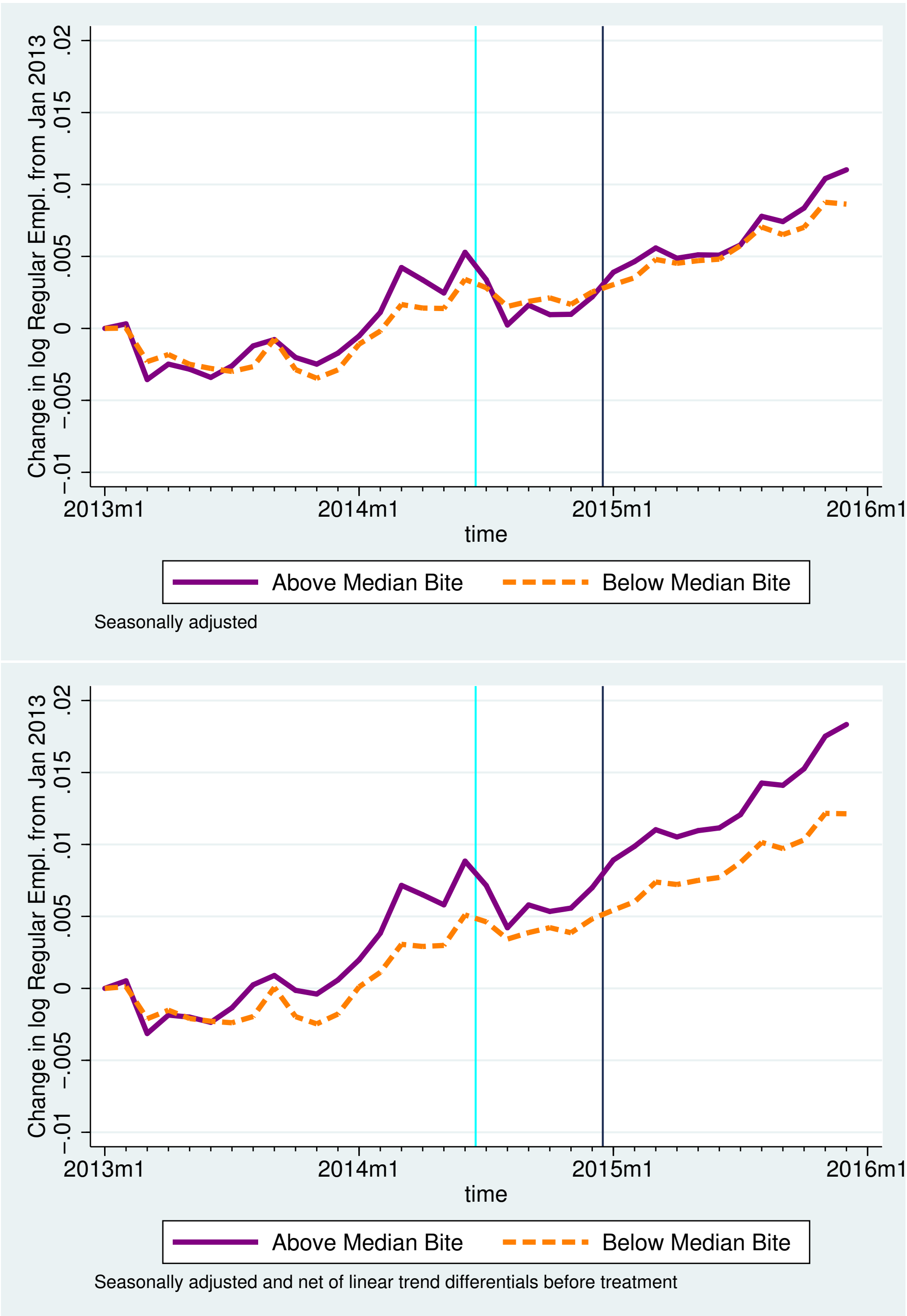


Figure 3.5: Regular Employment - Comparing Different Specifications

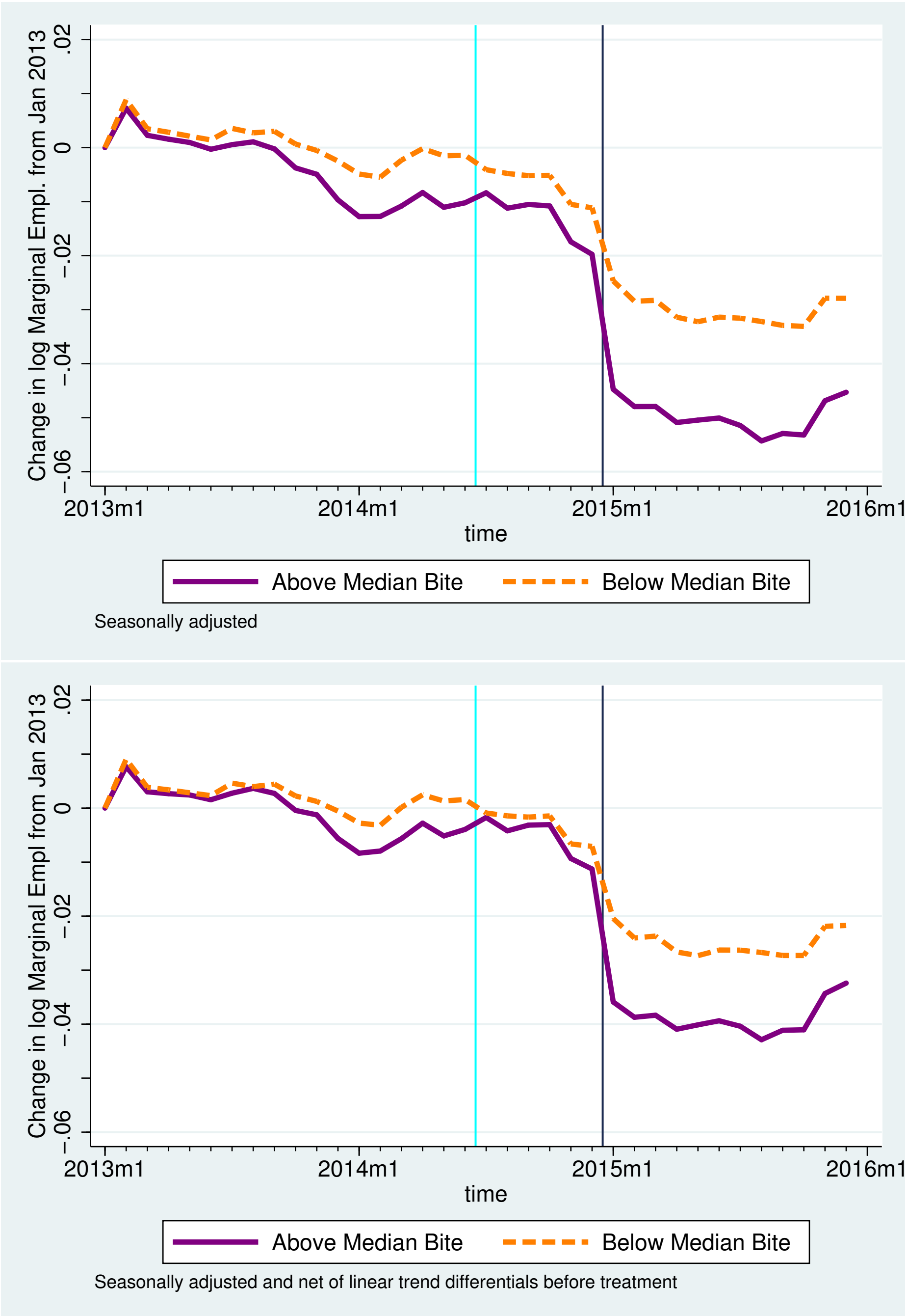


Figure 3.6: Marginal Employment - Comparing Different Specifications

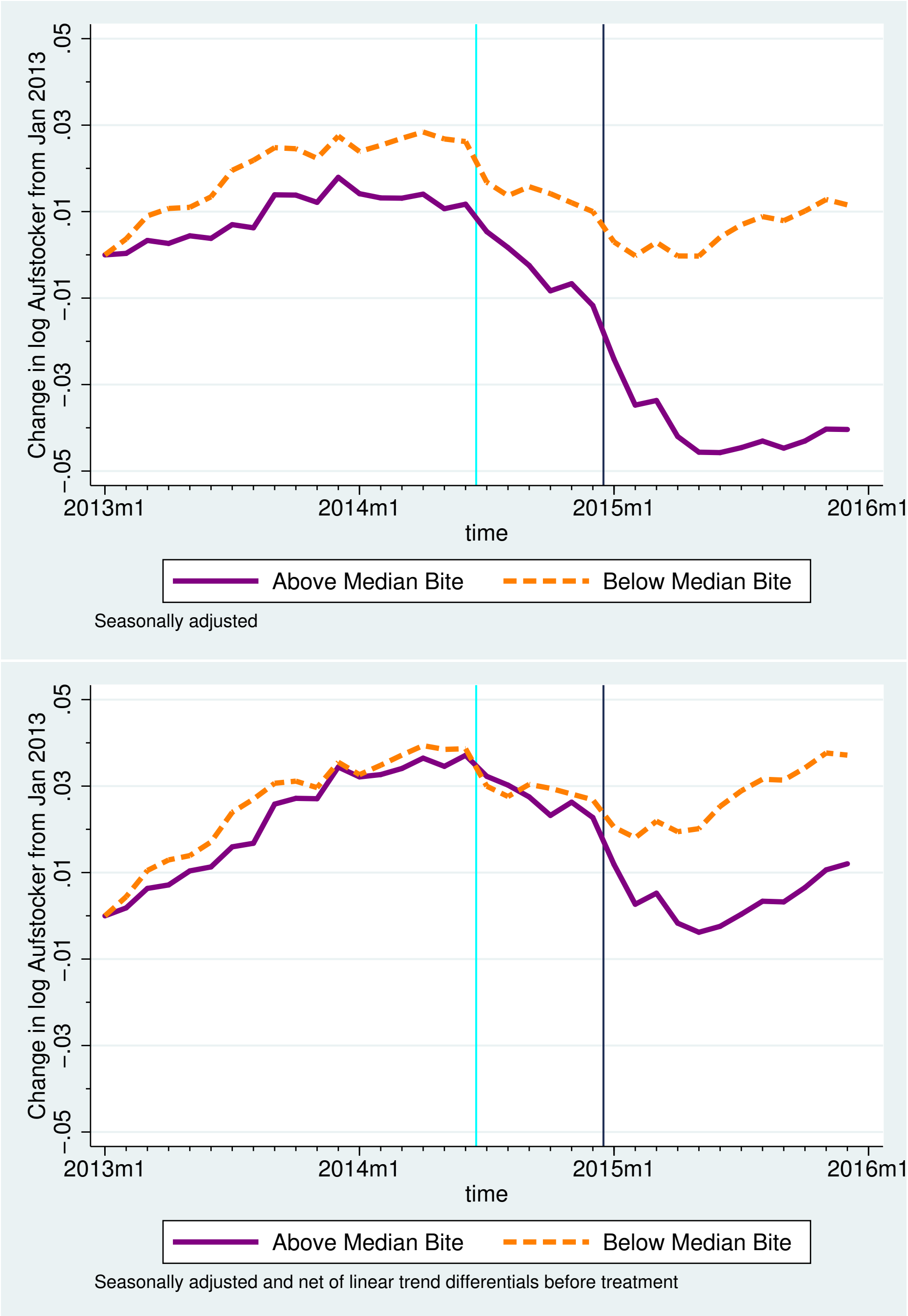


Figure 3.7: *Aufstocker* - Comparing Different Specifications

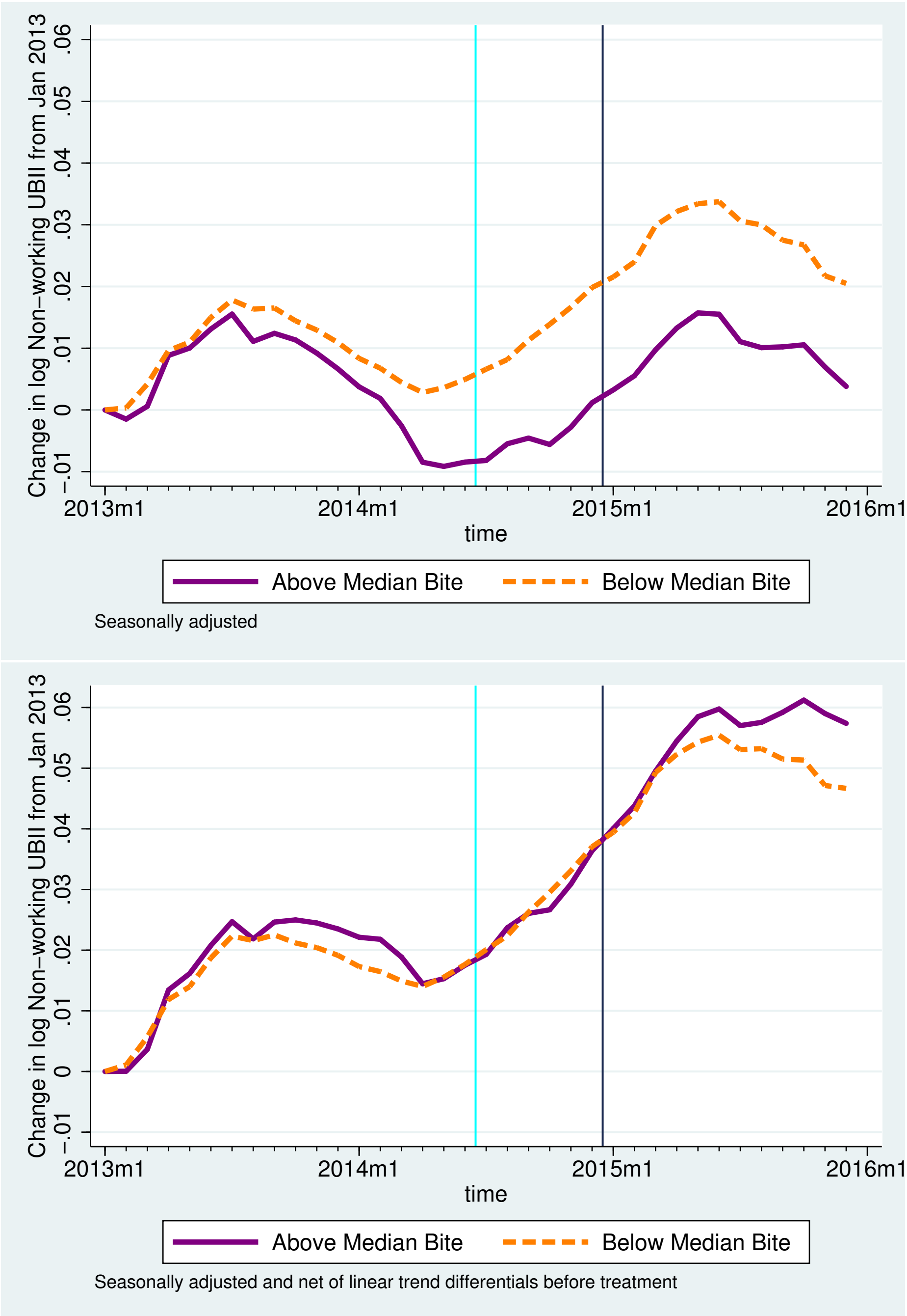


Figure 3.8: Non-working UBII - Comparing Different Specifications

3.4 Results

This section presents the results from fixed effects regressions as described in Section 3.3. As argued in the previous Section 3.3.4, the main specification will be in growth rates of the outcomes of interest (Equation 3.2). Note that in order to ease interpretation of the estimated coefficients, the county-specific bite of the minimum wage is normalized by its standard deviation and divided by 100. Thus, $\hat{\beta}$ gives the *percentage point change* of the growth rate of the respective outcome (regular or marginal employment, *Aufstocker*, non-working UBII) due to one additional standard deviation of the bite.[26] All tables will also report the overall, as well as the between and the within R2 measure. The between R2 represents the explained variation if the time-dimension would be collapsed on county level. The within R2 provides the goodness of fit for the mean detrended data, neglecting all variation across counties.

Employment Table 3.3 shows the effects of the minimum wage on employment outcomes. The upper panel displays the results for regular employment, the lower panel for marginal employment. Column (1) is the standard model from Equation 3.2, including time and county fixed effects, but no trend differential. As it was argued in the graphical analysis in Section 3.3.4, column (1) is the preferred specification for both, regular and marginal employment. The corresponding graphs are the upper panels in Figure 3.5 and 3.6. The specification in column (2) additionally includes $bite_i \cdot t \cdot \delta$, i.e. a linear trend differential proportional to the bite. This specification corresponds to the lower panels in Figure 3.5 and 3.6. As additional robustness checks, column (3) specifies a polynomial quadratic trend differential, column (4) a logarithmic one.[27]

For regular employment, all four specifications have a negative point estimate and thus indicate that the minimum wage reduced the growth rate of regular employment. The graphical analysis spoke in favor of the specification in column (1). The estimated coefficient implies that one standard deviation of the bite decreases the growth rate of employment by about 0.1 percentage points, significant at the 10% level. Given that the average bite is about two times the standard deviation, one could argue that such an effect does not appear to be economically significant. However, if the minimum wage would *permanently* depresses the growth rate of employment, the long-run effect could potentially be very large.

The choice of the functional form of the pre-treatment trend behavior appears to be influential, even though the differences across the bite coefficients are not always significant. Table 3.3 also reports the estimated trend coefficients; the linear and the log-trend differential are not statistically significant, while the quadratic trend differential

[26] In a robustness check, also the log-levels specification will be estimated. Here, $\hat{\beta}_L$ corresponds to the *percentage change* in the outcome variable due to one additional standard deviation of the bite.

[27] The corresponding graphs are shown in the Appendix in Figure 3.A8 and 3.A9.

is jointly highly significant.[28] Thus, on statistical grounds, one could argue that the specification from column (3) with a quadratic trend should be the preferred one. In that case, the effect of one standard deviation becomes very small and is no longer statistically different from zero at conventional levels of significance.

For marginal employment, the lower panel of Table 3.3 shows that effect on growth rate of marginal employment is more pronounced than the one on regular employment. This is in line with Figure 3.4 and previous evidence on the employment effects of the minimum wage. In the preferred specification, an additional standard deviation of the bite reduces the growth rate by 1.4 percentage points. This estimate is reduced to about 1 percentage point, if one controls for a pre-trend differential, regardless of the trend's functional form. Regarding the significance of the estimated trend differentials, all three trend differentials are statistically significant at least at the 5% level.[29] All three specifications including a trend differential indicate more or less the same effect of a 1 percentage point reduction of the growth rate. Thus, there is strong evidence that the minimum wage had a negative effect on marginal employment. Given that all point estimates for regular employment are negative, the results so far do not suggest that marginal jobs have been upgraded to regular jobs on a massive scale.

[28] The F-test of joint significance of the linear and quadratic trend differential has a test statistic of 6.0881 with an associated p-value of 0.0025.

[29] The F-test of joint significance of the linear and quadratic trend differential has a test statistic of 3.3733 with an associated p-value of 0.0353.

Table 3.3: Effects on Employment Outcomes

	(1)	(2)	(3)	(4)
		Regular Employment		
Bite·2015	-0.096*	-0.160***	-0.040	-0.170***
	(0.050)	(0.060)	(0.041)	(0.057)
t × Bite		0.001	0.009***	
		(0.001)	(0.003)	
t^2 × Bite			-0.000***	
			(0.000)	
log(t) × Bite				0.035
				(0.022)
Observations	14472	14472	14472	14472
R2 within	0.171	0.171	0.173	0.172
R2 between	0.177	0.177	0.177	0.177
R2 overall	0.097	0.058	0.001	0.000
		Marginal Employment		
Bite·2015	-1.359***	-0.995***	-1.133***	-1.065***
	(0.206)	(0.245)	(0.229)	(0.240)
t × Bite		-0.006***	-0.015	
		(0.002)	(0.010)	
t^2 × Bite			0.000	
			(0.000)	
log(t) × Bite				-0.138**
				(0.055)
Observations	14472	14472	14472	14472
R2 within	0.306	0.308	0.308	0.308
R2 between	0.110	0.110	0.110	0.110
R2 overall	0.249	0.249	0.231	0.196

Bite measured in December 2013 and normalized by one SD
Estimation including Time and County Fixed Effects
Cluster-robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Welfare Dependency While there already exist several studies about the employment effects, ex-post evidence on the effects of the minimum wage on the *Aufstocker* and welfare dependency is scarce. At the time of writing, there is only the descriptive study by Bruckmeier and Wiemers (2016), reporting an unusually large reduction of the *Aufstocker* at the turn of the year 2014/15 which is also present in the descriptive Figure 3.4 in Section 3.3.3.

Table 3.4 follows the same structure as Table 3.3, but this time for the two welfare-related outcomes. The number of observations is slightly smaller, since the number of welfare recipients are sometimes missing in the original data. Recall from the graphical exposition in Section 3.3.4 that for both welfare outcomes column (2) (linear trend differential) was preferred over the specification without any trend differential. For the *Aufstocker*, all four point estimates are negative and statistically significant at the 1% level. The preferred specification in column (2) indicates that one standard deviation of the bite reduces the growth rate of the *Aufstocker* by 1.4 percentage points. The point estimate from column (1) without and trend differential is substantially larger with 2.6 percentage points. This discrepancy was already indicated in the graphical exposition in Section 3.3.4, as the difference between the upper and lower panel of Figure 3.7. Concerning the estimated trend differentials, all three trend specifications are (jointly) statistically significant In sum, the results point without any doubt to a reduction of the the growth rate of the *Aufstocker*.

For the other welfare-related outcome, the growth rate of non-working UBII recipients, the results do not draw such a clear picture. The sign of the point estimate switches from negative to positive, if one includes any form of trend differential. The preferred estimate in column (2) with a linear trend differential indicates a small and significant increase in the growth rate of non-working UBII recipients by 0.3 percentage points due to one additional standard deviation of the bite. The other two trend-corrected specifications result in smaller and insignificant point estimates. All trend terms are highly significant. Thus, it is difficult to conclude which of the specifications is the most credible and consequently, whether there is a significant increase in the non-working welfare recipients. In any case, given that the strong effect for the *Aufstocker* is not matched by an equally striking effect on the non-working UBII recipients, the results from Table 3.4 suggest that the reduction in the *Aufstocker* is not entirely due to lost supplementary jobs, but potentially also due to increased labor earnings of the *Aufstocker*.

Table 3.4: Effects on Welfare Outcomes

	(1)	**(2)**	(3)	(4)
		Aufstocker		
Bite·2015	-2.617***	-1.460***	-1.249***	-1.745***
	(0.156)	(0.209)	(0.155)	(0.198)
t × Bite		-0.018***	-0.003	
		(0.003)	(0.012)	
t^2 × Bite			-0.000	
			(0.000)	
log(t) × Bite				-0.411***
				(0.074)
Observations	14298	14298	14298	14298
R2 within	0.238	0.248	0.249	0.247
R2 between	0.530	0.533	0.532	0.534
R2 overall	0.283	0.382	0.368	0.352
		Non-working UBII		
Bite·2015	-0.635***	0.321**	0.091	0.143
	(0.109)	(0.139)	(0.108)	(0.131)
t × Bite		-0.015***	-0.031***	
		(0.002)	(0.008)	
t^2 × Bite			0.000**	
			(0.000)	
log(t) × Bite				-0.367***
				(0.051)
Observations	14298	14298	14298	14298
R2 within	0.105	0.120	0.121	0.121
R2 between	0.446	0.446	0.446	0.446
R2 overall	0.111	0.294	0.296	0.279

Bite measured in December 2013 and normalized by one SD
Estimation including Time and County Fixed Effects
Cluster-robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Back-of-the-envelope The specification in growth rates and with a continuous treatment has the important drawback that it is difficult to translate the results into an easily understandable effect size. It is for instance not clear for which time horizon the minimum wage will affect the growth rates. It is unlikely that the minimum wage only has an impact in 2015. On the other hand, if the measured impact was *permanent*, it would be implausibly high in the long run. Concerning the treatment, it is not clear how many standard deviations should be the yardstick for the effect of the policy. Table 3.5 presents back-of-the-envelope calculations for the implied size of the short-run effect in 2015, using two standard deviations - which is about the size of the mean bite - as the preferred multiplier. The estimates should not be taken as *the* definite *treatment effect* of the minimum wage, but illustrate the order of magnitude of the implied effects. The calculations in Table 3.5 use the stock of the outcomes in December 2013. The $\hat{\beta}$-coefficients are those without adjustment for a trend-differential for employment outcomes (Column (1) in Table 3.3) and those with a linear trend-differential for the two welfare outcomes (Column (2) in Table 3.4).[30]

Table 3.5: Effects of Preferred Specifications

	Employment		Welfare	
	Regular	Marginal	*Aufstocker*	NW-UBII
Stock in 2013	29,883,573	7,438,102	1,298,297	3,016,337
$\hat{\beta}$	-0.096	-1.359	-1.460	0.321
Effect +1 SD	-28,688	-101,084	-18,955	9,682
Effect +2 SD	-57,376	-202,168	-37,910	19,365

Back-of-the-envelope calculation, short-run effects in 2015

For regular employment, the effect of two standard deviations corresponds to about 60,000 less jobs due to the minimum wage. Compared to the stock of employment of about 30 million employees, and the large predicted long-run effect of some of the ex-ante studies, this *short-run* effect is comparatively small. Additionally, the effect does not appear to be robust to the inclusion of pre-treatment trend differentials.[31] Note that this calculation does not imply that existing jobs are lost, but that the job creation dynamics are hampered. As it is displayed in the employment graphs in Figure 3.4, regular employment followed an upward trajectory in recent years. Knabe et al. (2016) also argue that the minimum wage did not destroy existing jobs but did reduce job

[30] Given that tests on the statistical significance of the pre-treatment trend differentials in Tables 3.3 and 3.4 did not provide a clear guidance, this choice is somewhat arbitrary. Especially the results for regular employment and non-working welfare recipients have to be taken with a pinch of salt. I decided to stick to the preferred specifications of the graphical analysis, since these are credible, but also parsimonious.

[31] For the specification with a quadratic trend differential, the effect would essentially zero.

creation. For marginal employment, the effect of two standard deviations of the bite is about 200,000 lost marginal jobs or about 150,000 for the robustness check including a trend differential For the level of marginal employment, the figures and the estimation results indicate an actual (and not only counter-factual) reduction due to the minimum wage, which is also in line with previous research, such as Groll (2015) or Garloff (2016).

For the welfare outcomes, the two standard deviations imply a reduction of about 38,000 *Aufstocker* and an partially offsetting increase in the number of non-working UBII recipients by about 19,000. This calculations suggest that roughly one half of the reduction in the *Aufstocker* was due to the loss of a supplementary job, instead of an increase in the household income. However, the effect on non-working UBII has to be taken with a pinch of salt, because the estimated coefficients fluctuate considerably. The very small absolute reduction of the *Aufstocker* due to the minimum wage also confirms the previous literature, which pointed to the limited effectiveness of minimum wages for reducing welfare dependency.

Differences between East and West Germany Table 3.6 uses only variation in the bite *within* East and West Germany.[32] The table presents only the preferred specification, i.e. without any pre-treatment trend differential for the two employment outcomes (column 1 and 2), but with a linear trend differential for the two welfare related outcomes (column 3 and 4). The Upper panel shows the results for West Germany, the lower panel for the East.

Concerning employment outcomes within West Germany, the result suggest positive job dynamics in counties more heavily affected by the minimum wage: The growth rate of marginal employment is reduced by about 0.6 percentage points, however, this reduction appears to be offset by an increase of the growth rate of regular employment of almost 0.3 percentage points. Give the relative magnitudes of the two types of employment, this pattern can be seen as evidence for upgrading of marginal into regular jobs due to the minimum wage. For East Germany however, both point estimates for the employment outcomes are negative and relatively large, even though they are not statistically significant at conventional levels.

For the welfare outcomes within West Germany, the point estimate for the *Aufstocker* is positive but not significant. For the non-working UBII recipients, the results indicate a statistically significant *reduction* of the growth by about 1 percentage point. The welfare effects in the West are difficult to reconcile. For East Germany however, the growth of the *Aufstocker* is reduced by 2.3 percentage points, while the growth of non-working UBII recipients increases by 0.9 percentage points. Thus, the overall effects on welfare dependency shown in Table 3.4 seem to be driven by changes in the East. Given the relative magnitudes of working and non-working welfare recipients, the estimates suggest that most of the *Aufstocker* in the East ended up in non-working welfare dependency.

[32]The corresponding maps of the bite are shown in the Appendix in Figure 3.A1.

On balance, the results from this sample split paint a rather positive picture of the short-run effects of the minimum wage in West Germany and a negative for the East. In the West, some marginal jobs seem to be upgraded due to the minimum wage. This dynamic is not present in the East, where some of the *Aufstocker* appear to have lost their supplementary jobs. These differences echo the findings of Knabe et al. (2014) and the concerns in the public debate whether a universally set minimum wage could be workable in the west, but too high for the east of Germany.

Table 3.6: Variation within West and East Germany

	Employment		**Welfare**	
	Regular	Marginal	*Aufstocker*	NW UBII
	(1)	(2)	(3)	(4)
		West Germany		
Bite·2015	0.280**	-0.659**	0.796	-1.010**
	(0.124)	(0.262)	(0.717)	(0.471)
Observations	11700	11700	11558	11558
R2 within	0.183	0.279	0.078	0.103
R2 between	0.003	0.017	0.050	0.034
R2 overall	0.095	0.208	0.066	0.068
		East Germany		
Bite·2015	-0.287	-0.955	-2.330***	0.900*
	(0.188)	(0.971)	(0.495)	(0.463)
Observations	2772	2772	2740	2740
R2 within	0.190	0.347	0.735	0.299
R2 between	0.088	0.002	0.057	0.169
R2 overall	0.117	0.264	0.639	0.224
Linear Trend	-	-	✓	✓
Time Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓

Bite measured in December 2013 and normalized by one SD

Cluster-robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Composition of the *Aufstocker* While there seems to be an overall reduction of the growth rate of the *Aufstocker*, it remains to be seen which group was affected most severely by the minimum wage. Table 3.7 shows the effects on four different *Aufstocker* subgroups, namely self-employed and Mini-, Midi- and Maxijobs. As for the parent category - *Aufstocker* - a linear trend differential is included. Note that the number of observations differs slightly for the self-employed due to data availability.

The effect on the self-employed is not statistically significant, which is not surprising, since a minimum wage should not directly affect self-employed *Aufstocker*. The three groups of dependently employed *Aufstocker* however, all feature statistically significant reductions of their growth rates. The relative effect is the largest for midi and maxi jobs. Especially for those with already relatively high earnings, it is plausible that some have left welfare dependency entirely.

Table 3.7: Composition of *Aufstocker*

	Self-empl. (1)	Mini Job (2)	Midi Job (3)	Maxi Job (4)
Bite·2015	-0.834	-0.786***	-2.676***	-2.357***
	(0.733)	(0.278)	(0.455)	(0.401)
Linear Trend	✓	✓	✓	✓
Time Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓
Observations	14294	14298	14298	14298
R2 within	0.012	0.296	0.108	0.068
R2 between	0.126	0.485	0.061	0.419
R2 overall	0.018	0.366	0.083	0.168

Bite measured in December 2013 and normalized by one SD

Cluster-robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Wages This section analyzes whether the introduction of the minimum wage actually affected labor earnings. Table 3.8 shows the results from OLS regressions where the bite measured in 2013 is related to the county-specific growth rate of labor earnings between 2013 and 2015 (upper panel) or to the level of monthly earnings in 2015, controlling for the average level in the county in 2013. Hence, the estimated coefficients in the upper panel can be read as the effect on the growth rate of earnings in percentage points, due to one additional standard deviation of the bite. In the lower panel, the coefficients provides the income change in Euro due to one additional standard deviation of the bite. All earnings information are taken from the wage statistics (Statistik der Bundesagentur für Arbeit, 2016c).

Table 3.8: Earnings and Earnings Growth between 2013 and 2015

	Avg income (1)	Up to 1400 (2)	Up to 1500 (3)	1500 to 2000 (4)	Above 2000 (5)
			Earnings Growth 2015-2013		
Bite	0.716***	0.419***	1.423***	-0.027	-0.891***
	(0.045)	(0.067)	(0.064)	(0.019)	(0.044)
Observations	402	402	402	402	402
R2	0.319	0.090	0.556	0.005	0.515
			Monthly Earnings 2015		
Bite	22.77***	8.27***	18.52***	-2.82***	-27.22***
	(2.761)	(1.032)	(1.013)	(0.498)	(1.911)
Observations	402	402	402	402	402
R2	0.996	0.847	0.896	0.683	0.995

Bite measured in December 2013 and normalized by one SD. Robust standard errors in parentheses.
Earnings Growth is $\frac{\text{Income 2015}-\text{Income 2013}}{\text{Income 2013}}$.
The effect on monthly earnings 2015 is conditional on average monthly earnings in 2013
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column (1) shows specifications which do not condition on the income level, in column (2) the outcomes are based on the average income of those earning up to € 1400, in column (3) the threshold is raised to € 1500, in column (4) of the average income between € 1500 and 2000 per month. Column (5) shows the outcomes based on the average income of those earning more than € 2000 per month. There are positive statistically significant effects for the first three columns, i.e. overall earnings and earnings that are likely affected by the minimum wage The largest relative effect is found column (3), i.e. for those earning slightly more than the minimum wage. This pattern provides strong evidence for a positive effect of the policy on wages. There is no significant earnings growth for those earning between € 1500 and 2000 per month. Also the absolute change for this group does not seem to be *economically* significant. Thus, there is no evidence for strong positive spillover effects across the wage distribution. The average income of those earning above € 2000, a group which is likely not affected by the minimum wage, is negatively related to the bite.

Robustness In the following, I will discuss the results from various robustness checks. All related tables are delegated to the Appendix. Tables 3.A1 contains the employment effects for the levels specification (Equation 3.1) and follows the same structure as Table 3.3, hence compares the four different specifications of the pre-treatment trend differential. Remember that in this specification, the estimated coefficients do not provide an effect on the growth rate in percentage points, but a percentage change of the outcomes. Concerning regular employment, the preferred specification implies a reduction of employment by about 0.8 percent due to an additional standard deviation of the bite. All four specifications show a negative effect, but controlling for a trend differential reduces the estimated coefficients considerably. The relative effect on marginal employment is more pronounced: In the preferred specification, an additional standard deviation of the bite reduces marginal employment by 1.8%. Recall from the graphical analysis in Section 3.3.4 that the common trend assumption did not appear to hold and thus might be misleading. Nevertheless, the levels specifications would point to similar conclusions as the specification in growth rates.

Tables 3.A2 provides the welfare counterpart in levels to Table 3.A1. The specification in levels with a linear trend differential indicates that for the *Aufstocker* one additional standard deviation of the bite reduces the stock by 2.7% . Also all other point estimates are negative and statistically significant, even though their validity is questionable, since the common trend assumption is likely not to hold. For non-working UBII recipients, not all point estimates have the same sign, even though the standard errors point to rather precise estimates. For non-working UBII, also in the log-levels specification, the point estimates oscillate wildly with the chosen specification of the pre-treatment trend differential.

Table 3.A3 summarizes the results from the alternative specification of the growth rate (Equation 3.3) with anticipation effects and adjustment over time. The employment outcomes are displayed in column (1) and (2). The two specifications do include any pre-treatment differential. The welfare-related outcomes are displayed in column (3) and (4) and are estimated including a linear trend differential. The anticipation period starts with July 2014, after the minimum wage bill was passed and hence consists of the six month in the second half of 2014. The adjustment period starts in January 2015 and includes all months until June 2015. The last six months of 2015 are grouped together. For regular employment, the first statistically significant effects arise form October 2014 onwards, for marginal employment from November 2014 onwards. For the welfare outcomes, there is no striking significant anticipation apart from small positive and significant effects just in the month of August 2014. On Balance, the found patterns do not point to any considerable anticipation effects. If one redefines the start of the treatment from January 2015 to October 2014 (the month with the first significant employment effects), and ignores the adjustment procedure over time, the effects remain largely unaffected.

Given that the identification rests on a difference-in-differences framework and that there are multiple time periods, it is natural to test the validity of the identification strategy using a placebo treatment. Table 3.A4 reports the preferred estimates for all four outcomes (no trend differential for the employment and a linear trend differential for the welfare outcomes) on a treatment that starts in January 2014 and ends in December 2014. The information from 2015 is discarded. Ideally, the estimated effects of this pseudo treatment would be close to zero and not statistically significant. Indeed, this is the case for regular employment and non-working UBII recipients. The coefficient for the *Aufstocker* indicates a change of the growth rate by 0.2 percentage points, even though it is not statstically significant at conventional levels. The growth rate of marginal employment is reduced by 0.3 percentage points and significant at the one percent level. This casts some doubts at the identification strategy. Nevertheless, the estimated pseudo-effects are much smaller than the preferred estimates and would be considered not economically significant.

The main analysis so far did not distinguish between female and male employees. Table 3.A5 shows variation within men and women and also another variant in which both gender types are used together. The latter corresponds to the "gender cell" specification in Garloff (2016) and only shows the preferred specifications of the pre-treatment trend behavior. In all three gender specification, there is no effect on regular employment, but always a reduction of the growth rate of marginal employment by about 1 percentage point. For the growth rate of the *Aufstocker* all three gender specifications find negative effects. However, the effect sizes differ considerably across samples. Concerning non-working UBII recipients, the effects vary from zero for the male sample to a positive effect of about 0.4 for women and -0.4 percentage points for the gender cells approach. Taken together, the results from Table 3.A5 confirm the discussion about the emplyoment welfare effects estimated in Table 3.3 and 3.4: The minimum wage has a negligible effect on regular, but a very robust negative effect on marginal employment and on the *Aufstocker*. The effect on the growth rate of non-working UBII is rather sensitive to the chosen estimation method and thus not very reliable.

Table 3.A6 shows the effects for alternative binary treatments, used for instance in the graphs assessing the validity of the common trend assumption. Treated and not-treated counties are displayed in Figure 3.A2. These specifications show strong significant negative effects on the growth rate of marginal employment and on the *Aufstocker*. The other two outcomes are not significant in this specification. Table 3.A7 repeats the same exercise, but this time with alternative definitions of the bite, namely with € 1400 or € 2000 as the threshold of monthly gross earnings. Changing the threshold for the bite only mildly affect the estimated coefficients and leaves the detected patterns unchanged.

Table 3.A8 uses a coarser level of aggregation, namely labor market regions instead of counties. The advantage of this approach is that it rests on variation across labor market regions which might be more relevant for the impact of the minimum wage. The regions are defined following Eckey et al. (2007), based on observed commuting patterns. Also this robustness check confirms strong significant negative effects on the growth rate of marginal employment and on the *Aufstocker*, but does not find significant effects for the other two outcomes.

To sum up, the vast majority of robustness checks confirm the existence of negative effects on marginal employment and on the *Aufstocker*. The effect on regular employment, which is quite small in the preferred specification frequently vanishes entirely, if one modifies the estimation strategy. Thus, one can conclude that there are only very small or even no considerable effects on regular employment. The effect on non-working UBII recipients does not appear to be robust.

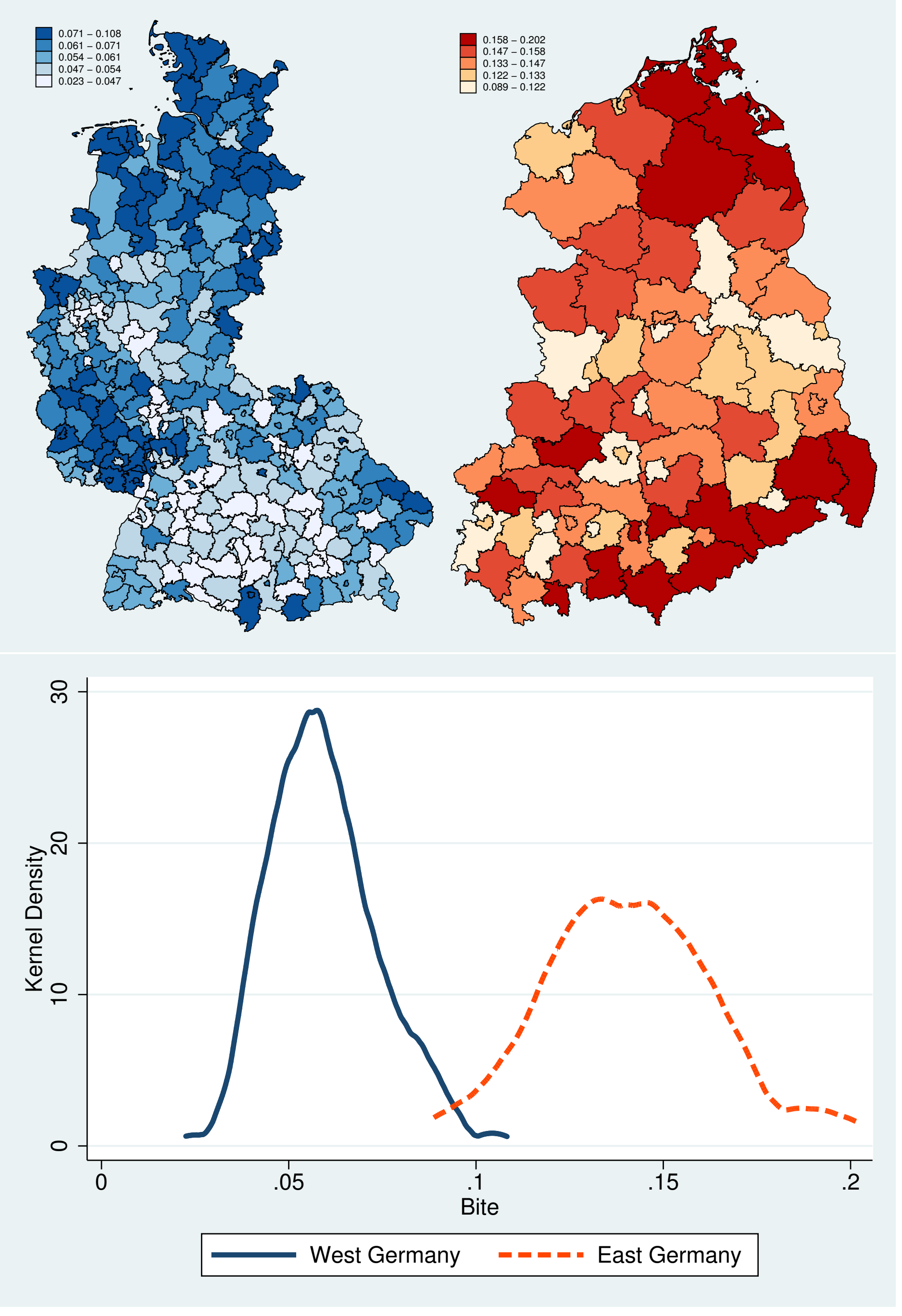
3.5 Discussion

This study examines the effects of the German statutory minimum Wage on employment and welfare dependency, using a difference-in-differences framework. The German labor market remained in a seemingly strong position after the introduction of the minimum wage with no striking *immediate* negative repercussions. However, this study finds evidence for a comparatively large reduction in marginal employment. Concerning regular employment, there is some evidence for an overall small negative effect, even though it does not appear to be very robust. For West Germany, there is some evidence that the loss of marginal employment is offset by conversions into regular employment. In general, the results confirm the findings of previous ex-post studies on the (modest) short-run employment implications of the statutory minimum wage.

Concerning welfare dependency, there is a reduction in the number of the *Aufstocker*, i.e. recipients of unemployment benefit II while working. However, as already argued by Müller and Steiner (2009); Bruckmeier and Wiemers (2014) this effect does not need to imply an improvement in the economic situation of the affected households, since withdrawal rates of the supplementary welfare payments are high. Nevertheless, there might be strong effects on subjective well-being, due to the elimination of welfare stigma (Hetschko et al., 2016) and a partial relief for social spendings. For West Germany there is no indication that the reduction in the growth rate of the *Aufstocker* was caused by them, loosing their job and ending up in non-working welfare receipts. For East Germany on the other hand, there is evidence that a considerable share of the *Aufstocker* did so.

The analysis only considers the *short-run* effects in the first year after the introduction of the minimum wage. Thus, the results cannot give a proper indication of the total effect or the impact of the minimum wage during the next economic recession and recovery. Additionally, the minimum wage might have some harmful medium to long-run effects in strongly affected regions due to location and investment decisions which have yet to take effect. Firms could invest in new machines which are less labor intensive or firms could decide to relocate to other areas due to a change in the relative prices for labor among regions. Nevertheless, it is worthwhile to study the *immediate* effects of the minimum wage directly after its introduction. The short-run loss of about 200,000 jobs in marginal employment is substantial. This finding at least casts some doubts at the sentiment that the minimum wage was free from side effects. Last but not least, if the detected reduction in the growth rate of regular employment turns out to be *permanent*, the resulting long-run effect on employment will be substantial.

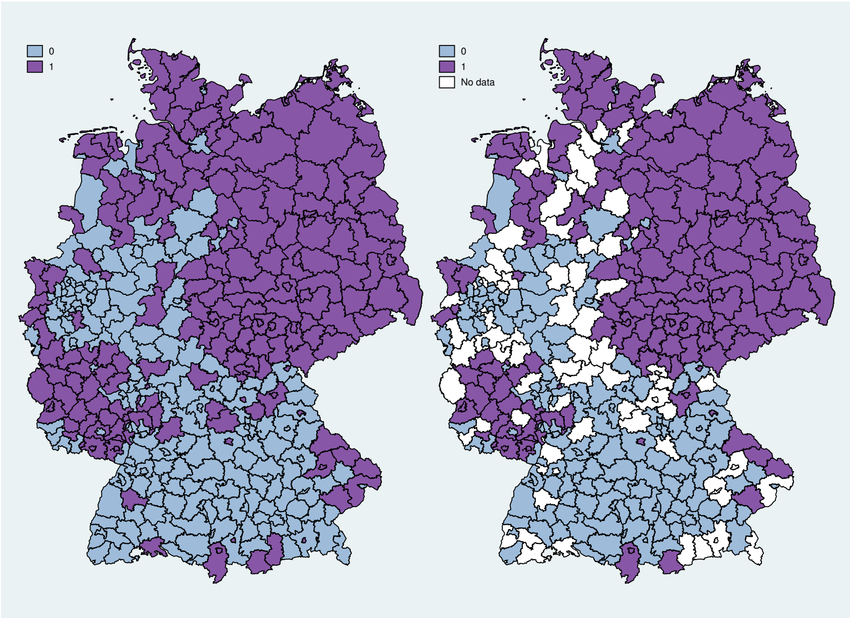
3.6 Appendix



Note: Measured in December 2013

Figure 3.A1: Bite within West and East Germany - Map and Kernel Density

Figure 3.A2: Alternative Binary Treatments



Note: 1: Treatment; 0: Control; Left: Above/below p50; Right: Above p60 - Below p40

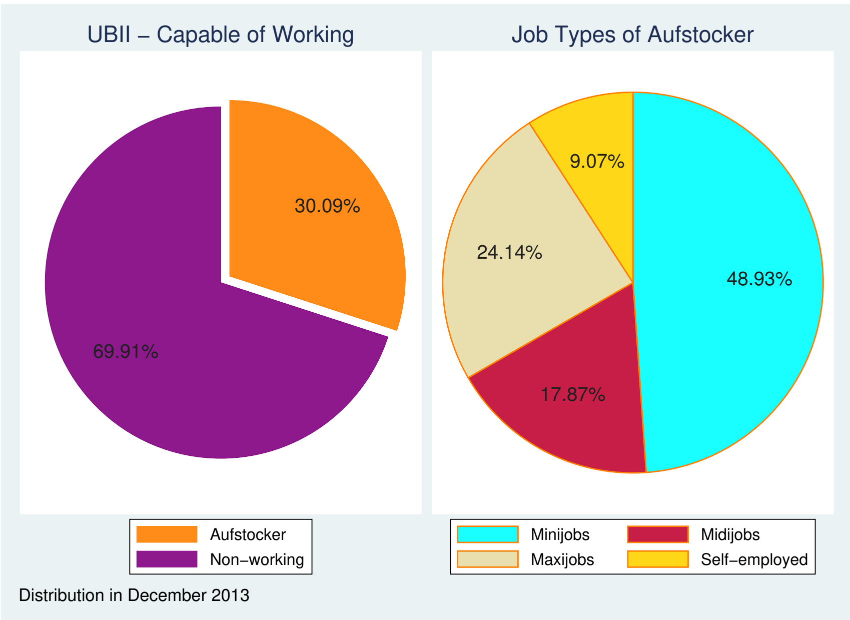


Figure 3.A3: Composition of UBII Recipients and Job Types of *Aufstocker*

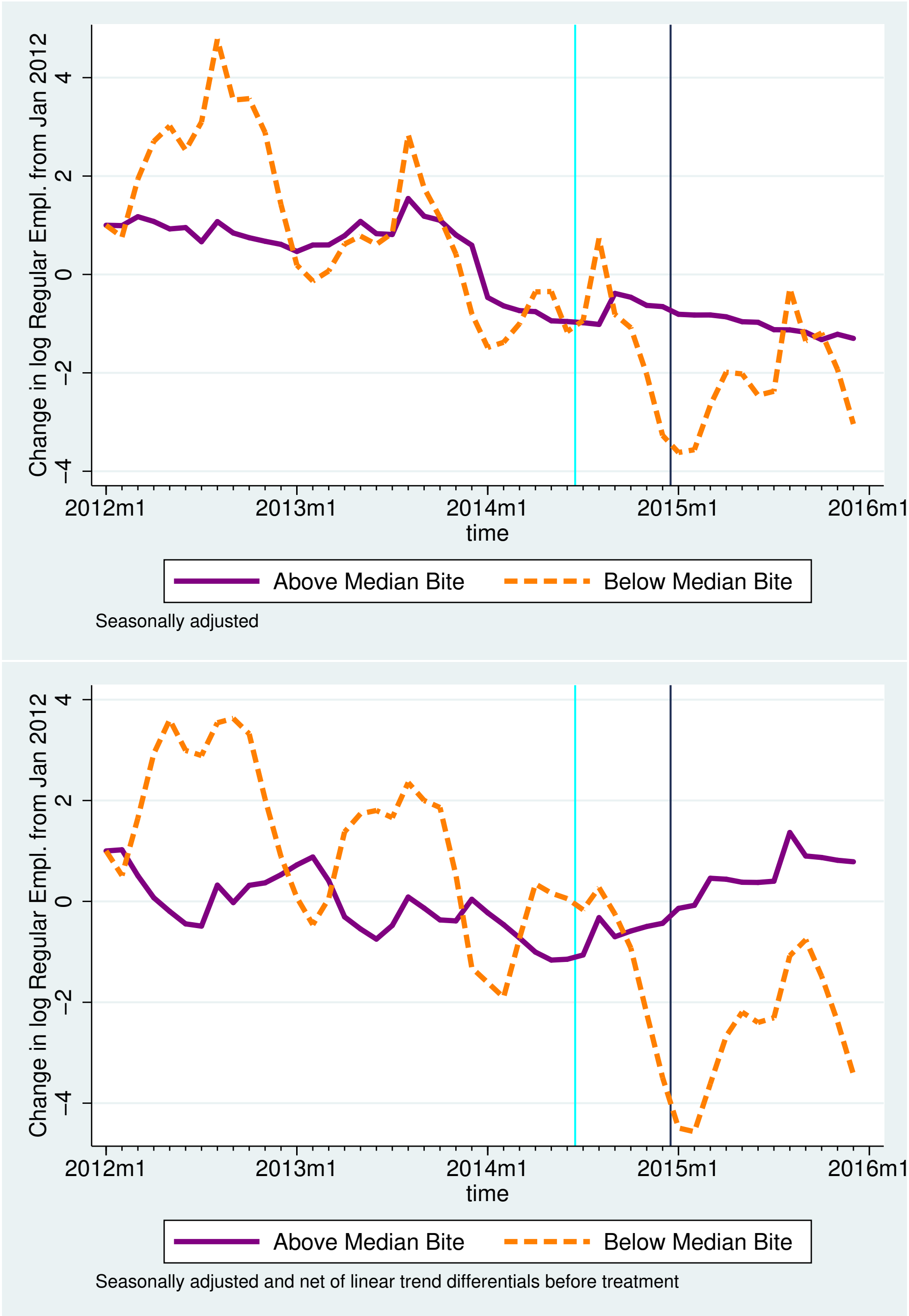


Figure 3.A4: Regular Employment - Deterministic Seasonality, With and Without Linear Trend Differential

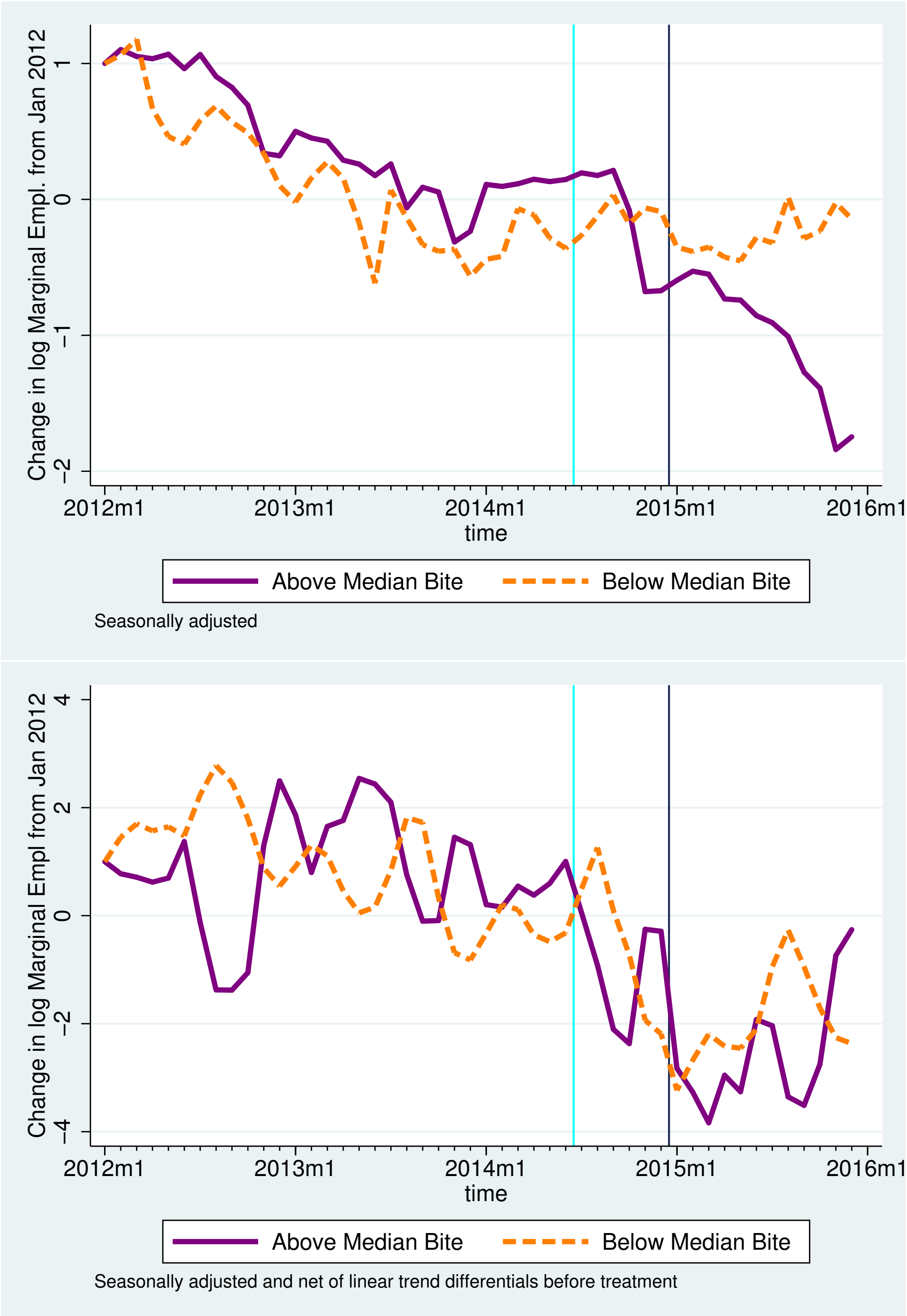


Figure 3.A5: Marginal Employment - Deterministic Seasonality, With and Without Linear Trend Differential

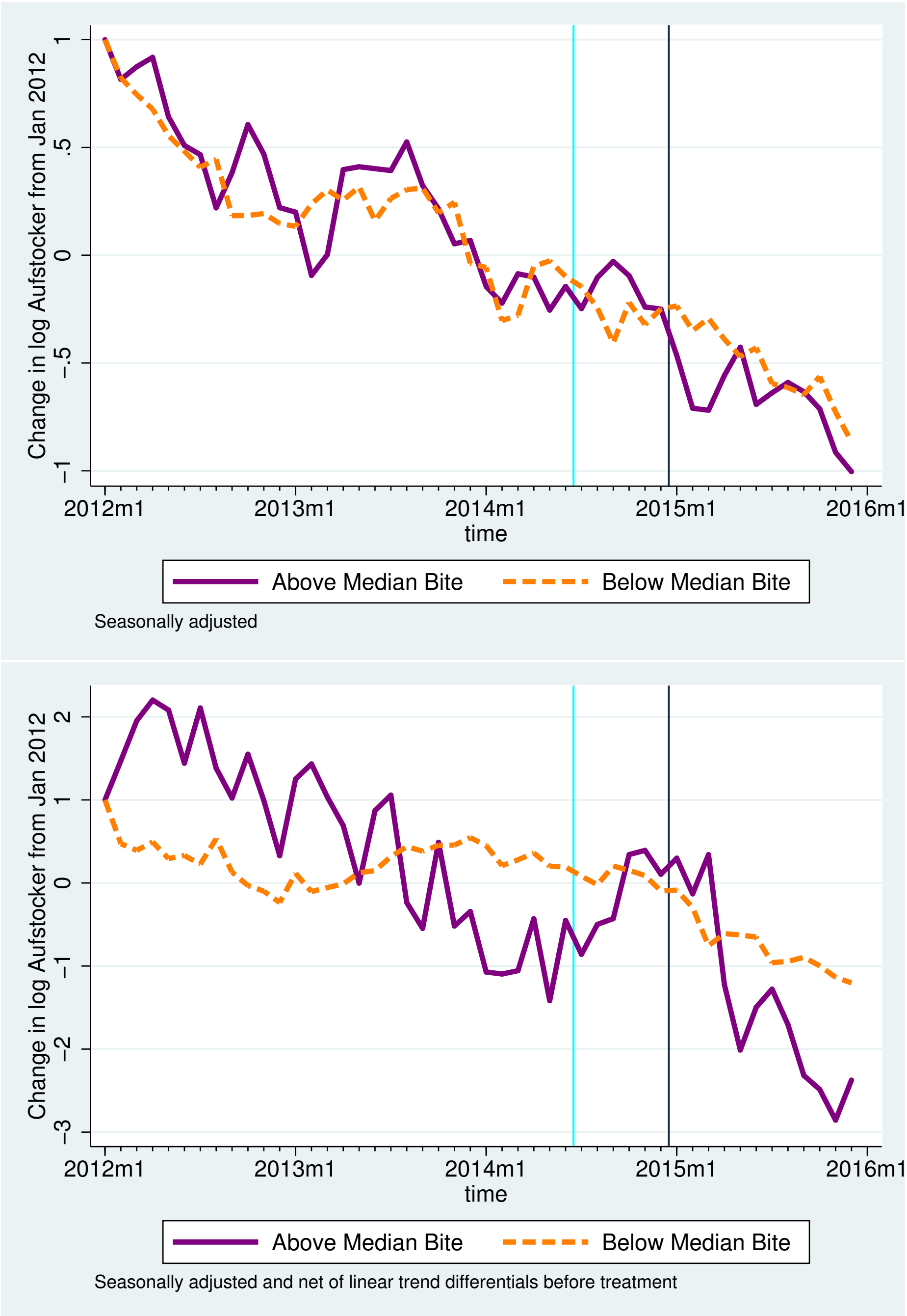


Figure 3.A6: *Aufstocker* - Deterministic Seasonality, With and Without Linear Trend Differential

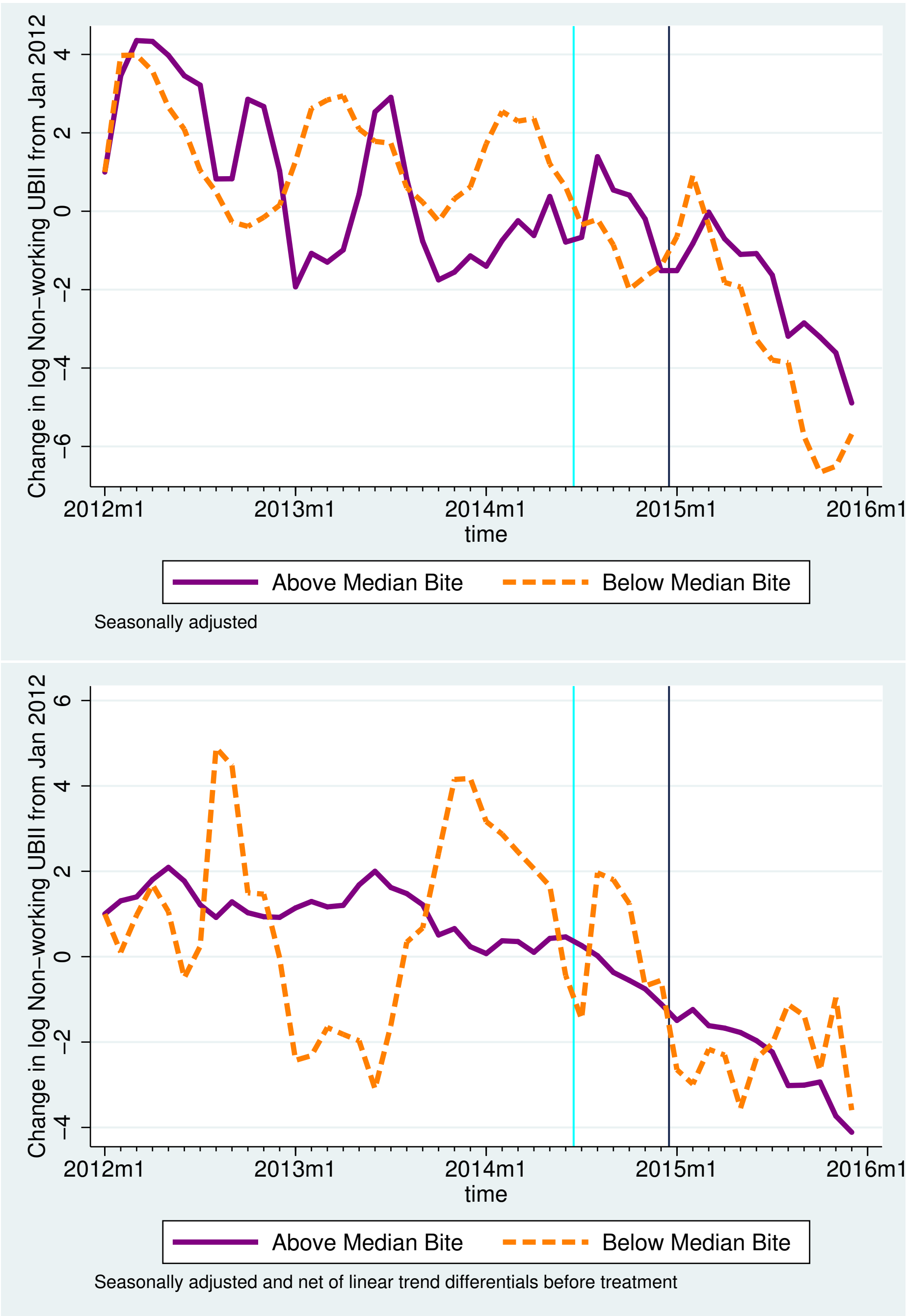


Figure 3.A7: Non-working UBII - Deterministic Seasonality, With and Without Linear Trend Differential

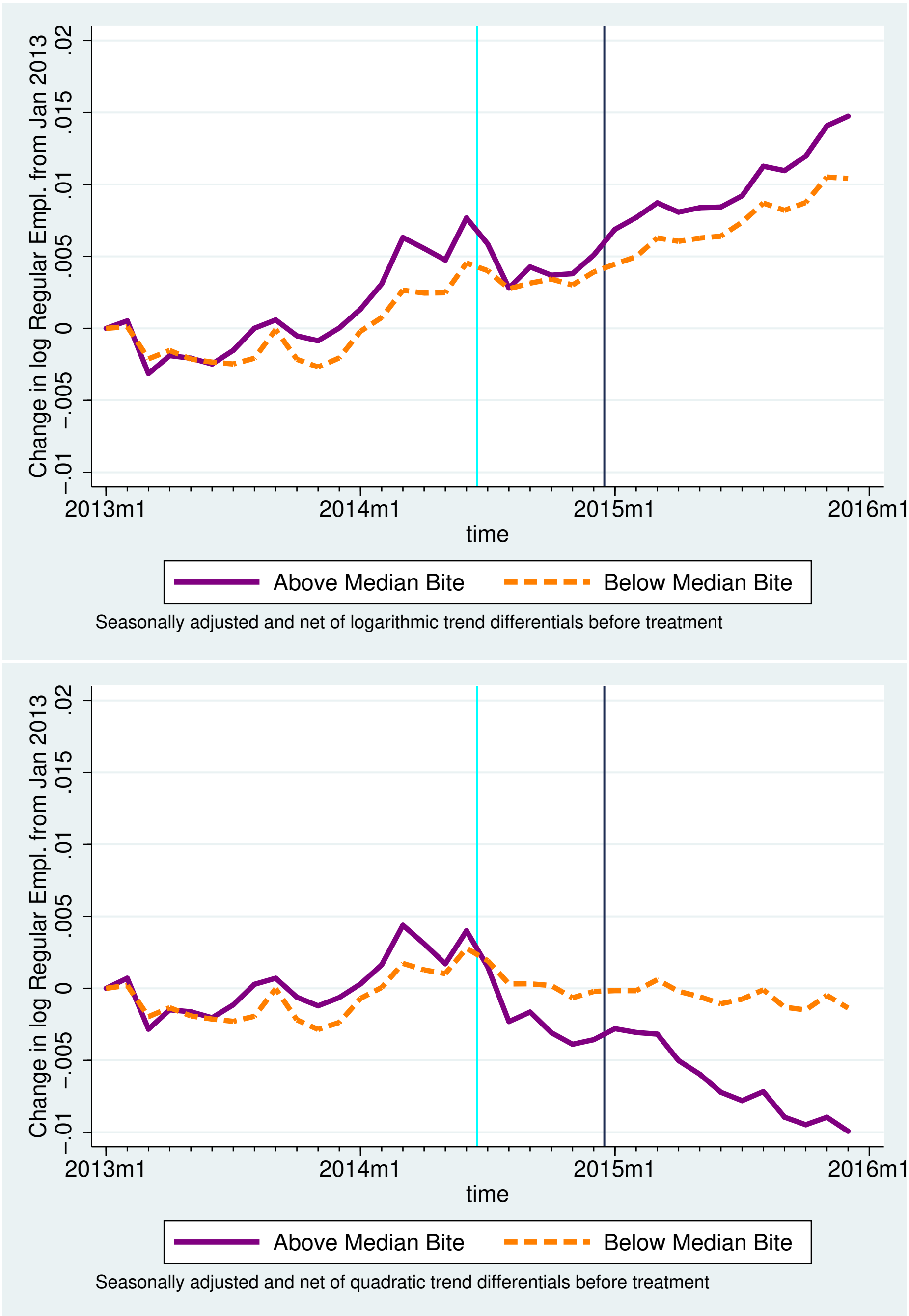


Figure 3.A8: Regular Employment - Further Trend Specifications

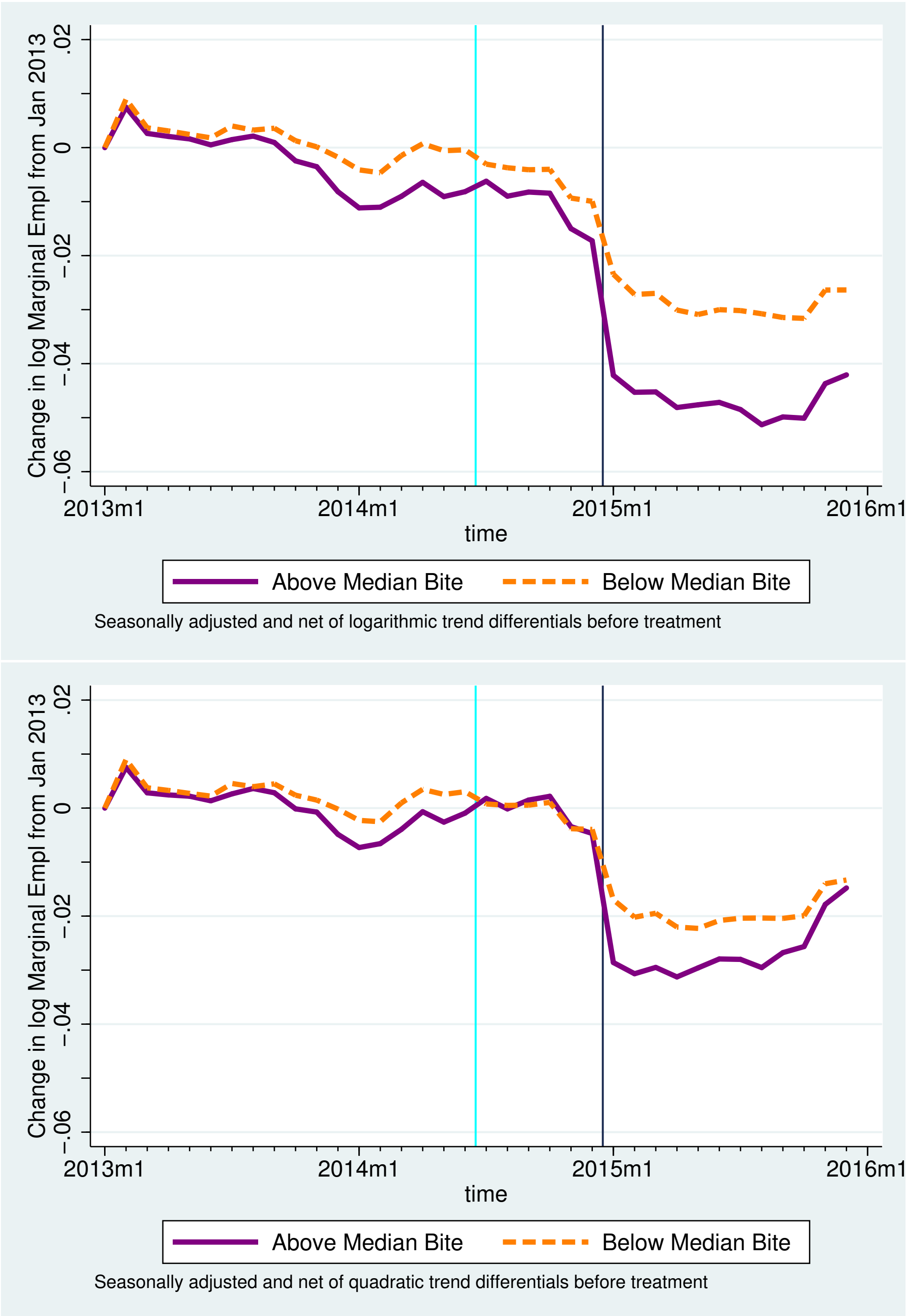


Figure 3.A9: Marginal Employment - Further Trend Specifications

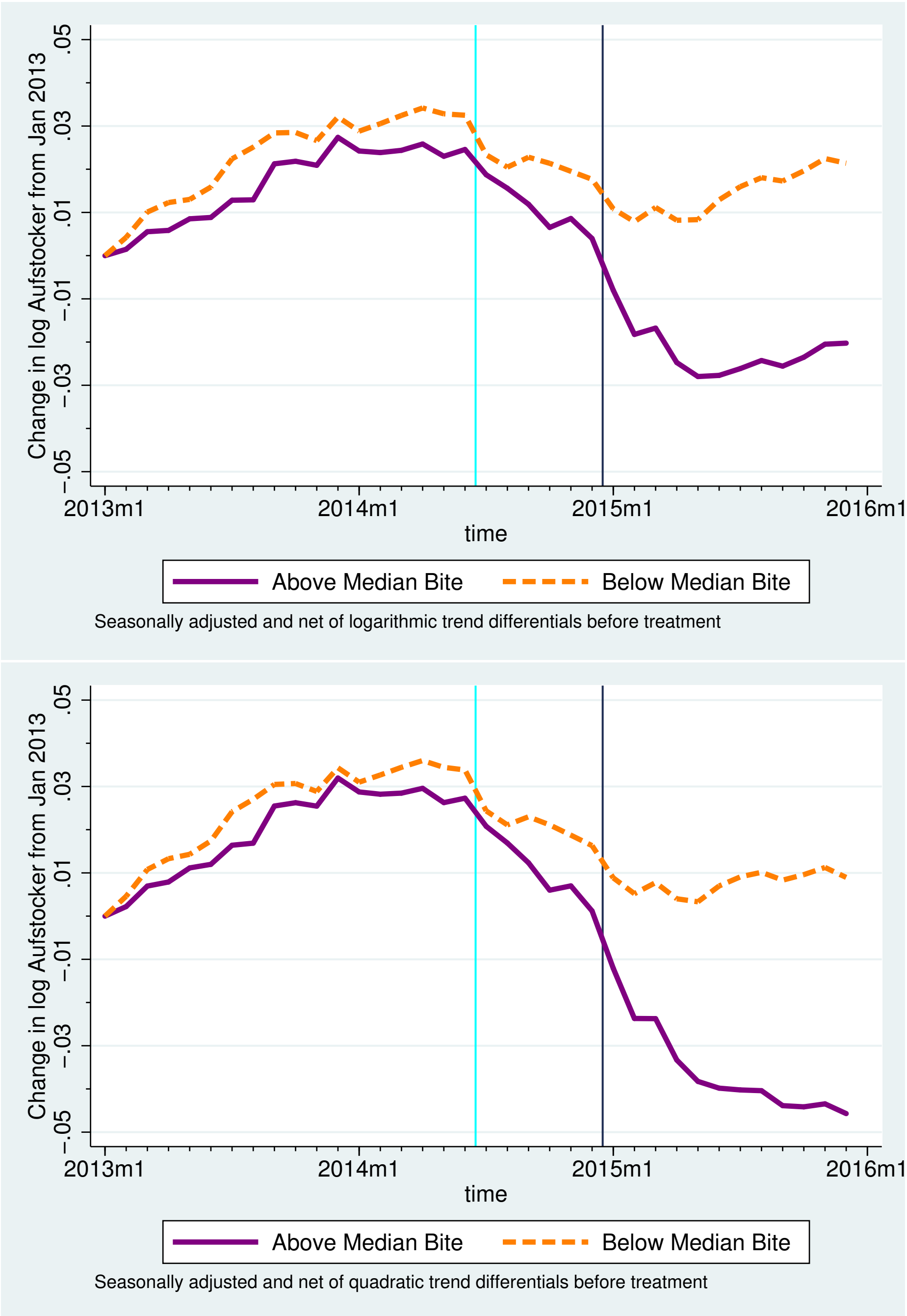
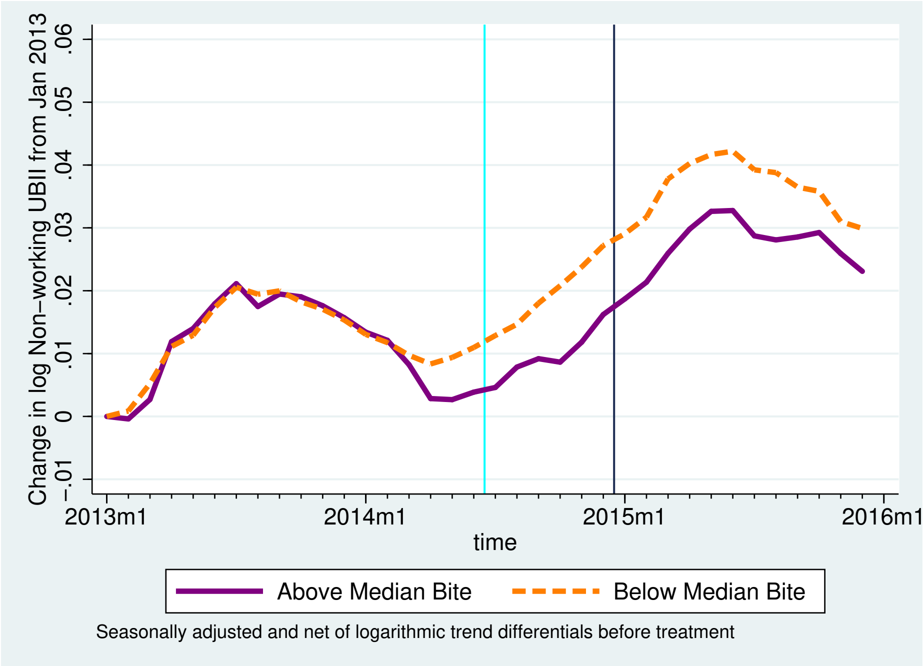


Figure 3.A10: *Aufstocker* - Further Trend Specifications



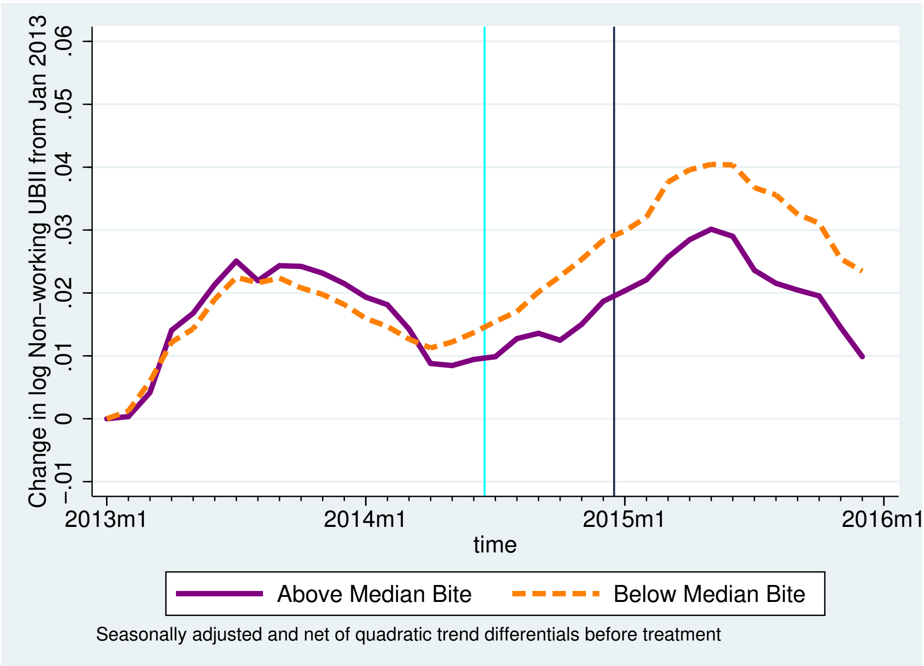


Figure 3.A11: Non-working UBII - Further Trend Specifications

Table 3.A1: Effects on Employment Outcomes - Levels Specification

	(1)	(2)	(3)	(4)
		Regular Employment		
Bite·2015	-0.826***	-0.186***	-0.222***	-0.519***
	(0.079)	(0.048)	(0.030)	(0.050)
$t \times$ Bite		-0.007***	-0.008***	
		(0.001)	(0.002)	
$t^2 \times$ Bite			0.000	
			(0.000)	
$\log(t) \times$ Bite				-0.078***
				(0.012)
Observations	19296	19296	19296	19296
R2 within	0.707	0.718	0.718	0.715
R2 between	0.044	0.044	0.044	0.044
R2 overall	0.003	0.009	0.010	0.014
		Marginal Employment		
Bite·2015	-1.825***	-1.343***	-1.057***	-1.604***
	(0.265)	(0.195)	(0.171)	(0.208)
$t \times$ Bite		-0.006**	0.001	
		(0.003)	(0.004)	
$t^2 \times$ Bite			-0.000***	
			(0.000)	
$\log(t) \times$ Bite				-0.056*
				(0.029)
Observations	19296	19296	19296	19296
R2 within	0.291	0.294	0.295	0.293
R2 between	0.201	0.201	0.201	0.201
R2 overall	0.018	0.045	0.032	0.058

Bite measured in December 2013 and normalized by one SD
Estimation including Time and County Fixed Effects
Cluster-robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A2: Effects on Welfare Outcomes - Levels Specification

	(1)	**(2)**	(3)	(4)
		Aufstocker		
Bite·2015	-6.347***	-2.699***	-1.452***	-4.662***
	(0.234)	(0.149)	(0.107)	(0.172)
$t \times$ Bite		-0.042***	-0.012**	
		(0.003)	(0.005)	
$t^2 \times$ Bite			-0.001***	
			(0.000)	
$\log(t) \times$ Bite				-0.433***
				(0.030)
Observations	19177	19177	19177	19177
R2 within	0.498	0.560	0.564	0.540
R2 between	0.071	0.071	0.071	0.071
R2 overall	0.004	0.027	0.021	0.037
		Non-working UBII		
Bite·2015	-4.069***	-0.650***	0.292***	-2.460***
	(0.203)	(0.113)	(0.096)	(0.145)
$t \times$ Bite		-0.040***	-0.016***	
		(0.002)	(0.003)	
$t^2 \times$ Bite			-0.001***	
			(0.000)	
$\log(t) \times$ Bite				-0.413***
				(0.023)
Observations	19177	19177	19177	19177
R2 within	0.410	0.504	0.507	0.476
R2 between	0.044	0.044	0.044	0.044
R2 overall	0.002	0.020	0.016	0.027

Bite measured in December 2013 and normalized by one SD
Estimation including Time and County Fixed Effects
Cluster-robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A3: Anticipation and Adjustment Effects

	Employment		Welfare	
	Regular	Marginal	*Aufstocker*	NW UBII
	(1)	(2)	(3)	(4)
Bite·07/2014	0.025	-0.145	0.218	0.003
	(0.041)	(0.153)	(0.137)	(0.090)
Bite·08/2014	-0.044	-0.242	0.370*	0.233**
	(0.044)	(0.153)	(0.195)	(0.111)
Bite·09/2014	-0.055	-0.212	0.229	0.196
	(0.047)	(0.155)	(0.230)	(0.130)
Bite·10/2014	-0.119**	-0.235	0.123	0.113
	(0.048)	(0.159)	(0.244)	(0.151)
Bite·11/2014	-0.111**	-0.395*	0.056	0.181
	(0.049)	(0.210)	(0.273)	(0.170)
Bite·12/2014	-0.105**	-0.525**	-0.067	0.293
	(0.052)	(0.215)	(0.281)	(0.186)
Bite·01/2015	-0.065	-1.484***	-0.462	0.340*
	(0.054)	(0.235)	(0.320)	(0.198)
Bite·02/2015	-0.042	-1.479***	-1.040***	0.409*
	(0.057)	(0.241)	(0.343)	(0.218)
Bite·03/2015	-0.077	-1.510***	-1.223***	0.320
	(0.060)	(0.245)	(0.381)	(0.232)
Bite·04/2015	-0.119**	-1.501***	-1.530***	0.460*
	(0.058)	(0.252)	(0.407)	(0.240)
Bite·05/2015	-0.154**	-1.430***	-1.636***	0.579**
	(0.062)	(0.257)	(0.430)	(0.255)
Bite·06/2015	-0.149**	-1.397***	-1.832***	0.607**
	(0.062)	(0.256)	(0.437)	(0.265)
Bite·HY2/2015	-0.126**	-1.398***	-1.664***	0.702**
	(0.058)	(0.223)	(0.474)	(0.291)
Observations	14472	14472	14298	14298
R2 within	0.172	0.308	0.252	0.121
R2 between	0.177	0.110	0.533	0.446
R2 overall	0.105	0.252	0.382	0.297
Linear Trend	-	-	✓	✓
Time Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓

Bite measured in December 2013 and normalized by one SD

Cluster-robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A4: Placebo

	Employment		Welfare	
	Regular	Marginal	*Aufstocker*	NW UBII
	(1)	(2)	(3)	(4)
Bite· 2014	0.072	-0.329***	0.222	0.084
	(0.049)	(0.125)	(0.146)	(0.085)
Observations	9648	9648	9533	9533
R2 within	0.092	0.039	0.078	0.109
R2 between	0.128	0.012	0.322	0.356
R2 overall	0.021	0.028	0.197	0.253
Linear Trend	-	-	✓	✓
Time Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓

Bite measured in December 2013 and normalized by one SD
Cluster-robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A5: Variation within Gender

	Employment		Welfare	
	Regular	Marginal	*Aufstocker*	NW UBII
	(1)	(2)	(3)	(4)
		Men		
Bite·2015	0.013	-1.184***	-0.858**	0.050
	(0.078)	(0.293)	(0.334)	(0.195)
Observations	14472	14472	14298	14298
R2 within	0.116	0.258	0.103	0.067
R2 between	0.103	0.121	0.394	0.393
R2 overall	0.060	0.214	0.189	0.160
		Women		
Bite·2015	-0.050	-1.301***	-1.418***	0.388***
	(0.051)	(0.181)	(0.230)	(0.140)
Observations	14472	14472	14297	14297
R2 within	0.203	0.277	0.198	0.055
R2 between	0.104	0.085	0.426	0.347
R2 overall	0.104	0.220	0.288	0.198
		Gender Cells		
Bite·2015	0.035	-0.968***	-1.608***	-0.445***
	(0.044)	(0.166)	(0.211)	(0.122)
Observations	28944	28944	28595	28595
R2 within	0.132	0.258	0.107	0.053
R2 between	0.017	0.200	0.211	0.274
R2 overall	0.062	0.211	0.140	0.126
Linear Trend	-	-	✓	✓
Time Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓

Bite measured in December 2013 and normalized by one SD
Cluster-robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A6: Binary Treatment

	Employment		Welfare	
	Regular	Marginal	*Aufstocker*	NW UBII
	(1)	(2)	(3)	(4)
	Binary Treatment			
Treated·2015	0.022	-1.320***	-1.398***	0.117
	(0.101)	(0.383)	(0.492)	(0.349)
Observations	14472	14472	14298	14298
R2 within	0.168	0.276	0.177	0.107
R2 between	0.050	0.032	0.235	0.173
R2 overall	0.080	0.207	0.204	0.142
	Robust Binary Treatment			
Treated·2015	0.037	-1.932***	-1.711***	-0.346
	(0.117)	(0.392)	(0.549)	(0.400)
Observations	11592	11592	11455	11455
R2 within	0.156	0.313	0.219	0.112
R2 between	0.072	0.084	0.289	0.239
R2 overall	0.072	0.247	0.253	0.177
Linear Trend	-	-	✓	✓
Time Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓

Bite measured in December 2013 and normalized by one SD
Cluster-robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A7: Alternative Bite Definitions

	Employment		Welfare	
	Regular	Marginal	*Aufstocker*	NW UBII
	(1)	(2)	(3)	(4)
		Bite 1500		
Bite·2015	-0.098*	-1.371***	-1.463***	0.334**
	(0.050)	(0.219)	(0.206)	(0.138)
Observations	14472	14472	14298	14298
R2 within	0.171	0.307	0.249	0.121
R2 between	0.182	0.114	0.546	0.456
R2 overall	0.097	0.251	0.388	0.301
		Bite 2000		
Bite·2015	-0.126**	-1.381***	-1.495***	0.343**
	(0.050)	(0.253)	(0.202)	(0.137)
Observations	14472	14472	14298	14298
R2 within	0.172	0.308	0.247	0.123
R2 between	0.168	0.102	0.571	0.479
R2 overall	0.102	0.249	0.397	0.314
Linear Trend	-	-	✓	✓
Time Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓

Bite measured in December 2013 and normalized by one SD

Cluster-robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A8: Labor Market Regions

	Employment		Welfare	
	Regular	Marginal	*Aufstocker*	NW UBII
	(1)	(2)	(3)	(4)
Bite·2015	-0.044	-1.799***	-1.689***	0.211
	(0.097)	(0.457)	(0.253)	(0.225)
Observations	5076	5076	5033	5033
R2 within	0.191	0.283	0.316	0.205
R2 between	0.122	0.141	0.458	0.355
R2 overall	0.106	0.239	0.391	0.278
Linear Trend	-	-	✓	✓
Time Fixed Effects	✓	✓	✓	✓
Labor Market Region FE	✓	✓	✓	✓

Bite measured in December 2013 and normalized by one SD
Cluster-robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4 Congestion Pricing: A Mechanism Design Approach

4.1 Introduction

Congestion is an ever-present nuisance to many people living in the world's urban areas and has significant economic costs. Congestion adds about 5 cents per mile to the cost of travel, making it the single largest source of traffic-related externalities (Parry et al., 2007).[1] Schrank et al. (2015) estimate delay and fuel costs of $160bn due to congestion for the USA in 2014. From an economist's point of view, congestion is an externality problem since an individual driver does not take the effect of her journey on other drivers' travel time into account.

The classic solution is a Pigouvian tax set at a level to ensure that each driver internalizes the marginal cost of the increased travel time of other drivers. Much of the existing literature on congestion pricing is concerned with finding the right level of such a congestion charge given observable characteristics.[2] In practice this requires the regulator to have a reliable estimate of the value of travel time. Generally one would expect these values to differ substantially across drivers. Indeed Small et al. (2005) as well as Steimetz and Brownstone (2005) have empirically demonstrated the existence of substantial heterogeneity in the value of time from both observed and unobserved sources. This prevalence of unobserved heterogeneity suggests that traffic regulation could be improved if one could find a way to observe valuations of travel time directly, without trying to estimate them from observable characteristics.

In this paper we propose an application of mechanism design as a novel approach to deal with demand uncertainty in traffic regulation. The literature on mechanism design has already provided many insights on practical problems where agents have private information. Examples of these include the sale of items (Myerson, 1981) where agents have private information on their valuations of the item; or the provision of public goods (Clarke, 1971) where agents have private information about their own valuations of the public good. This includes cases where agents have private information about their costs for reducing emissions (Montero, 2008). Using mechanism design allows us to target three important issues. First, we can analyze the implications of the regulator's lack of detailed knowledge concerning the value of time. Second, we can use more general payment schemes that do not impose specific functional forms. Third, we obtain a

[1] Table 2 in Parry et al. (2007) summarizes the costs of traffic-related externalities. The estimated external cost of local pollution caused by automobiles is 2 cents per mile and 3 cents per mile for accident-related externalities. Externalities of greater greenhouse gas emissions is 0.3 cents per mile.

[2] Some examples of observable characteristics are time of day, income, measures of traffic density etc.

better understanding about the conditions under which a Pigouvian congestion tax is appropriate to deal with traffic congestion.

In our model each driver has private information regarding their value of time. We study a route-choice problem and illustrate our results using a simple origin-destination model similar to Mayet and Hansen (2000) in which travelers can use one of two roads. In the main text we focus on the simpler case in which only one of the two roads may be congested.[3] The externality of a driver on the others arises from slowing down other drivers on the congested road. A mechanism designer, who can be seen as either a local government authority, a regulator, or a private company administering the roads, can design mechanisms involving pricing schemes and allocations to implement efficient road usage. We also consider a mechanism designer who wants to maximize revenue in an extension of our model.

In addition to road congestion, this setup also captures congestion in other applications. Container ships that travel from Shanghai to New York can either use the shorter route through the Panama Canal or the longer one around Cape Horn. Also in air travel, airlines face congested hubs that make delays or missed connection flights more likely.

We characterize the efficient allocation in this model. For a finite number of drivers, the efficient allocation depends on the *realized* values of time. Since these realized values are ex-ante unknown to the traffic regulator, also the efficient allocation is ex-ante uncertain. The regulator asks each driver to report their value of time and then allocates drivers efficiently to different roads. We derive a payment scheme which ensures truthful reports of drivers' valuations *and* implements the efficient allocation. This payment scheme makes sure that drivers have the right incentives to report their values of travel time truthfully, and—with these information—implements the efficient allocation. This payment scheme is an application of the well-known Vickrey-Clarke-Groves (VCG) mechanism[4] and reflects the externality that the presence of the driver imposes on others.

Intuitively, our mechanism differs from a classic Pigouvian price in the following way: a Pigouvian price allows drivers to pay for the fast road. If, however, many drivers have sufficiently large values of time, they will pay this price but the resulting allocation makes the fast road inefficiently slow. Conversely, when values of time are relatively low for most drivers, a single Pigouvian price that does not adjust immediately to these realizations leads to an underused road which is inefficiently fast. This illustrates that the efficient allocation depends on the realized values of time. Our mechanism deals with these issues in two ways. First, in addition to paying for usage of the fast road, our mechanism allows drivers to reduce travel time on the fast road by making additional payments.

[3] Note, however, that the uncongested alternative does not need to be a road. It could also represent a network of alternative travel options, including several roads or public transport. We also study the case with two congestible roads at a later point.

[4] This mechanism is inspired by the seminal works of Vickrey (1961), Clarke (1971) and Groves (1973). VCG-type mechanisms are used in practice for example by facebook to sell advertisement space in their AdAuction system, or in ebay auctions.

Second, if many drivers have relatively low values of travel time, our mechanism learns about that by the drivers' reports and adjusts the price to enter the fast road, inducing efficient road usage in this case as well. In equilibrium, reporting your value of travel time is analogous to deciding out of a menu of desired travel time-price combinations. We will illustrate this in some examples below.

Our procedure opens the possibility of using second-degree price discrimination *within* a congestible road to extract drivers' value of time.[5] Even if the regulator does not have any prior knowledge about the distribution of the values of time, our mechanism instantaneously provides the real time valuations of drivers. Thus, compared to using estimated values of time our approach is flexible and adaptive. Regulators can deal with aggregate uncertainty and can respond quickly to changes in demand.

This result has important implications for the optimal functional form of congestion pricing. We show that even when the regulator knows the exact distribution of the values of travel time, a single Pigouvian tax does not generally implement the efficient allocations when the number of drivers is finite. Previous models of congestion pricing took as granted a given pricing structure for the regulator.[6] Each driver faced a price that could depend on observables, such as which road was used, time of the day, type of car, current or previous speed-flow on the road etc. However, unobservable characteristics related to the values of travel time were not taken into account. Applying mechanism design allows us to show that this pricing structure is not optimal in general. We show that with a finite number of drivers, more sophisticated payment schemes are required to implement the efficient allocation. Therefore the lack of knowledge concerning drivers' value of time does not impede the implementation of efficient allocations.

We also derive conditions under which the welfare loss of using a single Pigouvian price becomes negligible. In this way, we contribute to understanding the conditions under which a Pigouvian congestion tax is appropriate, and when it could be improved upon. A single congestion charge can implement efficient traffic if the regulator already knows the efficient allocation of traffic. However, if she is uncertain about the efficient allocation, she could use a more complex price schedule, allowing her to learn about the efficient allocation while implementing it at the same time.

Our mechanism might have seemed impractical in the past as there was no way to easily communicate the value of time to a central decision-maker. There was also no way for the decision maker to force drivers to take particular routes. Modern communication

[5]The literature already considers price discrimination between roads. For an example see Mayet and Hansen (2000). We show that this approach is generally not sufficient to implement efficient allocations between roads.

[6]As some examples that will be discussed in detail below, consider Vickrey and Sharp (1968), Vickrey (1969), Bernstein and El Sanhouri (1994), Verhoef et al. (1996), Arnott and Kraus (1998) Mayet and Hansen (2000), Verhoef and Small (2004).

technology, such as smart phones and GPS, and the advent of self-driving cars imply that these practical problems may soon be overcome.[7]

Instead of estimating demand, regulators could use trial-and-error methods proposed for instance by Vickrey (1993); Downs (1993). Intuitively, these methods use observable, realized speed-flow relationships to adjust existing pricing schemes. This leads to a trial-and-error algorithm that converges to socially optimal congestion charge.[8] Our mechanism is designed to enable even an uninformed traffic regulator to implement road pricing using price discrimination to elicit reports of drivers about their value of travel time. In this way, our mechanism is in principle very flexible and able to adjust quickly to changes in traffic demand.

The following section 4.2 discusses previous papers on congestion pricing and mechanism design problems with externalities. Section 4.3 introduces our basic model and solves for the efficient allocation. Section 4.4 derives a payment schedule that implements the efficient allocation and considers the limit case when the number of drivers becomes large. We also discuss simple examples that include actual price schedules for drivers. Section 4.5 studies the case when the mechanism designer maximizes revenue rather than welfare. Section 4.6 extends our results to the case when there are two congestible roads. Section 4.7 discusses currently used congestion pricing schemes in light of our results. Section 4.8 discusses congestion problems arising in the other environments and concludes.

4.2 Literature

Up to our knowledge, we are the first who study congestion externalities as a mechanism design problem. The existing literature on road congestion either assumes Pigouvian taxes or focuses on cases with infinitely many "small" drivers where Pigouvian taxes are efficient. We establish that Pigouvian prices are not efficient in the presence of aggregate uncertainty, i.e. the case with a finite number of drivers. We characterize efficient payment schemes for these cases.

The idea to use prices to implement efficient road usage dates back to Pigou (1920) and Knight (1924) and later gained popularity among economists. Vickrey and Sharp (Vickrey and Sharp, 1968; Vickrey, 1969) are still regarded as the founding fathers of transport economic theory (Verhoef, 2000). The central idea of this literature, the *Pigouvian approach*, is to implement efficient road usage by internalizing the social cost of

[7]With self-driving cars travelers could simply indicate their desired destination and their desire for reaching it on time. A self-driving car would electronically transmit this information to a central authority, which calculates an efficient travel schedule for the self-driving cars. Singapore for instance is already experimenting with self-driving cabs (Land Transport Authority, 2016).

[8]For a survey on the existing literature see Yang et al. (2005), Tsekeris and Voß (2009), or de Palma and Lindsey (2011)

congestion via a tax or price regulation: the surcharge for road usage is set to equate the marginal social cost at the efficient level.

Subsequent empirical and theoretical research has identified several problems with this approach, namely information requirements and the users' heterogeneity in their value of travel time. It has been shown that the value of travel time varies in the course of the day and hence also the demand for road usage; moreover, there is considerable heterogeneity across users (for an overview, see Small, 2012). Based on data on the use of pay lanes Small et al. (2005) estimate the distribution of the value of time of commuters in California choosing between a tolled express lane and a free alternative lane. This study finds a median value of time of around $23 with substantial heterogeneity unrelated to observable factors. Steimetz and Brownstone (2005) use commuters' choices on the California Interstate 15 to characterize the heterogeneity in the value of time by observable characteristics. They find that while the mean value of time is $30 per hour, the value of time of different drivers ranges between $7 and $65 per hour.[9]

Several theoretical papers analyze the congestion pricing problem assuming a commonly known value of travel time, identical across all drivers. For example, Bernstein and El Sanhouri (1994) and Verhoef et al. (1996) analyze the problem of optimally setting congestion charges in a network with two roads, in which only one of the roads can be tolled. The value of travel time is implicitly normalized to unity for all drivers. The heterogeneity present in those papers concerns mainly the overall value of a trip. In our paper, in contrast, the heterogeneity of drivers concerns the value of travel time.

Another strand of the literature has studied optimal congestion pricing when there is heterogeneity in the value of time. Closely related to our paper is Mayet and Hansen (2000), who also consider a model in which there are two roads, only one of which may be congested. Like in our model the heterogeneity of drivers concerns valuation of travel time, rather than the value of a trip. However they restrict the regulator to setting a single toll for using the congestible road. Small and Yan (2001) consider a model in which there are only two types of drivers, one with a high value of time and another with a low value of time. They highlight that because of the heterogeneity, there is some welfare gain from having roads with different travel times, as drivers with a high value of time will be willing to pay more to reach their destination faster. Verhoef and Small (2004) also compare the social optimum to the congestion charges chosen by a private, profit-maximizing road operator. Arnott et al. (1994) analyze the choice of an optimal time-varying toll in a model with a heterogeneous value of travel time and random departure times. The number of drivers of each type in this model is known ex ante. Arnott and Kraus (1998) distinguish between anonymous and non-anonymous congestion charges and investigate under which conditions an anonymous congestion charge is optimal, when drivers can have varying values of time. Unlike our model, the departure times may vary across drivers.

[9] Some empirical research also studies the value of reliability of time. See Concas and Kolpakov (2009) for an overview.

One feature of those papers is that they assume simple Pigouvian taxes or infinitely many "small" drivers. In contrast, we study explicitly the role of aggregate uncertainty over the number of drivers on each road and over the optimal level of road usage. When setting a single fixed congestion charge the mass of drivers using a road is precisely determined. Thus there is no role for the optimal pricing scheme in eliciting information on what the optimal level of road usage is. The mechanism only determines which driver uses which road. In contrast, in our mechanism the reports by the drivers will also determine the optimal number of drivers on each road. In the limit of our model, aggregate uncertainty disappears so that we also recover the optimality of a single congestion charge. In general this single congestion charge (for each observable type) is not optimal. This point has not been recognized in the earlier literature.

Most of the aforementioned models on congestion and ours study a route choice problem. Another strand of the traffic-congestion literature, such as Vickrey (1967) or Arnott et al. (1990) considers a different source of congestion in a bottleneck problem. In these models congestion pricing aims at influencing drivers' departure time to regulate traffic.

Methodically, we build on the literature on mechanism design. Many of these papers focus on auctions in which buyers have private values for the items to be sold or look at the optimal provision of public goods such as Vickrey (1961), Clarke (1971), Groves (1973) or Myerson (1981).

Jehiel et al. (1996) study a single unit auction in which a buyer is privately informed about the payoff received by other buyers when she is assigned the item. Jehiel et al. (1999) study a similar single unit auction in which a buyer has private information about her own payoff from owning the item as well as from others owning the item.[10] The paper is therefore closer to ours, in the sense that a driver in our setting has private information about her valuation when another driver is added to a road. In both papers, agents cannot escape the externality, i.e. even agents who do not participate in the auction suffer (or benefit) from it. The seller can then threaten agents who do not receive the good with allocating it to an agent that causes a negative externality. In this way, the seller can induce payments by agents that do not receive the good. In our setting however, the externality only affects drivers on the congestible road such that only drivers assigned to this road make payments to the mechanism designer. Another difference of our paper is that we do not restrict the mechanism designer to selling a single good. In principle, each agent can be assigned to the congestible road.

VCG-type mechanims have also been used to study efficient solutions to environmental externalities (Montero, 2008). He looks at the problem of emissions abatement where polluters are privately informed about their cost of abatement. Traditionally proposed solutions, such as a tax on emissions or an emissions trading scheme are not efficient

[10]Haghpanah et al. (2013) study a more general auction where externalities are common knowledge and described by a social network. The networks they consider are more general, but they focus on positive externalities and on methods to approximate solutions in this problem.

mechanisms in this context. Montero (2008) proposes instead a VCG-type mechanism to give polluters an incentive to report their cost of abatement truthfully and to implement the efficient level of abatement.

4.3 Model

There are n drivers that simultaneously want to reach some common destination D, starting from a common starting point O, and a mechanism designer. The mechanism designer can represent a local government or a toll authority. Drivers can take one of two roads A or B. Road A is in principle the faster road but becomes congested as more drivers use it. Road B is an uncongestible alternative, for instance a bypass road or public transport.[11] Assuming an uncongested alternative simplifies the exposition of results but is not crucial for the results, as we will show in an extension in Section 4.6. On road A the travel time increases with the number of drivers.

We assume that drivers differ by the value they attach to the time spent traveling. This information is summarized for each driver i in the parameter $\theta_i \in \Theta_i \subseteq \mathbb{R}_+ \setminus \{0\}$. For some results we assume additionally that all θ_i are independently and identically distributed according to the well-behaved cumulative distribution function $F(\theta_i)$[12]. For most of our analysis the assumption on the distribution of the value of time $F(\cdot)$ is not necessary as the efficient mechanism induces the revelation of each drivers' value of time independent of distributional assumptions. However the distribution of drivers' value of time is needed when we consider limit cases and revenue maximization. We let $\theta \in \Theta \equiv \times_i \Theta_i$ be the n-dimensional vector of all drivers' valuation of time. We assume that for all i, j, $\theta_i \neq \theta_j$. Given well-behaved distribution functions, this case is expected to occur with certainty. We will denote by θ_{-i} the vector of all valuations except that of driver i.[13]

Since there is no congestion on road B we normalize the utility of a driver who travels on road B to be zero, excluding the transfer p, i.e. $u_i^B(\theta_i, k, p) = -p$. The utility of a driver who travels on road A is given by:

$$u_i^A(\theta_i, k, p) = v(\theta_i, k) - p$$

where k is the total number of drivers, including i, traveling on the fast road and p is a transfer payment to the mechanism designer. We assume that for all θ_i and $k \leq n$, $v_{\theta_i} > 0$, that is $v(\cdot,\cdot)$ increases in θ.[14] This implies that drivers with a higher value of time

[11]Another interpretation is that a driver may decide not to travel at all. We will use this interpretation for the discussion of revenue maximization.

[12]Well-behaved means that $F(\cdot)$ is continuously differentiable with a strictly positive derivative $f(\cdot)$ and such that $\frac{1-F(\theta_i)}{f(\theta_i)}$ is weakly decreasing in θ_i.

[13]With a slight abuse of notation we denote by $F(\theta)$ the joint distribution of the vector of valuations and $F(\theta_{-i})$ the joint distribution of all valuations except for that of driver i.

[14]This formulation allows for utility to be non-linear in the valuation of travel time.

value the trip from O to D more highly. We further assume that there is congestion on road A, in the sense that more other drivers using the fast road reduce the value a driver obtains from the trip, i.e. for all θ_i and $k \leq n$, $v(\theta_i, k+1) - v(\theta_i, k) \leq 0$. Additionally, to interpret higher values of θ_i as a greater value of time, we assume that for drivers with a higher value of time the impact of congestion is greater. This means that for all θ_i and $k \leq n$, $v_{\theta_i}(\theta_i, k+1) - v_{\theta_i}(\theta_i, k) \leq 0$. Finally we assume for all θ_i, $k \leq n$ and $k' < k$, $v(\theta_i, k+1) - v(\theta_i, k) \leq v(\theta_i, k'+1) - v(\theta_i, k')$. This means that the increase in travel time caused by congestion increases in the level of congestion.[15]

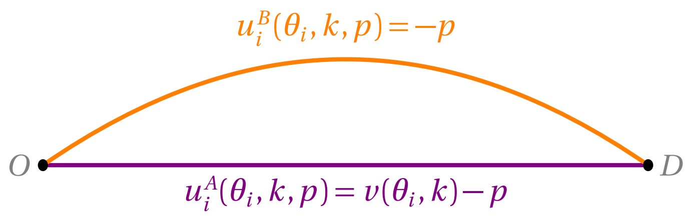


Figure 4.1: A Congestible and an Uncongestible Road

An underlying assumption of the congestion function is that drivers do not care about the identity of other drivers on road A, but only about their number. We believe this is a natural assumption in the context of transportation, and it simplifies the externality problem significantly.

To make the problem interesting we assume that for all θ_i, $v(\theta_i, 1) > 0$ and $v(\theta_i, n) < 0$ so that it is always optimal to have a positive number of drivers, strictly less than n, using road A. We will refer to road A as the fast road, while we call road B, which will yield a lower utility in equilibrium, the slow road. The allocation of agent i is given by $x_i \in \{0, 1\}$, where $x_i = 1$ means that i is allocated to the fast road, while $x_i = 0$ means that i uses the slow road. We let $x \in X = \{0, 1\}^n$ denote the overall allocation. The number of drivers using the fast road is therefore given by $\sum_i x_i$.

Ultimately drivers should only care about their travel time, but not about the number of other drivers. Because of congestion effects, the number of drivers on a road directly affects the travel time on that road. Therefore we can derive the reduced-form utility function specified above from an underlying model in which drivers only care about travel time. Too see this, suppose that $\tilde{v}(\theta_i, t)$ captures the utility to driver i from traveling from O to D in t units of time. By assumption the slow road B is not congestible, implying that the travel time on B is constant. If the time it takes to travel on road A is given by

[15]Congestion pricing is most relevant for medium levels of congestion where traffic is still flowing but at a lower speed. The model captures these situations. For larger levels of congestion, traffic comes to a standstill. While these situations are clearly important, a congestion-pricing scheme cannot resolve them, but only prevent them.

$C(k)$ then we can derive $v(\theta_i, k)$ as $\tilde{v}(\theta_i, C(k))$.[16] Since more drivers on the fast road A imply a longer travel time, we will often refer to travel time when discussing the effects of congestion.

The mechanism designer maximizes total welfare, i.e. the sum of all drivers' utility plus the total revenue collected.[17] Hence the objective function of the mechanism designer is given by:

$$W = max_{[x,p]} \sum_{i=1}^{n} \left[x_i u_i \left(\theta_i, \sum_{j=1}^{n} x_j, p_i \right) + p_i \right] \tag{4.1}$$

In the absence of a mechanism, drivers would be free to travel on both roads- without payments. In that case the value of travel on the fast road would be close to zero. Otherwise, more drivers would start using the fast road, thereby increasing travel time on this road. Furthermore, there is a coordination problem as nearly identical travel times mean that there is no sorting according to the value of time in terms of road usage. In contrast, the efficient allocation both solves the coordination problem, as drivers with a high value of time are allocated to the fast road and ensures that the fast road is indeed faster than the slow road.

Before discussing the incentive problem that arises because each driver's value of time is private knowledge, we will analyze the mechanism designer's problem with perfect information.

First-Best Allocation: We can cancel out the transfers and write the mechanism designer's problem as follows:

$$W = max_{[x]} \sum_{i=1}^{n} \left[x_i v \left(\theta_i, \sum_{j=1}^{n} x_j \right) \right] \tag{4.2}$$

First, at an optimum the benefit of using the fast road needs to be positive. Otherwise, welfare could be increased by having just one driver using it. Second, at the optimum drivers will be sorted according to their value of time, with high-value drivers using the fast road, while low value users use the slow road. Suppose not, so that there is a driver who travels on the slow road with a higher value of time than a driver traveling on

[16]Our assumptions on $v(\theta_i, k)$ can then be obtained by assuming that $\tilde{v}_t < 0$ and $\tilde{v}_{\theta_i} > 0$. We can obtain for all $k' < k$, $v(\theta_i, k+1) - v(\theta_i, k) \leq v(\theta_i, k'+1) - v(\theta_i, k')$ if $C(\cdot)$ is an increasing function with increasing differences and the derivative of $\tilde{v}_t$ is small enough. Furthermore the cross-partial derivative of $\tilde{v}(\cdot,\cdot)$ needs to be strictly negative to ensure that $v_{\theta_i}(\theta_i, k+1) - v_{\theta_i}(\theta_i, k) \leq 0$. This means that for drivers with a higher value of time θ_i, increases in travel time reduce utility by a greater amount.

[17]We consider the case when the mechanism designer wants to maximize total revenue collected in Section 4.5.

the fast road. Switching their allocations leaves the number of drivers on the fast road unchanged, but since $v_\theta > 0$, total welfare increases.

Lemma 1. *Let $x^{FB}(\theta)$ be the allocation that solves Equation 4.2 for a given θ. Then it must satisfy:*

- *For all θ_i, $v(\theta_i, \sum_{j=1}^n x_j^{FB}(\theta)) > 0$.*
- *If $\theta_i > \theta_j$, then $x_i^{FB}(\theta) \geq x_j^{FB}(\theta)$.*

Proof. For the first part suppose the utility from using the fast road is weakly negative under x^{FB}. Then consider letting only a single driver i use the fast road. By assumption for that driver we have $v(\theta_i, 1) > 0$, while all other drivers obtain zero. Hence welfare increased, which contradicts the optimality of x^{FB}.

For the second part suppose otherwise. Then there is at least one pair of drivers $\{i, j\}$ such that $\theta_i > \theta_j$ and $x_i^{FB}(\theta) < x_j^{FB}(\theta)$. But then one could replace $x_i^{FB}(\theta)$ and $x_j^{FB}(\theta)$ by $x_i^*(\theta) := x_j(\theta)$ and $x_j^*(\theta) := x_i(\theta)$ which would lead to change in welfare of

$$v\left(\theta_i, \sum_{j=1}^n x^{FB}(\theta)\right) - v\left(\theta_j, \sum_{j=1}^n x_j^{FB}(\theta)\right) = \int_{\theta_j}^{\theta_i} v_{\tilde{\theta}}\left(\tilde{\theta}, \sum_{j=1}^n x_j^{FB}(\theta)\right) d\tilde{\theta} > 0.$$

Hence welfare increased, which contradicts the optimality of x^{FB}. □

Lemma 1 says that the first-best allocation has a simple structure: all drivers with a value of θ sufficiently high will use the fast road while the remainder will not. To find the efficient allocation, we only need to find the optimal number of drivers using the fast road as a function of θ. We define $\theta^{(k)}$ to be the k^{th} highest value of θ from among the n drivers. Suppose there are $k \in \{0, 1, ..., n-1\}$ drivers on the fast road and consider adding another driver to it. The change in welfare resulting from this reallocation is given by:

$$\Delta_k(\theta) \equiv v(\theta^{(k+1)}, k+1) + \sum_{i=1}^k \left(v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k)\right)$$

The expression Δ_k characterizes the key trade-off in determining the efficient allocation. The first term appearing in Δ_k is the benefit of allocating the driver with the $(k+1)$-highest value of time to the fast road. The value of the trip for this driver is given by $v(\theta^{(k+1)}, k+1)$. The second term captures the cost of increased congestion for the first k drivers from another driver on the fast road. For efficient allocations the first term will generally be positive. The second term always enters negatively due to the assumption that the fast road is congestible, i.e. $v(\cdot,\cdot)$ is decreasing in its second argument. The next Lemma characterizes some useful properties of $\Delta_k(\theta)$ that we will use in the proof of Proposition 1.

Lemma 2. *For all $\theta \in \Theta$ and all $k \in \{0, 1, ..., n-2\}$, we have that $\Delta_k(\theta) > \Delta_{k+1}(\theta)$. Furthermore $\Delta_0(\theta) > 0$.*

Proof. For the first result consider the difference in the first terms of $\Delta_k(\theta)$ and $\Delta_{k+1}(\theta)$ which is given by:

$$\begin{aligned} & v(\theta^{(k+2)}, k+2) - v(\theta^{(k+1)}, k+1) \\ + \sum_{i=1}^{k} \Big[\big(v(\theta^{(i)}, k+2) - v(\theta^{(i)}, k+1) \big) - & \big(v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k) \big) \Big] \\ & + v(\theta^{(k+1)}, k+2) - v(\theta^{(k+1)}, k+1) \end{aligned}$$

From the definition of $\theta^{(k)}$ we have that $\theta^{(k)} > \theta^{(k+1)}$. Furthermore we have that $v(\cdot,\cdot)$ is strictly increasing in its first argument and decreasing in its second argument, so that the first difference is strictly negative. Next consider the second term which is given by:

$$\sum_{i=1}^{k} \Big[\big(v(\theta^{(i)}, k+2) - v(\theta^{(i)}, k+1) \big) - \big(v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k) \big) \Big]$$

This term is weakly negative, which follows from our assumption that for all θ_i, k and $k' < k$, $v(\theta_i, k+1) - v(\theta_i, k) \leq v(\theta_i, k'+1) - v(\theta_i, k')$. Finally consider the third term. This term is also weakly negative, which follows from $v(\cdot, k)$ being decreasing in k. Hence it follows that $\Delta_k(\theta) > \Delta_{k+1}(\theta)$.

We can verify the value of $\Delta_0(\theta)$ from the definition of $\Delta_k(\theta)$ by setting $k = 0$. □

Lemma 2 shows how the trade-off from adding another driver evolves as more drivers use the fast road. First, the benefit of adding one driver falls, since the value of time of the marginal driver decreases. Second, the costs which an additional driver imposes on the other users of the fast road increase, because more drivers are affected by increased congestion. Both effects go in the same direction, so that the welfare gain of each additional driver falls. When there is no driver on the fast road, then there is no cost of adding the first driver. The following Proposition follows directly from Lemma 2 and gives a necessary and sufficient condition for the first-best allocation.

Proposition 1. *There exists some $k^* < n$ such that for all $i = \{1, ..., n\}$ at the optimum $x_i(\theta) = 1$ if and only if $\theta_i \geq \theta^{(k^*)}$. The value of k^* is given by:*

$$k^* = \max_{\Delta_k(\theta) \geq 0} k + 1 \tag{4.3}$$

The Proposition characterizes the efficient number of drivers k^*. We will often refer to x^{FB} as the efficient allocation that leads to the efficient number of drivers k^*. The logic behind Proposition 1 follows from Lemma 2: as the number of drivers on the fast road increases, the benefit of adding another driver shrinks. Welfare changes are strictly positive when the driver with the largest valuation is the only driver on the fast road, but become smaller as more drivers use it. By assumption there is an interior solution, i.e. $v(\cdot, n) < 0$ implies that $k^* < n$.

In the next Lemma, we derive comparative statics of the efficient allocation with respect to θ, the vector of the drivers' value of time.

Lemma 3.

1. *Comparative statics of the efficient number of drivers k^*:*
   a) *If $\theta_i' \geq \theta_l(\theta_{-i})$, then $k^*(\theta_i, \theta_{-i})$ is a weakly decreasing function in θ_i.*
   b) *If $\theta_i' < \theta_l(\theta_{-i})$, then $k^*((\theta_i, \theta_{-i}))$ is constant for $\theta_i < \theta'_{-i}$, increases by one at $\theta_i = \theta_i'$ and for $\theta_i > \theta'_{-i}$ is weakly decreasing.*
2. *The utility $x_i^{FB} v(\theta_i, \sum_{j=1}^n x_j^{FB})$ of driver i at the efficient allocation is weakly decreasing in θ_i, $\forall i$.*

Proof. For the results we need the effect of a change in a single θ_i on the optimal allocation, holding θ_{-i} fixed. Let $\theta_{-i}^{(k)}$ be the k^{th} highest value of time among all drivers except driver i. Consider the auxiliary problem in which we set $\theta_i = 0$, but driver i still needs to be allocated. In that case, it is clear that $x_i = 0$ is optimal. Denote the efficient allocation in this case by k_{-i}^*, which is a function of θ_{-i}. Refer to the driver with the k_{-i}^* highest value of time as driver $l(\theta_{-i})$. Denote his associated value of time by $\theta_l(\theta_{-i}) = \theta_{-i}^{k_{-i}^*}$. We denote by θ_i' the value of θ_i such that adding driver i to the fast road when k_{-i}^* are allocated to it results in no welfare change. Thus, θ_i' is defined by:

$$v(\theta_i', k_{-i}^* + 1) - \sum_{j=1}^{k_{-i}^*} \left[v(\theta^{(j)}, k_{-i}^* + 1) - v(\theta^{(j)}, k_{-i}^*) \right] = 0$$

We consider two cases. First, suppose $\theta_i' \geq \theta_l(\theta_{-i})$. Then for $\theta_i \leq \theta_l(\theta_{-i})$ we have that k^* does not vary in θ_i and neither does the allocation x^{FB}. When $\theta_i > \theta_l(\theta_{-i})$ it follows that $x_i^{FB} = 1$ while $x_l^{FB} = 0$. As θ_i increases, k^* falls. In this case, θ_i is allocated to the fast road, and the cost of adding other drivers increases as the value of time of driver i increases. Therefore k^* is a decreasing function of θ_i.

Second, suppose $\theta_i' < \theta_l(\theta_{-i})$. For $\theta_i < \theta_i'$ we have $k^* = k_{-i}^*$. For $\theta_i \in [\theta_i', \theta_l(\theta_{-i})]$ it is optimal to add driver i to the fast road without removing any other driver from it. Therefore we have $k^* = k_{-i}^* + 1$. For $\theta_i > \theta_l(\theta_{-i})$, k^* falls as θ_i increases. Since θ_i is allocated to the fast road, any increase in θ_i increases the cost of adding other drivers to it, implying that k^* will fall. □

Intuitively, if θ_i increases and becomes just large enough to be assigned to the fast road, driver i will be either added to the set of drivers on the fast road or she replaces someone else. At the point when driver i is added to the fast road, k^* can only be non-decreasing in θ_i. This implies that the number of drivers on the fast road is only non-decreasing in θ_i at that point. Nonetheless, the utility at the efficient allocation for driver i is strictly lower due to Lemma 1, since for a lower θ_i that particular driver is assigned to the slow road. When driver i is assigned to the fast road and θ_i increases further, k^* decreases and the utility obtained by driver i increases.

4.4 Implementing the First-Best Allocation

In this section we consider the problem of allocating drivers to roads as a mechanism design problem. We first derive an efficient and incentive-compatible mechanism and illustrate it with a simple example. Then we consider the congestion pricing problem as the number of drivers increases and present simulation results.

An important consideration in the mechanism design literature is incentive compatibility, i.e. designing mechanisms in a way that drivers always report their private information truthfully.[18] As a first step, we apply the dominant-strategy *revelation principle* of Gibbard (1973), which allows us to study a large class of mechanisms by focusing on a smaller subclass. By the revelation principle every complicated mechanism involving potentially very large message spaces can be replaced by a simpler mechanism that only asks drivers to directly report their type truthfully. Thus, instead of studying complicated mechanisms where in equilibrium a type can be inferred from a message, it is without loss of generality to study a direct mechanism. In such a mechanism, each driver will be asked to report her private information, namely her valuation for travel time. The mechanism designer then allocates drivers to roads based on this information. More precisely, a direct mechanism is a function associating to each θ an allocation, x and a transfer function p, where p is the transfer paid by a driver to the mechanism designer.[19] In short, a mechanism is a mapping from reports of the drivers' valuation of time to an allocation and transfers, i.e. $\left[x(\theta), p(\theta)\right] : \theta \rightarrow X \times \mathbb{R}^n$.

We apply the concept of dominant-strategy incentive compatibility. The mechanism designer requires each driver to prefer truth-telling given all possible valuations of the other drivers. Hence the mechanism works regardless of what driver i believes about the distribution of driver j's valuation of travel time and the mechanism designer obtains the exact valuations of the n drivers without requiring precise information ex-ante.

Definition 1. *A direct mechanism* $[x, p]$ *is dominant-strategy incentive compatible if for all* $\theta_i, \hat{\theta}_i \in \Theta_i$ *and* $\theta_{-i} \in \Theta_{-i}$ *it satisfies*

$$U_i(\theta_i;\theta_{-i}) \equiv x_i(\theta)v(\theta_i, \sum_{j=1}^{n} x_j(\theta)) - p_i(\theta) \geq x_i(\hat{\theta}_i, \theta_{-i})v(\theta_i, \sum_{j=1}^{n} x_j(\hat{\theta}_i, \theta_{-i})) - p_i(\hat{\theta}_i, \theta_{-i})$$
$$\equiv U_i(\hat{\theta}_i, \theta_i; \theta_{-i}) \qquad \text{(DIC)}$$

A mechanism that satisfies DIC makes it optimal for a driver with type θ_i to report this value, rather than any other value $\hat{\theta}_i$ for all other possible reported values θ_{-i} of the other drivers. Requiring that truth-telling is optimal for all possible realizations of other drivers' values θ_{-i} implies that truth-telling is optimal for driver each i no matter what

[18] In our context, truthfully reporting private information means that each driver actually chooses the option which the mechanism designer wants them to take.

[19] Note that transfer functions are the same for all drivers. Hence drivers remain anonymous within the mechanisms we consider. On anonymous congestion charges, c.f. Arnott and Kraus (1998).

her beliefs are about other drivers' valuations.[20] Mechanisms that satisfy DIC are robust to incorrect beliefs of the mechanism designer and do not require detailed knowledge about the distribution of values by the mechanism designer. This is an attractive feature in our application.

The mechanism designer maximizes welfare given by equation 4.1 subject to the DIC constraints. We say that an allocation function $x(\theta)$ is *implemented in dominant strategies* by payment rules $p_i(\theta)$ if they satisfy the incentive-compatibility constraints. Since the mechanism designer maximizes total welfare, there is no revenue-raising motive. Note that the private information held by the drivers affects other drivers only indirectly through the resulting allocation.

In the following discussion, we make use of some extra notation. We denote by $\theta^{(k^*)}$ the k^*-highest value of time, where k^* is as defined in Proposition 1. We suppress the dependence of k^* on θ for simplicity. Let $k^*_{-i}(\theta_{-i})$ be the value of k^* as in Proposition 1 excluding driver i. As before we let $\theta_{-i}^{(k)}$ be the k^{th}-highest value of time in the problem excluding driver i among the $n-1$ remaining drivers. Again, we suppress the dependence of k^*_{-i} on θ_{-i} for simplicity. For notational ease we define for each $i \in \{1,...,n\}$, the following two sets:

$$\Omega_i^+(\theta) \equiv \{j \neq i | \theta_j \geq \theta^{(k^*)}\}$$
$$\Omega_i^0(\theta) \equiv \left\{j \neq i | \theta_j \in \left(\theta_{-i}^{(k^*_{-i})}, \theta^{(k^*)}\right)\right\}$$

Ω_i^+ is the set of drivers, excluding i, that use the fast road irrespective of driver i's allocation. These drivers are affected by driver i only through changes in congestion. Similarly the set of drivers assigned to the fast road if driver i is allocated to the slow road is denoted by Ω_i^0. These drivers are affected by driver i through the impact of his allocation on theirs. If the set Ω_i^0 is empty, then driver i's report does not affect the assignment of other drivers to the fast road, but may affect the level of congestion on it. Note that both Ω_i^+ and Ω_i^0 depend on the vector of the drivers' value of time, θ. Note also that a consequence of Lemma 3 is that $k^* - k^*_{-i} \leq 1$.

Proposition 2. *The following payment rule implements the first-best allocation, $x^{FB}(\theta)$ and specifies for all $i \in \{1,...,n\}$:*

$$p_i^{FB}(\theta) = -\sum_{j\in\Omega_i^+} \left[v(\theta_j, k^*) - v(\theta_j, k^*_{-i})\right] + \sum_{j\in\Omega_i^0} v(\theta_j, k^*_{-i}) \tag{4.4}$$

Proof. The payment rule is an application of the Vickrey-Clarke-Groves mechanism. The basic logic behind such a mechanism is to make each driver residual claimant of total welfare and thereby let them internalize the mechanism designers problem. The

[20]Alternatively one could consider the incentives of drivers to report truthfully given their beliefs about the types of other drivers. However in our application it is unlikely that the mechanism designer knows these beliefs.

incentives for truth-telling are unaffected by an added term which does not depend on a driver's own report, but it may depend on the reports of all the other drivers. In our model, this implies that the VCG payment rule has the following form:

$$p_i^{VCG}(\hat{\theta}_i, \theta_{-i}) = -\sum_{j=1;j\neq i}^{n} x_j(\hat{\theta}_j, \theta_{-i}) v(\theta_j, k^*(\hat{\theta}_i, \theta_{-i})) + h_i(\theta_{-i})$$

Since $k^*(\theta)$ is by definition the function that maximizes welfare for each $\theta \in \Theta$, driver i maximizes her utility, given by:

$$\begin{aligned} U_i(\hat{\theta}_i, \theta_i; \theta_{-i}) &= x_i(\hat{\theta}_i, \theta_{-i}) v(\theta_i, k^*(\hat{\theta}_i, \theta_{-i})) \\ &+ \sum_{j=1;j\neq i}^{n} x_j(\hat{\theta}_i, \theta_{-i}) v(\theta_j, k*) - h_i(\theta_{-i}) \end{aligned}$$

This means that driver i faces the mechanism designer's problem 4.1, so that reporting $\hat{\theta}_i = \theta_i$ is optimal. The function $h_i(\theta_{-i})$ is constant in $\hat{\theta}_i$ and therefore does not affect i's incentives. We choose $h_i(\theta_{-i})$ as the sum of utilities of all drivers but i in the hypothetical scenario in which driver i was excluded. This induce the desirable property that the uncongestible road is for free. We denote by $x_{j,-i}^{FB}$ the optimal allocation of driver j when driver i is excluded.

$$h_i(\theta_{-i}) \equiv \sum_{j=1;j\neq i}^{n} x_{j,-i}^{FB}(\theta_{-i}) v(\theta_j, k_{-i}^*)$$

Since $k_{-i}^*(\theta_{-i})$ solves the allocation problem as if the value of time of driver i was zero, driver i will then be allocated to the slow road. Hence the payment above is the surplus of the other drivers under the optimal allocation given that i is assigned to the slow road. Note that it may happen that $\theta_{-i}^{(k_{-i}^*)} > \theta^{(k^*)}$. In this case, the set Ω_i^0 is empty and only the first term in 4.4 remains. Substituting our choice of $h_i(\theta_{-i})$ into p_i^{VCG} and applying the definitions of Ω_i^+ and Ω_i^0 leads to 4.4. □

The payment schedule 4.4 consists of two components. We call the first term the congestion effect: $-\sum_{j\in\Omega_i^+}\left[v(\theta_j, k^*) - v(\theta_j, k_{-i}^*)\right]$. It captures how driver i's report changes the travel time on the fast road. This effect may be either positive or negative. It will be negative (meaning that driver i has to pay less) under the condition in Part 1(a) of Lemma 3. In that case a higher report of θ_i reduces k^*, the number of drivers on the fast road. All drivers remaining on the fast road will benefit from reduced congestion. The reduction in the price paid by driver i reflects the value of this reduced congestion of the other drivers. The first component will be positive (meaning driver i has to pay more) only if the valuations of time θ imply $k^* - k^*_{-i} = 1$.[21] In that case the report of driver i increases the number of drivers and thereby the congestion on the fast road. When $k^* = k^*_{-i}$, the congestion effect is zero. In that case driver i replaces another driver on the fast road.

The second component is the reallocation effect: $\sum_{j \in \Omega_i^0} v(\theta_j, k^*_{-i})$. It captures how the report of driver i induces a reallocation of other drivers from the fast road to the slow road. For high values of θ_i, it becomes efficient to reduce congestion on the fast road. This is accomplished by reallocating drivers with a lower value of time to the slow road. It follows from Lemma 1 that those who are reallocated obtain lower utility. The reallocation effect reflects these costs. In contrast, drivers allocated to the slow road pay nothing. By being allocated to the slow road, both the congestion and the reallocation effect are zero.[22]

Remark 1. *The payment schedule $p_i^{FB}(\theta)$ is a weakly increasing step function.*

The payment schedule 4.4 depends on θ_i only through its effect on the efficient allocation $k^*(\theta)$. By definition, k^* can only take a finite number of values. It is constant almost everywhere but there are jumps at a finite number of points, whenever the number of drivers allocated to the fast road changes. Therefore the payment schedule faced by driver i is constant almost everywhere and has a finite number of jumps. Furthermore by Lemma 3 the utility of driver i is weakly increasing in θ_i. This implies that the payment schedule of driver i is weakly increasing. If it were not, there would be cases in which driver i could misreport her valuation to obtain both a lower travel time and a lower payment. This would violate incentive compatibility. Most importantly, a single Pigouvian price does not implement the efficient allocation and is therefore not generally optimal. The following example illustrates the step function of the price schedule.

An Example In order to illustrate the pricing schedule, consider a problem in which only the value of time of one driver is unknown. There are five drivers $i = 0, 1, 2, 3, 4$ simultaneously traveling from a common origin to a common destination. For this particular example we derive the utility of driver i from the following fundamentals. Driver i's utility as a function of travel time t is $v - \theta_i t - p$. There is a fast road A and an uncongestible alternative B. It takes a driver half a minute to travel on road A if there are no other drivers. For each driver on road A, the average travel time for all drivers on that road increases by half a minute, that is $t_A(k) = 0.5 + 0.5k$. On Road B it takes two and a half minutes to travel from the origin to the destination irrespective of the number of drivers using it, i.e. $t_B = 2.5$. Plugging this in the functions of travel time and normalizing

[21] This corresponds to case 1(b) in Lemma 3 at the point where k^* is increasing in θ_i. Intuitively, this happens when a driver i's report causes no other driver to be allocated from the congestible road to the slow road.

[22] Note that this follows mainly from our assumption that there is no congestion on the slow road. This assumption will be relaxed in Section 4.6.

net utility on road B to zero, that substracting $v - \theta_i t_B$, leads to utility on road A as a function of the number of drivers on this road, i.e. $U_i^A = v(\theta_i, k) - p = \theta_i(2-0.5k) - p$, and $U_i^B = -p$.

It is common knowledge that the value of time is given by $\theta_i = 11 - i$ for $i \in \{1,2,3,4\}$. The valuation of time of driver $i = 0$, given by $\theta_0 \in (0, \infty)$ is unobservable private information. Let $\mathscr{A}^*(\theta_0)$ be the set of drivers on road A at the efficient allocation, given by:

$$\mathscr{A}^*(\theta_0) = \begin{cases} \{1,2\} & \theta_0 < 8.\overline{63} \\ \{0,1,2\} & 8.\overline{63} \leq \theta_0 < 9.8 \\ \{0,1\} & 9.8 \leq \theta_0 < 32 \\ \{0\} & 32 \leq \theta_0 \end{cases}$$

The efficient allocation depends on the value of θ_0 which we refer to as aggregate uncertainty. Knowing θ_0 the mechanism designer could set a single price as a function of θ_0 such that drivers use roads efficiently. When the mechanism designer does not know θ_0, a mechanism which sets the price as a function of the report of θ_0 gives driver 0 an incentive to lie about her value of time. However the efficient allocation can be implemented by letting driver 0 face the following payment schedule for a travel time $t \in \{0.5, 1, 1.5, 2.5\}$:

$$P^*(t) = \begin{cases} 0 & t = 2.5 \\ 9.5 & t = 1.5 \\ 14.4 & t = 1 \\ 30.4 & t = 0.5 \end{cases}$$

The difference of this payment schedule to setting a single price is that it allows to charge driver 0 different prices for different travel times, while a single price mechanism only charges for use of the fast road irrespective of the number of other drivers on the fast road. This payment schedule is constructed in a way that the prices faced by driver 0 capture the externalities this driver imposes on the other drivers and gives driver 0 incentives to report θ_0 truthfully. Figure 4.2 displays the resulting price schedule and the efficient allocation as a function of driver 0's value of travel time. A similar table can be computed if the valuations of the other drivers are initially unknown to the regulator. We discuss an example of this case in Section 4.4.1.

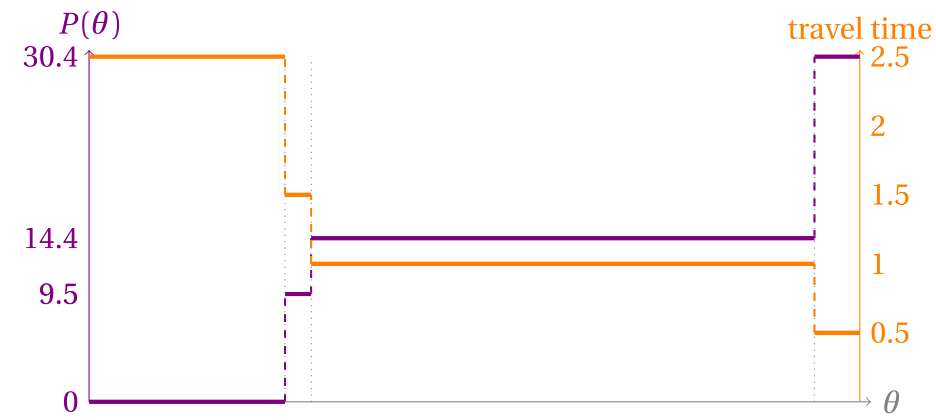


Figure 4.2: Efficient Allocation and Pricing

VCG mechanisms are generally known to be not collusion proof. Several drivers might benefit from jointly misreporting their valuation of travel time. We believe, however, that this kind of collusion is not very relevant in transportation. To have a sizeable effect, a number of drivers would have to coordinate their behavior and possibly agree on transfers. A policymaker could simply outlaw any legally binding agreements of that sort. This would make it very costly for drivers to coordinate their behavior.

4.4.1 Example with Two Drivers

To illustrate, consider an example with two drivers and $v(\theta_i, k) = \theta_i(4-k)$, while utility from using the slow road is constant at zero. This could be derived from fundamentals as follows. The utility of traveling on any road is given by $S - \theta_i t$, where S is the benefit from reaching D and t is the travel time. If we assume that travel time on the fast road A is given by $t = C(k) = k$ and travel time on the slow road B is constant at $t = 4$, we can derive the utility function above.[23] If $\theta_i > \theta_j$ it is efficient for both drivers to use the fast road if and only if $-2(\theta_i + \theta_j) > -\theta_i - 4\theta_j$. This simplifies to $\theta_i < 2\theta_j$. Otherwise only θ_i is efficiently allocated to the fast road. The efficient allocation is

$$(x_1, x_2)(\theta) = \begin{cases} \{1,1\} & if \quad \theta_1 \in [\frac{1}{2}\theta_2, 2\theta_2] \\ \{1,0\} & if \quad \theta_1 > 2\theta_2 \\ \{0,1\} & if \quad \theta_1 < \frac{1}{2}\theta_2 \end{cases}$$

Figure 4.3 shows that when the values of time are similar (i.e. we are in the violet area around the 45 degree line), it is efficient for both of them to use the fast road. When the relative difference in the values of time is large, it is optimal to assign only the driver with the higher value of time to the fast road.

[23] The benefit of reaching the destination S cancels out.

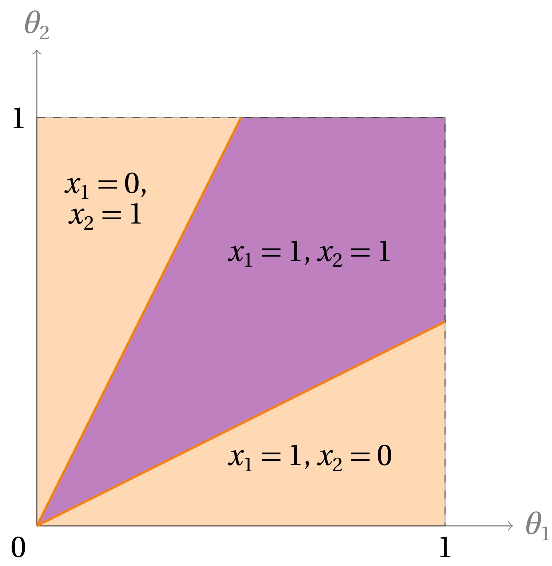


Figure 4.3: Optimal Allocations with Two Drivers

We apply Proposition 2 to get the price schedule that implements this allocation for driver 1:

$$p_1(\theta) = \begin{cases} 0 & if \quad \theta_1 < \frac{1}{2}\theta_2 \\ \theta_2 & if \quad \theta_1 \in [\frac{1}{2}\theta_2, 2\theta_2] \\ 3\theta_2 & if \quad \theta_1 > 2\theta_2 \end{cases}$$

The corresponding travel time for driver 1 is given by the following function:

$$t_1(\theta) = \begin{cases} 4 & if \quad \theta_1 < \frac{1}{2}\theta_2 \\ 2 & if \quad \theta_1 \in [\frac{1}{2}\theta_2, 2\theta_2] \\ 1 & if \quad \theta_1 > 2\theta_2 \end{cases}$$

These schedules also illustrate that we can express the price schedule of driver 1 $p_1(\theta)$ also in terms of travel time $t_1(\theta)$, leading to a schedule of prices per travel time.

Figure 4.4 plots the optimal price schedule faced by driver 1 for two values of θ_2. Note that the price paid by driver 1 is not monotone in θ_2.

Figure 4.5 plots the optimal travel time of driver 1 for two different values of θ_2 as a function of θ_1. Unlike the payment schedule, the travel time is a monotone function of the value of θ_2. The payment schedule is independent of the distribution of θ.

4.4.2 Congestion Pricing in the Limit

Our results so far suggest that each driver should face a price schedule, depending on that driver's valuation and all other driver's reported types. To reconcile with the Pigouvian

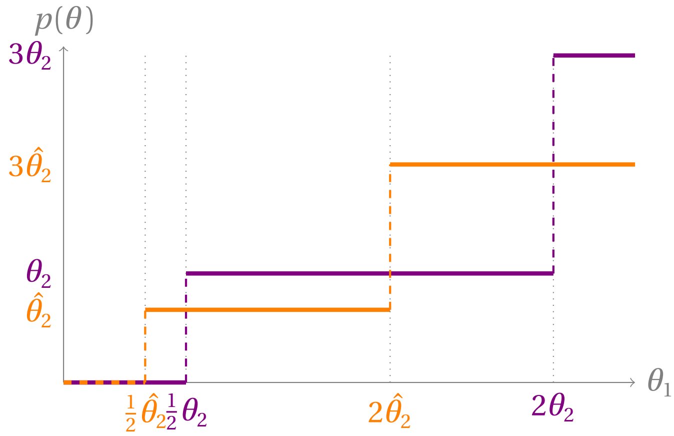


Figure 4.4: Payment Schedule for Two Different Values of θ_2

approach we consider a variant of our model in which the number of drivers goes to infinity and the congestion caused by each driver goes to zero. For simplicity we assume that $v(\theta_i, k) = \theta_i(S - b_n k)$, where $b_n = b/n$ and b, S are a positive constants. S can be interpreted as the benefit of using the fast road A relative to the slow road B in the absence of congestion on the fast road. $b_n \cdot k$ is the travel time on road A. The constant b can be interpreted as measuring the strength of congestion, so that $b_n = b/n$ ensures that the effect each driver has on the level of congestion goes to zero in the limit. This assumption implies that as the number of drivers increases, the effect of each driver on congestion becomes "small". In many settings this assumption is likely to be violated, but it is necessery to bridge the gap between the Pigouvian approach and our more

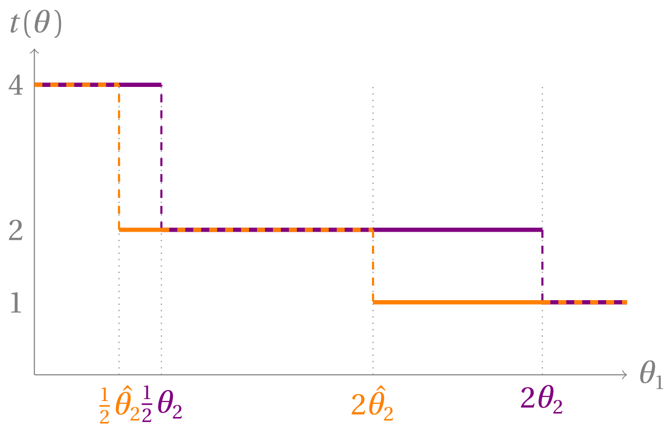


Figure 4.5: Travel Time for Two Different Values of θ_2

general setup. Since in reality, and especially in cities, space is limited, traffic would come to a stand still if the number of drivers gets too large.

We assume $S-b<0$, to ensure that at the efficient allocation not all drivers will use the fast road in the limit. We also normalize the objective function by dividing through by the number of drivers n:

$$W = max_{[x_i(\theta),p(\theta)]}\frac{1}{n}\sum_{i=1}^{n}[u_i(p_i,x;\theta_i)+p_i] \tag{4.5}$$

For each n we can use the results of Lemmas 1 and 2 as well as Proposition 1 to find the optimal solution. We consider now the probability limit of the value of $\Delta_k(\theta)$ as n converges to infinity, while letting k go to infinity, such that $\lim_{n\to\infty}\frac{k}{n}=q\in[0,1]$.

$$\operatorname*{plim}_{n\to\infty}\Delta_k = \operatorname*{plim}_{n\to\infty}\theta^{(k+1)}\left(S-\frac{k+1}{n}b\right)-\frac{b}{n}\sum_{i=1}^{n}\theta_i 1\left(\theta_i>\theta^{(k+1)}\right) \tag{4.6}$$

Note that $\operatorname{plim}_{n\to\infty}\theta^{(k+1)}$ is simply the $(1-q)^{th}$ quantile of the distribution function $F(\cdot)$, which we will denote by $\theta^q\equiv F^{-1}(1-q)$. The probability limit of the second term is:

$$\operatorname*{plim}_{n\to\infty}\frac{b}{n}\sum_{i=1}^{n}\theta_i 1\left(\theta_i>\theta^{(k+1)}\right)=b\int_{\theta^q}^{\overline{\theta}}\theta\, dF(\theta)$$

Therefore, we have that:

$$\operatorname*{plim}_{n\to\infty}\Delta_k=\theta^q\left(S-qb\right)-b\int_{\theta^q}^{\overline{\theta}}\theta\, dF(\theta)\equiv\Delta_q \tag{4.7}$$

The welfare change Δ_q is strictly decreasing in q. We denote by q^* the efficient level of q, given by the unique solution of $\Delta_q=0$. Let θ^* be such that a fraction q^* of all drivers that have a greater value of time than θ^*. The optimality condition in the limit is then given by:

$$\theta^*\left(S-q^*b\right)=b\int_{\theta^*}^{\overline{\theta}}\theta\, dF(\theta) \tag{4.8}$$

Consistent with the Pigouvian approach a deterministic share of drivers efficiently uses the fast road in the limit. Therefore, we can set a single congestion charge to implement the efficient allocation in the limit. The following Proposition summarizes the preceding discussion:

Proposition 3. *In the limit, a unique Pigouvian price p^* for the use of the fast road implements the efficient allocation. This price is given by:*

$$p^* = b \int_{\theta^*}^{\overline{\theta}} \theta \, dF(\theta) \tag{4.9}$$

The efficient congestion charge equals the value of the marginal increase in travel time from a small increase in the number of drivers using the fast road. The Pigouvian approach holds only if there are infinitely many small drivers. Intuitively, each driver has a negligible effect on the other drivers in the limit. To set the Pigouvian congestion charge correctly, the mechanism designer needs to know the distribution of the value of time. In contrast, the optimal mechanism of Proposition 2 does not require any prior knowledge of this distribution.

4.4.3 Simulations

The limit results of the previous section suggest that an appropriately set Pigouvian price maximizes welfare. There are two impediments to setting the Pigouvian price optimally. First, in practice the number of drivers is finite. Given that we focus on a static problem in which all drivers use the fast road simultaneously, it is likely that the number of drivers in applications is low. Especially in cities, traffic is possible only if the number of drivers is not too large since otherwise it turns into a standstill. In other traffic-related applications like shipping or air travel, the number of participants is frequently quite low as well. In these case, one cannot rely on limit results and the more complex price schedule that we introduce in this paper is optimal.

Second, setting the Pigouvian price optimally requires knowledge of the distribution of the value of time. As this is usually not known by the policy maker, the price will be set either as a function of the policy maker's prior or may need to be estimated. Even the most flexible estimation method cannot take unobservable factors into account.

Both issues are no longer relevant when the mechanism designer asks drivers directly about their valuation of travel time and sets payment schedules that induce truthful revelation of these information.

To illustrate potential problems with mechanism designers' priors, suppose the common distribution of the value of time depends on an unobservable parameter α, so that each θ_i is distributed according to $F(\theta_i;\alpha)$. We assume that α is distributed uniformly over the unit interval. This implies that from the mechanism designer's view, the values of θ_i are not independent. The mechanism designer's prior is then $f_p(\theta_i) \equiv \int_0^1 f(\theta_i;\alpha) d\alpha$. The mechanism designer sets the Pigouvian price based on this prior as follows:

$$p = b \int_{\theta^*}^{\overline{\theta}} \theta \, dF_p(\theta)$$

This will not yield the optimal price $p^*(\alpha)$ that the mechanism designer would set if he knew α. In contrast, our mechanism always achieves an efficient allocation. So far aggregate uncertainty resulted from a finite number of drivers. The mechanism designer's lack of knowledge over α is another source of aggregate uncertainty. Therefore, aggregate uncertainty implies that a Pigouvian price based on priors is not efficient even with a continuum of drivers.

We analyze the performance of Pigouvian prices using a simulation. We calculate the normalized welfare loss[24] resulting under the Pigouvian price. We also consider welfare losses under pricing errors (±20%). The results are summarized in Figure 4.6. The purple curve indicates the welfare loss for the correct Pigouvian limit price. This welfare loss vanishes as the number of drivers increases. But for a small number of drivers there is a welfare loss. The welfare loss when a price is set at 20% below the correct Pigouvian price is in orange. It is clear that the welfare loss under this price converges to a strictly positive value.

When there are few drivers the lower price gives a lower welfare loss than the correct Pigouvian price. This can result from realizations of the valuations of time such that none of the drivers is willing to pay the price for the fast road. In that case a lower price may induce at least one driver to use the fast road, which always dominates no driver using the fast road. If we restricted attention to a single price, the optimal price under this constraint depends on n. The blue curve shows the welfare loss when the price is set at 20% above the optimal Pigouvian price. Again, the welfare loss does not vanish as the number of drivers increases.

The simulation[25] highlights that even when there are infinitely many small drivers our mechanism can achieve significant welfare gains by acquiring the information needed to set the optimal price. While the Pigouvian price is optimal in the limit, it is not optimal for each realization of the values of time with finitely many drivers.

[24]We calculate the welfare loss as follows: $Loss = \frac{W^* - W(p)}{W^* - W(0)}$, where W^* is maximum welfare, $W(p)$ gives welfare when a price of p is set and $W(0)$ is welfare resulting from a price of 0 and all drivers' travel time is given by $\overline{t}$. We compare the loss in welfare from setting a price p to the welfare loss from not having a mechanism. Thus welfare losses are normalized by the maximal welfare gain from a mechanism.

[25]The simulation compares welfare under the efficient limit price (Pigouvian price) ±20% to maximum welfare in %. For each n, 10,000 random draws were taken. Parameters are set as follows: $b = 15$, $\overline{t} = 14$, $\underline{t} = 0$. We choose a lognormal distribution for $F(\cdot)$ with a median of 21.46 \$/h and interquartile range of 10.47 \$/h, taken from Small et al. (2005), Table 3. Hence the natural logarithm of the value of time is distributed with a mean of 3.07 and a variance of 0.13. The efficient limit price under this set-up is given by 186\$, which implies that around 40.2% of drivers use the fast road in the limit.

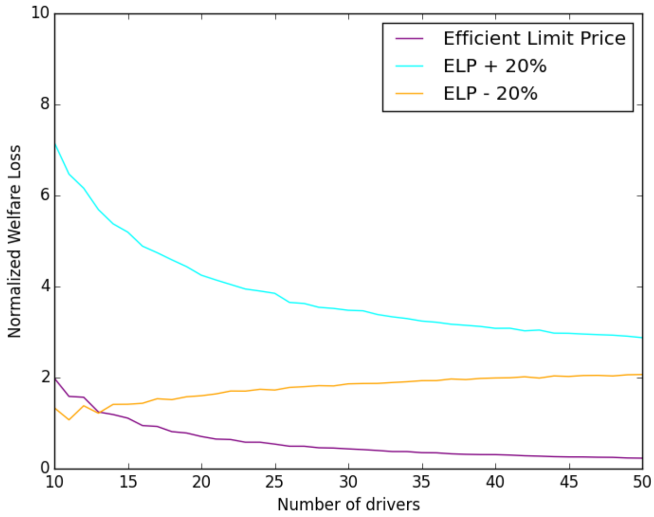


Figure 4.6: Welfare Loss of Single Price Mechanisms in the Limit

4.5 Revenue Maximization

Until now the mechanism designer wanted to maximize total surplus. Instead governments could use congestion charges to raise funds or a private company could be the mechanism designer. In this section the mechanism designer maximizes revenue given her beliefs about the distribution of θ. For simplicity, we assume correct beliefs $F(\cdot)$, and that $F(\cdot)$ has positive mass only on the interval $\Theta = [\underline{\theta}, \overline{\theta}]$.

We denote by $U(\hat{\theta}_i, \theta_i; \theta_{-i})$ the utility to driver i when her true type is θ_i, she reports $\hat{\theta}_i$ and the value of time of the other drivers is given by the vector θ_{-i}. Hence we have:

$$U(\hat{\theta}_i, \theta_i; \theta_{-i}) \equiv x_i(\hat{\theta}_i, \theta_{-i}) v(\theta_i, \sum_{j=1}^{n} x_j(\hat{\theta}_i, \theta_{-i})) - p_i(\hat{\theta}_i, \theta_{-i})$$

The mechanism designer thus faces the following incentive and participation constraints:

$$U_i(\theta_i; \theta_{-i}) \equiv U_i(\theta_i, \theta_i; \theta_{-i}) \geq U_i(\hat{\theta}_i, \theta_i; \theta_{-i}), \forall i \textit{ and } \hat{\theta}_i, \theta_i \in \Theta_i, \theta_{-i} \in \Theta_{-i}. \tag{4.10}$$

$$U_i(\theta_i; \theta_{-i}) \geq 0, \forall i \textit{ and } \theta_i \in \Theta_i, \theta_{-i} \in \Theta_{-i}. \tag{4.11}$$

The mechanism designer maximizes total expected revenues

$$W = max_{[x(\cdot),p(\cdot)]}\mathbb{E}_\theta\left[\sum_{i=1}^n p_i\right]$$
$$= max_{[x(\cdot),U(\cdot)]}\sum_{i=1}^n \mathbb{E}_\theta\left[x_i(\theta)v\left(\theta_i,\sum_{j=1}^n x_j(\theta)\right)-U_i(\theta)\right] \tag{4.12}$$

subject to the incentive constraints 4.10 and participation constraints 4.11.

We solve this problem as in the classic optimal auction problem in Myerson (1981). The proof of the following Lemma is standard and therefore omitted.

Lemma 4. *The incentive constraints in equation 4.10 are equivalent to:*

$$U_i(\theta_i;\theta_{-i}) = U_i(\underline{\theta}_i;\theta_{-i}) + \int_{\underline{\theta}_i}^{\theta_i} x_i(\tilde{\theta}_i,\theta_{-i})v_{\tilde{\theta}_i}\left(\tilde{\theta}_i,\sum_{j=1}^n x_j(\tilde{\theta}_i,\theta_{-i})\right)d\tilde{\theta}_i \tag{4.13}$$

and

$$v_{\theta_i}\left(\theta_i,\sum_{j=1}^n x_j(\theta_i,\theta_{-i})\right) \tag{4.14}$$

is weakly increasing in θ_i.

Profit maximization requires that the participation constraint is binding for the lowest type, i.e. $U_i(\underline{\theta}_i;\theta_{-i}) = 0$ for all $\theta_{-i}\in\Theta_{-i}$. We can then insert the first expression of Lemma 4 into the mechanism designers' objective function and simplify in the usual way by using integration by parts to get the following expression:

$$max_{[x(\cdot)]}\sum_{i=1}^n \mathbb{E}_{\theta_{-i}}\left[\int_{\underline{\theta}}^{\overline{\theta}}\left(v(\theta_i,\sum_{j=1}^n x_j(\theta)) - \frac{1-F(\theta_i)}{f(\theta_i)}v_{\theta_i}(\theta_i,\sum_{j=1}^n x_j(\theta))\right)x_i(\theta)f(\theta_i)d\theta_i\right] \tag{4.15}$$

Substituting the *virtual value* $w_i \equiv v(\theta_i,\sum_{j=1}^n x_j(\theta)) - \frac{1-F(\theta_i)}{f(\theta_i)}v_{\theta_i}(\theta_i,\sum_{j=1}^n x_j(\theta))$, we obtain a maximization problem that corresponds to 4.2 subject to the additional monotonicity requirement of Lemma 4. This implies that we can use our earlier results concerning the implementation to the problem of revenue maximization. This holds as long as the allocated value of travel is weakly increasing in θ_i for all $\theta_{-i}\in\Theta_{-i}$ and as long as virtual values are declining in $k=\sum_{j=1}^n x_j(\theta)$, the number of drivers on the fast road. When $v(\theta,\cdot)$ is linear in θ_i, a sufficient condition is that virtual valuations are weakly increasing in θ_i, which is the case for example if the distribution of the value of time is such that $\frac{1-F(\theta_i)}{f(\theta_i)}$ is weakly decreasing in θ_i.

To obtain the optimal number of drivers on the fast road, the mechanism designer trades off the additional virtual valuation for travel time of adding another driver to the fast road with the reduced virtual valuation of all drivers already allocated to this road. Thus, the algorithm is the same as in Proposition 1 but using virtual valuations v_i instead of θ_i for all drivers i. We assume that valuations are distributed identically across drivers. This implies that a driver with a higher virtual value also has a larger value of time, so that a similar sorting of drivers by value of time occurs. Hence keeping the number of drivers fixed, the revenue maximizing mechanism efficiently selects drivers to drive on the fast road.[26] The payment schedules that implements the revenue maximizing allocation can be obtained in the usual way by substituting the resulting allocation of drivers to roads into equation 4.13 and rearranging.

This procedure is very similar to the one studied before, but there is an important difference: knowledge of virtual valuations requires the mechanism designer to have some belief concerning the distribution $F(\cdot)$. As a result, the revenue maximizing mechanism will only maximize revenues if the mechanism designer knows the true distribution of the value of time. Instead, the VCG mechanism is efficient irrespective of the distribution of values of time.

4.5.1 Example with Two Drivers Continued

We illustrate the differences between welfare and revenue maximization using the simple example with two drivers from Subsection 4.4.1. We now need to make an assumption about the distribution of the value of time of the drivers. Specifically we assume that $F(\theta_i)$ is the uniform distribution over the unit interval, so $\theta_i \sim U(0,1)$ for $i = 1,2$. Substituting into equation 4.15 yields:

$$max_{[x_i(\theta)]} \sum_{i=1}^{n} \int_0^1 (2\theta_i - 1) \left[x_i(\theta)(4 - \sum_{j=1}^{n} x_j(\theta)) \right] f(\theta_i) d\theta. \tag{4.16}$$

A necessary condition for $x_i(\theta) = 1$ is that $\theta_i \geq 0.5$. If only one driver satisfies this condition, she is the only driver on the fast road. If both drivers satisfy the condition, the driver with the higher value of time is allocated to the fast road. It remains to specify when the second driver is allocated to the fast road. Let j be the driver with the lower value of time. The effect on profits of adding driver j to the fast road is given by $-(2\theta_{-j} - 1) + 2(2\theta_j - 1)$. The first term captures the virtual value of driving on the fast road for driver $-j$. The second term represents the additional virtual value of driver j from switching from not driving to driving. Driver j is added to the fast road if this term is positive. This characterizes the optimal allocation, which we depict in Figure 4.7:

[26] If drivers differed in their distribution of the value of time, this would no longer be true.

$$
(x_1^*, x_2^*)(\theta) = \begin{cases} \{0,0\} & if \quad \theta_1, \theta_2 < 0.5 \\ \{1,1\} & if \quad \theta_1, \theta_2 \geq 0.5 \ \& \ \theta_1 \in (0.5\theta_2 + 0.25, 2\theta_2 - 0.5) \\ \{1,0\} & if \quad \theta_1 \geq \max(0.5, 2\theta_2 - 0.5) \\ \{0,1\} & if \quad \theta_2 \geq \max(0.5, 2\theta_1 - 0.5) \end{cases}
$$

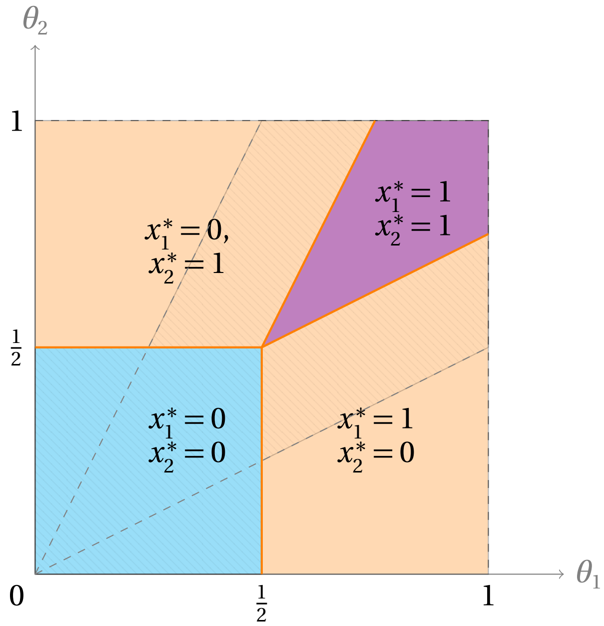


Figure 4.7: Revenue Maximizing Allocation with Two Drivers

As in Figure 4.3, the violet area shows values of (θ_1, θ_2) for which both drivers use the fast road. In orange areas only one driver uses the fast road, while in the blue area no one uses the fast road. The dashed lines in Figure 4.7 indicate the solution for the efficient allocation from Figure 4.3. Dashed areas highlight values of (θ_1, θ_2) in which the revenue maximizing allocation does not coincide with the efficient allocation. Notice that the number of drivers on the fast road is weakly lower in the revenue optimal allocation compared to the efficient allocation. Therefore the driving time on the fast road is also weakly lower in the revenue optimal allocation compared to the efficient allocation.

Given this allocation we can use the drivers' incentive constraints to derive the payment schedule that implements the revenue maximizing allocation:

$$
p_i^*(\theta) = \begin{cases} 0 & if \quad \theta_i < \max\{0.5, 0.5\theta_{-i} + 0.25\} \\ \theta_{-i} + 0.5 & if \quad \theta_{-i} \geq 0.5 \ \& \ \theta_i \in [0.5\theta_{-i} + 0.25, 2\theta_{-i} - 0.5) \\ 3\max\{0.5, \theta_{-i}\} & if \quad \theta_{-i} \in [0, 0.75) \ \& \ \theta_i \geq \max\{0.5, 2\theta_{-i} - 0.5\} \end{cases}
$$

Figure 4.8 shows in violet the pricing function for driver i for different values θ_{-i}. For comparison, we show in orange the payment schedule that implements the efficient allocation.

The qualitative features of the revenue maximizing payment schedule are similar to the efficient payment schedule. For values of θ_i below 0.5, the price charged by the efficient mechanism is larger than the revenue-maximizing price.[27] The mechanism designer does not allow drivers with a low value of time to use the fast road, because it allows her to charge higher prices for values of time above 0.5. Note that payments are not monotone in the reported value of time of the other driver.

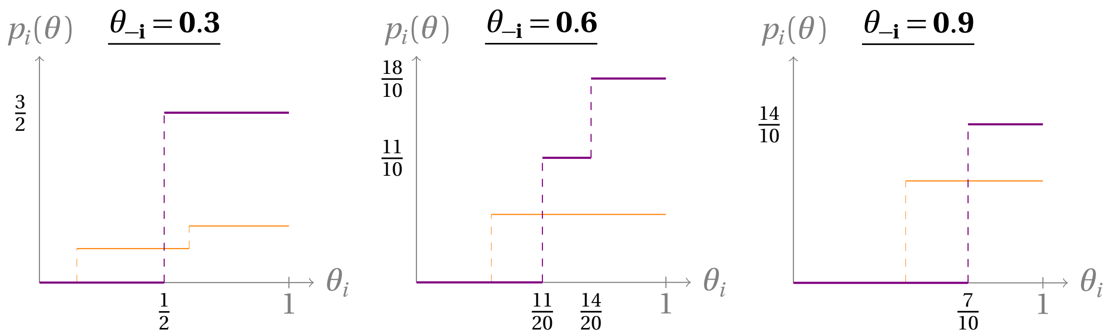


Figure 4.8: Payment Schedules for Driver i Given Different Values of θ_{-i}.

4.6 Two Identical Roads

In this section we relax the assumption of a non-congestible alternative and study the case of two identical roads. This covers the relevant case of two lanes of the same road. For example, in California there are special high-occupancy vehicle lanes parallel to normal lanes.[28] As before, there are two roads A and B. The utility of driving on each of the two roads, ignoring transfer payments, is given by $v(\theta_i, k)$. There is an odd number n of drivers, such that $n = 2m + 1$ and $m \in \mathbb{N}$. This ensures that at the efficient allocation features a fast and a slow road.[29] As before, we refer to road A as the fast road, which means that fewer drivers use road A. The conditions on $v(\cdot,\cdot)$ are the same as before.

First-Best Allocation with Two Identical Roads We begin again by characterizing the efficient allocation that maximizes total surplus 4.1. We denote the efficient allocation

[27] This feature also occurs in the context of auctions. When comparing the optimal auction of Myerson (1981) below the reservation price with a standard auction without a reserve price.

[28] See for example Small et al., 2005

[29] We make this assumption for simplicity. The solution for an even number of drivers is similar, although for some realizations of θ, the utility on both roads will be the same.

as a function of the drivers' valuations by $x^{FB,2}(\theta)$, where the superscript 2 refers to the case of two identical roads.

Lemma 5. *Let $x^{FB,2}(\theta)$ be the efficient allocation. Then it must satisfy:*

- $\sum_i x_i^{FB,2}(\theta) < n - \sum_i x_i^{FB,2}(\theta)$.
- *For all i, j and $\theta \in \Theta$ such that $\theta_i > \theta_j$, we have $x_i^{FB,2}(\theta) \geq x_j^{FB,2}(\theta)$.*

The first point states that one road will be less congested than the other. This follows directly from assuming an odd number of drivers and full participation. Clearly, if all drivers are allocated to one of the roads, one road has to be faster. Since there is a slow road and a fast road, it is again optimal to put drivers with a high value of time on the fast road, starting with the highest valuations. This is the intuition behind the second point. The proof for the second point is analogous to the proof of Lemma 1. Lemma 5 implies that the faster road has the drivers with the k highest valuations and the smaller portion of drivers, i.e. $k \leq m$.

Next, we need to determine the efficient number of drivers on the fast road. Consider adding another driver to it. By Lemma 5 it can only be optimal to add the driver with the $(k+1)^{th}$ highest valuation to the fast road. Then welfare changes by:

$$\Delta_k^2(\theta) = \left[v(\theta^{(k+1)}, k+1) - v(\theta^{(k+1)}, n-k)\right] + \sum_{i=1}^{k}\left[v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k)\right] + \sum_{i=k+1}^{n}\left[v(\theta^{(i)}, n-k-1) - v(\theta^{(i)}, n-k)\right]$$

The first term is the benefit for the driver that was previously using the slow road and is now transfered to the fast one. The second term reflects the loss to drivers on the fast road from sharing the road with another driver. The trade-off between these two effects is also present in the base model without congestion on the slow road and expressed in $\Delta_k(\cdot)$. The third term is new and represents the benefit to drivers on the slow road from having one fewer driver on it. The next Lemma gives the properties of $\Delta_k^2(\theta)$.

Lemma 6. *For all $\theta \in \Theta$ and all $k \in 1, \ldots, m$, we have that $\Delta_k^2(\theta) > \Delta_{k+1}^2(\theta)$. Furthermore $\Delta_0^2(\theta) > 0$ and $\Delta_m^2(\theta) < 0$.*

The Proof of Lemma 6 is similar to that of Lemma 2.

Proof. We can write $\Delta_0^2(\theta) = v(\theta^{(1)}, 1) - v(\theta^{(1)}, n) + \sum_{i=2}^{n}\left[v(\theta^{(i)}, n-1) - v(\theta^{(i)}, n)\right]$. Clearly, this expression is positive. For $k = m$ we have that $\Delta_m^2 = \sum_{i=1}^{m}\left[v(\theta^{(i)}, m+1) - v(\theta^{(i)}, m)\right] + \sum_{i=m+2}^{n}\left[v(\theta^{(i)}, n-m-1) - v(\theta^{(i)}, n-m)\right]$. Given that $n - m = m + 1$ and $v_{\theta_i}(\theta_i, k+1) \leq v_{\theta_i}(\theta_i, k)$ this expression is negative since the first sum contains the m largest values of θ, while the second term contains the m smallest values of θ. To see that $\Delta_k^2(\theta)$ is decreasing in k, consider the first difference, which is given by:

$$\begin{aligned}\Delta_{k+1}^2(\theta)-\Delta_k^2(\theta)=&[v(\theta(k+1),k+2)-v(\theta^{(k+2)},n-k-1)]\\&-[v(\theta(k+1),k+1)-v(\theta^{(k+1)},n-k)]\\&+\sum_{i=1}^{k+1}[v(\theta^{(i)},k+2)-v(\theta^{(i)},k+1)]\\&-\sum_{i=1}^{k}[v(\theta^{(i)},k+1)-v(\theta^{(i)},k)]\\&+\sum_{i=k+3}^{n}[v(\theta^{(i)},n-k-2)-v(\theta(i),n-k-1)]\\&-\sum_{i=k+2}^{n}[v(\theta^{(i)},n-k-1)-v(\theta(i),n-k)]\end{aligned}$$

Adding $v(\theta(k+2),k+1)-v(\theta^{(k+2)},n-k)$ to line 1 and subtracting it from line 2 yields a negative term. Rearranging lines 3 and 4 as well as 5 and 6 also yields negative terms after rearranging and making use of the assumptions about $v(\cdot,\cdot)$. □

Using the results of Lemma 6 the efficient allocation is the following.

Proposition 4. *There exists some k^{**} such that for all $i=\{1,...,n\}$ at the optimum $x_i(\theta)=1$ if and only if $\theta_i \geq \theta^{(k^{**})}$. The value of k^{**} is given by:*

$$k^{**}=\max_{\Delta_k^2(\theta)\geq 0} k+1 \tag{4.17}$$

The proof follows the same lines as the proof of Proposition 1 making use of Lemmas 5 and 6.

Implementing the First Best Having characterized the first-best allocation, we now find a payment schedule that ensures truth-telling of drivers in dominant strategies. We maximize the mechanism designer's objective function 4.1 subject to DIC and additionally take into account the congestion effects on the slow road. We again need to define a few new terms, similarly to Proposition 2. We denote by k_{-i}^{**} the optimal number of drivers on the fast road in the auxiliary problem when driver i's value of time is ignored, i.e. set equal to zero. We denote by $\theta_{-i}^{(k)}$ the k^{th}-highest value of time among all drivers excluding driver i.

We define the following sets:

$$\Omega_i^+(\theta) \equiv \left\{ j \neq i | \theta_j \geq \theta^{(k^{**})} \right\}$$
$$\Omega_i^0(\theta) \equiv \left\{ j \neq i | \theta_j \in \left(\theta_{-i}^{(k_{-i}^{**})}, \theta^{(k^{**})} \right) \right\}$$
$$\Omega_i^-(\theta) \equiv \left\{ j \neq i | \theta_j \leq \theta^{(k_{-i}^{**})} \right\}$$

The sets Ω_i^+ and Ω_i^0 are defined similarly as in the base model. The set Ω_i^- is the set of those drivers that are allocated to the slow road irrespective of the allocation of driver i.

Proposition 5. *With two identical roads, the first-best allocation $x^{FB,2}(\theta)$ is implemented by the following payment rule, which specifies for all $i \in 1, ..., n$:*

$$p_i^{FB,2}(\theta) = -\sum_{j \in \Omega_i^+} \left[v(\theta_j, k^{**}) - v(\theta_j, k_{-i}^{**}) \right] + \sum_{j \in \Omega_i^0} \left[v(\theta_j, n - k^{**}) - v(\theta_j, k_{-i}^{**}) \right] + \sum_{j \in \Omega_i^-} \left[v(\theta_j, k_{-i}^{**}) - v(\theta_j, -k^{**}) \right] \tag{4.18}$$

Proof. The proof follows along the same lines as the proof of Proposition 2. □

The first two terms are the congestion effect -$-\sum_{j \in \Omega_i^+} \left[v(\theta_j, k^{**}) - v(\theta_j, k_{-i}^{**}) \right]$ - and the reallocation effect - $\sum_{j \in \Omega_i^0} \left[v(\theta_j, n - k^{**}) - v(\theta_j, k_{-i}^{**}) \right]$ - analogous to the payments in Proposition 2. The congestion effect captures the externality on drivers on the fast road of adding driver i to this road. The reallocation effect corresponds to the externality that driver i imposes on those drivers that were on the fast road initially, but are reassigned to the slow road when driver i is considered in the optimization problem. We call the third term - $\sum_{j \in \Omega_i^-} \left[v(\theta_j, k_{-i}^{**}) - v(\theta_j, -k^{**}) \right]$- the de-congestion effect. It is new relative to equation 4.4. The effect represents the externality on the drivers that remain on the slow road when driver i is added to the optimization problem. The de-congestion effect has the opposite sign than the congestion effect.

When the slow road is non-congestible, moving driver i from the slow road to the fast road or adding more drivers to the slow road has no externality on drivers on the slow road. In the case of two identical roads the number of drivers allocated to the slow road now also has a congestion effect and hence affects the travel time on the slow road B. Adding more drivers to the fast road decreases the travel time on the slow road. We again normalize payments such that those drivers who do not affect the final allocation pay zero transfers. When there is congestion on both roads, this however means that drivers on the slow road might pay a positive transfer. Intuitively, drivers need to pay whenever they affect the allocation, since only in that case they impose an externality on others.

More precisely, consider a driver i who use the slow road at the efficient allocation. When driver i is not considered in the maximization, the benefit of adding another driver to the fast road is lower than when the effect on driver i is also considered. This is because when driver i is not considered, she implicitly has a value of $\theta_i = 0$, meaning she does not care about congestion. This may lead to more drivers on the slow road when i is not considered. Hence driver i might influence the final allocation with her report even though it does not affect her own road allocation. While driver i uses the slow road in both circumstances, the travel time on the slow road will be lower when she is considered. This means that drivers now pay for faster travel on *both* roads.

As before, the efficient payment schedule depends on θ_i only through its effect on k^{**} which implies that the payment schedule is again a weakly increasing step function. Also with two identical roads, the *single* Pigouvian price is not an efficient mechanism.

4.7 Congestion Pricing in Practice

In this section we discuss currently used or proposed road and congestion pricing schemes around the world and explain how they relate to our mechanism. We suggest modifications to some of these schemes that bring them closer to our proposed mechanism and thereby achieve efficiency gains. Moreover, we explore how our mechanism relates to incentives and special regulations regarding ride sharing and challenges of initial incomplete implementation.

The majority of road-pricing schemes currently in place do not contain an explicit congestion-pricing element in the sense of price discrimination with respect to the level of congestion. Many of these systems regulate long-distance traffic on highways. Take for instance per usage or per distance charges (e.g. road pricing for trucks in Germany, or highway fees in France and Italy, etc.) or a vignette to have the permission to use a road network (e.g. the Austrian or the Swiss highway system). These systems might indirectly affect congestion by reducing overall demand for car rides. But since prices do not adjust to congestion levels or travel time, they do not directly address congestion. Some systems deal with the congestion externality more directly by charging higher prices during rush hours or when road usage is high.

Road-pricing schemes in urban areas more frequently aim at reducing congestion. A particularly simple form of congestion pricing is *cordon area congestion pricing*, i.e. a fee to enter a specific (congestion prone) area of the city. In many cases, prices vary to achieve other objectives, for example environmental ones (e.g. in London or Milan).[30] If the fee structure of road prices is not flexible with respect to traffic conditions, the resulting level of congestion is unlikely to be efficient. Even though Leape (2006) argues that the introduction of the London congestion charge is "a triumph of economics" and led to time savings and better reliability of transport in general, he only finds small positive net benefits. Prud'homme and Bocarejo (2005) on the other hand argue that the

[30] The London congestion charge for instance is a daily price (currently £11.50) for entering one zone in the inner city center which will be charged between 7:00 a.m. and 6:00 p.m.. However, there are various discounts for cars with low emissions, vans, residents, etc (Transport for London, 2017).

economic benefits of the congestion charge represent less than 60% of the economic costs.

More complex and flexible systems are in place in Singapore, and the Swedish cities Gothenburg and Stockholm (The Swedish Transport Agency, 2017; Land Transport Authority, 2017) and are more adequate to reducing the actual congestion externality. In these systems tolls are charged automatically, prices depend on observable characteristics and may vary over the course of the day and at different places.[31] In our mechanism, price discrimination by the value of travel time increases the amount and the quality of the information available. Today's existing systems are already highly complex and could easily extended to additionally incorporate unobservable characteristics, such as the value of travel time. In practice, the Swedish or the Singaporean systems could be extended by offering a menu of travel times with respective prices to reveal the value of travel time via a simple smartphone app or on-board navigation systems.

To the best of our knowledge there is no congestion-pricing system in place which uses price discrimination to gather information about the drivers' value of travel time. Nonetheless, the idea of our approach is partially implemented in an ad-hoc way without using monetary transfers. For instance, ambulances have a high value of travel time in case of accidents since any delay might cause a loss of life. In such cases, regulation gives ambulances a privileged use of roads. Similar rights are frequently awarded to convoys transporting heads of state or heavy cargo. These privileges increase travel time for other drivers. In terms of our model, these regulations are appropriate if the value of time for the privileged drivers (i.e. a convoy or an ambulance) is large enough to render the reduced travel time for other drivers efficient.

A concern regarding the practical implementation of congestion-pricing schemes is that not all drivers might participate in the mechanism right from the start. The systems in Sweden and Singapore shows that full implementation for any vehicle is feasible. But even if this was not so, we do not believe that partial implementation is a fundamental obstacle. For instance, our system could initially be implemented only on designated fast lanes, where non-participating drivers are banned. Alternatively, non participants could be charged based on their effective travel time, assuming that they would have chosen the actual price and time combination offered by the mechanism designer. This would be similar to the system of Stockholm for non-participants. As noted earlier, charging without detailed notification is a common practice in the Stockholm (The Swedish Transport Agency, 2017).

Another area of transportation research is ride sharing (see Furuhata et al., 2013 for a recent survey), which gained interest due to technological progress. Ride sharing refers to any action in which travelers share a vehicle to go from their origins to their destinations. Public transport is usually regarded as a cheap but inconvenient form due

[31] In Sweden, cameras make pictures of number plates and record place and daytime. Payments are subsequently based on these information. In Singapore, drivers are obliged to have transmitters in their vehicles and payments even vary with respect to traffic conditions, which are estimated based on retrospective data and current traffic flows, similar to a real time weather forecast.

to the fixed routes and schedules. More flexible solutions could be obtained by so-called dynamic ride sharing, i.e. a real-time matching of travelers with similar itineraries. With the advent of new technologies such as GPS and smart-phones, dynamic ride sharing gained considerable commercial and academic attention.[32] Our model does not feature ride-sharing. However, the introduction of any per vehicle fees increases the incentives to share rides, which could further reduce the extent of congestion problems in practice.[33]

Many cities try to increase incentives to engage in car pooling and ride sharing, e.g. the designation of special bus lanes, which can only be used by buses and taxis. In the context of our model, this could be efficient, since a bus usually carries several people, implying a large value for time of the bus overall. Similarly, the High-Occupancy Vehicle Lanes in California allow faster travel for cars which seat more passengers. The assumption is that vehicles carrying more people implicitly have a higher value for time. In contrast to our optimal mechanism, these allocations are implemented through rigid rules rather than flexible pricing schemes. Rules such as free access to fast roads for cars with a minimum number of drivers are not necessarily desirable. Such a rule increases the number of cars on the fast road without explicitly taking valuations into account. With our mechanism implemented, such rules might no longer be needed.

4.8 Discussion & Conclusion

This paper shows that the traditional Pigouvian approach of internalizing social costs of congestion by setting a *single* congestion charge applies only when there are infinitely many drivers and their impact on congestion is negligible. The generally optimal solution involves charging drivers a variety of different prices. These prices depend on the desired travel time, i.e. on the number of other drivers using the same road.

One major advantage of applying mechanism design to congestion problems is that it obviates the need to conduct detailed econometric studies to estimate the distribution of the value of time. Moreover, our mechanism can be adapted to incorporate other externalities such as local- or global pollution and accidents.[34] The mechanism design approach requires that each driver may communicate instantaneously with the mechanism designer. Given modern communication technology, we do not believe this is a major issue.

A basic lesson from our mechanism is that there are potential welfare gains from introducing discrimination *within roads* with respect to travel time. We give simple

[32] Kleiner et al. (2011) for instance, evaluate the potential of a VCG mechanism for the efficient matching of drivers to vehicles in such a dynamic ride-sharing setting.

[33] When Singapore introduced congestion pricing in 1975, the number of cars entering the center decreased by 41.6% initially and 22.9% in the long-run, (Morrison, 1986). Interestingly, Leape (2006) does not report any effects of the London congestion charge on car pooling.

[34] For a general overview on traffic related externalities see Parry et al. (2007). For externalities related to car accidents see Edlin and Karaca-Mandic (2006)

examples of such pricing schedules in Section 4.4. Our results indicate that modern traffic management systems, such as those in Singapore or Sweden, could be improved in such a way by offering a menu of travel times with respective prices.

Requiring drivers to directly report their value of time may be impractical because such mechanisms might be hard to explain to people. We approach this issue by designing a mechanism in which it is a dominant strategy for each driver to report her value of travel time truthfully. Thus, independently of what a driver believes about other traffic participants, whether correct or wrong, it is optimal to reveal her information truthfully. While this deals with potentially wrong beliefs, drivers might misunderstand a price schedule that is too complex. If this is the case, a regulator would like to design a simpler price schedule. It may be easier to use solutions that allow drivers to choose from a simplified menu of prices and arrival times. For example people could be offered a choice of three categories: fast, normal and slow. In that case they would be charged a premium for choosing faster options. The trade-offs involved in finding good user interfaces, would need to be investigated further before such congestion pricing mechanisms are implemented.

Pricing schedules could be further simplified by relaxing the implementation requirements from dominant-strategy incentive compatibility to Bayesian incentive compatibility. The price schedule of an agent would then no longer depend on the realizations of other agents but only on their distribution. This would induce continuously increasing price schedules for each driver. A caveat is that for Bayesian incentive compatibility to work, drivers need to have correct beliefs about the distribution of values of travel time of other drivers.

When inducing price discrimination into a road-pricing scheme, a municipality has to deal with acceptability of a new system. In relation to a classic non-discriminating Pigouvian pricing scheme, the question who benefits and who does not depends on the distribution of values of travel time. Acceptance could be encouraged in our price scheme because it allows for cross-subsidization between drivers who pay different prices. In the main model, we set the price of the uncongetible alternative to be zero. But this is not necessary for incentives and a municipality could use the revenues from drivers of fast roads to subsidize uncongestible alternatives like public transport, i.e. to set a negative price for the slower road. As long as price differences are unaffected, such a cross subsidy would not affect the allocation in our mechanism.

Drivers who participate in our mechanism on a regular basis might have to pay different prices each time they use these roads. Drivers might not like this uncertainty over prices. This uncertainty is already a feature of many existing traffic pricing schemes, such as the one in Stockholm or Gothenburg, but it can be eliminated by selling tickets for multiple uses or a period of time at a price equal to the expected total transfer. The mechanism would then take the value of time of that driver as constant for the time the ticket is valid. Naturally, this would lead to inefficiencies when the valuation of time varies a lot.

One other potential concern is that congestion pricing mechanisms as envisioned here would provide too much information on citizens' travel behavior. However there are ways in which congestion pricing could be implemented without collecting detailed personal information. Charges for traveling in an autonomous vehicle could technically be depersonalized.

Congestion problems arise not only in the case of road or traffic-related pricing but in many other applications that are of interest to economists and policymakers. One application arises in the context of data routing. Internet service provider (ISP) allocate bandwidth among content providers (CP) such as video-call, email or streaming services. There is a congestion problem since more transmission capacity for one CP has a negative impact on the other CPs. Additionally, some CPs require a fast internet connection to offer their service, but this sensitivity of their service to the connection speed is likely to be private information and not observed by the ISP. VCG mechanisms might also be used to efficiently route data. In the context of data transmission, we cannot assume that each CP is identical in terms of how much bandwidth capacity is required for service provision. Thus CPs can have heterogeneous effects on the overall level of congestion. This is not captured in our model so far. Nevertheless, the qualitative features of our mechanism should continue to hold. In particular, it is doubtful that a single Pigouvian price is an efficient mechanism when the number of CPs is finite.

Effects similar to congestion are also prevalent in keyword search auctions on the internet, which have first been analyzed by Varian (2007) and Edelman et al. (2007). These papers assumed that the total number of clicks of an ad depended solely on the position on which the ad is shown. However the attention and hence the number of clicks an ad receives on a website likely depends on the total number of other ads displayed in the same impression. Hence advertisers might be willing to pay to ensure that fewer other advertisers are shown on the same impression. In practice however the type of externalities that arise in this context are likely to be more complex than those arising in the application to traffic congestion pricing. Jeziorski and Segal (2015) empirically analyze clicking behavior of consumers. They find that the number of times an ad is clicked depends to a significant and economically meaningful extent on the identity of other ads shown both in higher and lower positions.

Before applying the results of this paper in practice one would need to extend the model to cover more complex road networks and intertemporal issues. In these cases - with a finite number of drivers - characterizing efficient allocations in simple expressions is generally impossible, since it is a discrete optimization problem. Nonetheless, efficient allocations exist in more complex networks as well, especially when the number of drivers and roads are finite. Importantly, this is not a problem of existence or implementation. Computational methods can in principle find efficient allocations and once these allocations are known, VCG-type mechanisms can be used to implement them. This should not be a major obstacle for implementation since modern traffic control systems are already today very sophisticated.

In this model we focus on the interaction of congestion and route choice. Other congestion models also consider a bottleneck problem where traffic participants face the decision to drive now or later.[35]. Studying our mechanism in settings that incorporate both causes of congestion is an interesting direction of future research.

[35]For examples, see Vickrey (1967) or Arnott et al. (1990).

List of Tables

List of Figures

Bibliography

Adda, J., C. Dustmann, and K. Stevens (2017). The Career Costs of Children. *Journal of Political Economy 125*(2), 293–337.

Adsera, A. (2004). Changing Fertility Rates in Developed Countries. The Impact of Labor Market Institutions. *Journal of Population Economics 17*(1), 17–43.

Adsera, A. (2005). Vanishing Children: From High Unemployment to Low Fertility in Developed Countries. *American Economic Review 95*(2), 189–193.

Adsera, A. and A. Menendez (2011). Fertility Changes in Latin America in Periods of Economic Uncertainty. *Population Studies 65*(1), 37–56.

Aiyagari, S. R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *Quarterly Journal of Economics 109*(3), 659–684.

Akyol, M., M. Neugart, and S. Pichler (2013). Were the Hartz Reforms Responsible for the Improved Performance of the German Labour Market? *Economic Affairs 33*(1), 34–47.

Altonji, J. G. (1986). Intertemporal Substitution in Labor Supply: Evidence from Micro Data. *Journal of Political Economy 94*(3), S176–S215.

Anderson, T. W. and C. Hsiao (1981). Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association 76*(375), 598–606.

Angrist, J. D. and A. B. Krueger (1999). Empirical Strategies in Labor Economics. *Handbook of Labor Economics 3*, 1277–1366.

Antonczyk, D., B. Fitzenberger, and K. Sommerfeld (2010). Rising Wage Inequality, the Decline of Collective Bargaining, and the Gender Wage Gap. *Labour Economics 17*(5), 835–847.

Antonczyk, D., B. Fitzenberger, and K. Sommerfeld (2011). Anstieg der Lohnungleichheit, Rückgang der Tarifbindung und Polarisierung. *Journal for Labour Market Research (Zeitschrift für ArbeitsmarktForschung) 44*(1-2), 15–27.

Apps, P. and R. Rees (2004). Fertility, Taxation and Family Policy. *Scandinavian Journal of Economics 106*(4), 745–763.

Arellano, M. and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies 58*(2), 277–297.

Arellano, M. and O. Bover (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics 68*(1), 29–51.

Aretz, B., M. Arntz, and T. Gregory (2013). The Minimum Wage Affects Them All: Evidence on Employment Spillovers in the Roofing Sector. *German Economic Review 14*(3), 282–315.

Aretz, B., M. Arntz, T. Gregory, and C. Rammer (2012). Der Mindestlohn im Dachdeckerhandwerk: Auswirkungen auf Beschäftigung, Arbeitnehmerschutz und Wettbewerb. *Journal for Labour Market Research (Zeitschrift für ArbeitsmarktForschung) 45*(3-4), 233–256.

Arni, P., W. Eichhorst, N. Pestel, A. Spermann, and K. F. Zimmermann (2014). Der gesetzliche Mindestlohn in Deutschland: Einsichten und Handlungsempfehlungen aus der Evaluationsforschung. *Journal of Applied Social Science Studies (Schmollers Jahrbuch) 134*(2), 149–182.

Arnott, R., A. de Palma, and R. Lindsey (1990). Departure Time and Route Choice for the Morning Commute. *Transportation Research Part B: Methodological 24*(3), 209–228.

Arnott, R. and M. Kraus (1998). When are Anonymous Congestion Charges Consistent with Marginal Cost Pricing? *Journal of Public Economics 67*(1), 45–64.

Arnott, R., A. d. Palma, and R. Lindsey (1994). The Welfare Effects of Congestion Tolls with Heterogeneous Commuters. *Journal of Transport Economics and Policy 28*(2), 139–161.

Arroyo, C. R. and J. Zhang (1997). Dynamic Microeconomic Models of Fertility Choice: A Survey. *Journal of Population Economics 10*(1), 23–65.

Askitas, N. and K. F. Zimmermann (2009). Google Econometrics and Unemployment Forecasting. *Applied Economics Quarterly 55*(2), 107–120.

Auer, W. and N. Danzer (2016). Fixed-Term Employment and Fertility: Evidence from German Micro Data. *CESifo Economic Studies 62*(4), 595–623.

Autor, D. H. (2003). Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics 21*(1), 1–42.

Bachmann, R., T. K. Bauer, J. Kluve, S. Schaffner, and C. M. Schmidt (2008). *Mindestlöhne in Deutschland: Beschäftigungswirkungen und fiskalische Effekte.* RWI Materialien 43. Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Essen.

Baltagi, B. H., E. Bratberg, and T. H. Holmås (2005). A Panel Data Study of Physicians' Labor Supply: The Case of Norway. *Health Economics 14*(10), 1035–1045.

Bauer, T. K., J. Kluve, S. Schaffner, and C. M. Schmidt (2009). Fiscal Effects of Minimum Wages: An Analysis for Germany. *German Economic Review 10*(2), 224–242.

Baum, C., M. Schaffer, and S. Stillman (2016). IVREG2: Stata Module for Extended Instrumental Variables/2SLS and GMM Estimation (v10). Statistical Software Components, Boston College Department of Economics.

Becker, G. S. (1960). An Economic Analysis of Fertility. In *Demographic and Economic Change in Developed Countries*, pp. 209–240. Columbia University Press.

Bell, D. N. and D. G. Blanchflower (2013a). How to Measure Underemployment? *Peterson Institute for International Economics Working Paper 13-7*.

Bell, D. N. and D. G. Blanchflower (2013b). Underemployment in the UK Revisited. *National Institute Economic Review 224*(1), F8–F22.

Bell, L. A. and R. B. Freeman (2001). The Incentive for Working Hard: Explaining Hours Worked Differences in the US and Germany. *Labour Economics 8*(2), 181–202.

Benito, A. and J. Saleheen (2013). Labour Supply as a Buffer: Evidence from UK Households. *Economica 80*(320), 698–720.

Bernardi, L., A. Klärner, and H. Von der Lippe (2008). Job Insecurity and the Timing of Parenthood: A Comparison Between Eastern and Western Germany. *European Journal of Population/Revue Européenne de Démographie 24*(3), 287–313.

Bernstein, D. and I. El Sanhouri (1994). *Congestion Pricing with an Untolled Alternative*. Unpublished Paper. Massachusetts Institute of Technology `https://users.cs.jmu.edu/bernstdh/Web/its/papers/untolled-alternative.pdf` (Retrieved 22 July 2017).

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Bhaumik, S. K. and J. B. Nugent (2011). Real Options and Demographic Decisions: Empirical Evidence from East and West Germany. *Applied Economics 43*(21), 2739–2749.

Biewen, M. and A. Juhasz (2012). Understanding Rising Income Inequality in Germany, 1999/2000–2005/2006. *Review of Income and Wealth 58*(4), 622–647.

Block, M. K. and J. M. Heineke (1973). The Allocation of Effort under Uncertainty: The Case of Risk-averse Behavior. *Journal of Political Economy 81*(2), 376–385.

Bloom, D. E., D. Canning, and J. Sevilla (2001). Economic Growth and the Demographic Transition. *National Bureau of Economic Research Working Paper 8685*.

Blundell, R. and S. Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics 87*(1), 115–143.

Blundell, R., M. Graber, and M. Mogstad (2015). Labor Income Dynamics and the Insurance from Taxes, Transfers, and the Family. *Journal of Public Economics 127*, 58–73.

Blundell, R., L. Pistaferri, and I. Preston (2008, December). Consumption Inequality and Partial Insurance. *American Economic Review 98*(5), 1887–1921.

Blundell, R. and I. Preston (1998). Consumption Inequality and Income Uncertainty. *Quarterly Journal of Economics 113*(2), 603–640.

BMAS (2015). *Bestandsaufnahme: Einführung des allgemeinen Mindestlohnes in Deutschland.* Bundesministerium für Arbeit und Soziales, Berlin [Federal Ministry of Labour and Social Affais]. `http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/2015/bestandsaufnahme-mindestlohn.pdf` (Retrieved 22 July 2017).

Bonin, H. and R. Euwals (2001). Participation Behavior of East German Women after German Unification. *IZA Discussion Paper 413.*

Bönke, T., M. Giesecke, and H. Lüthen (2015). The Dynamics of Earnings in Germany: Evidence from Social Security Records. *Ruhr Economic Papers 582.*

Boockmann, B., R. Krumm, M. Neumann, and P. Rattenhuber (2013). Turning the Switch: An Evaluation of the Minimum Wage in the German Electrical Trade Using Repeated Natural Experiments. *German Economic Review 14*(3), 316–348.

Borjas, G. J. (1980). The Relationship Between Wages and Weekly Hours of Work: The Role of Division Bias. *Journal of Human Resources 15*(3), 409–423.

Bossler, M. (2016). Employment Expectations and Uncertainties ahead of the New German Minimum Wage. *IAB-Discussion Paper 3/2016.*

Bossler, M. and H.-D. Gerner (2016). Employment Effects of the New German Minimum Wage: Evidence from Establishment-level Micro Data. *IAB-Discussion Paper 10/2016.*

Boysen-Hogrefe, J. and D. Groll (2010). The German Labour Market Miracle. *National Institute Economic Review 214*, R38–R50.

Brenke, K. (2012). Geringe Stundenlöhne, lange Arbeitszeiten. *DIW-Wochenbericht 79*(21), 3–12.

Brown, C. (1999). Minimum Wages, Employment, and the Distribution of Income. *Handbook of Labor Economics 3*, 2101–2163.

Brown, T. M. (1952). Habit Persistence and Lags in Consumer Behaviour. *Econometrica 20*(3), 355–371.

Bruckmeier, K. and J. Wiemers (2014). Begrenzte Reichweite: Die meisten Aufstocker bleiben trotz Mindestlohn bedürftig. *IAB-Kurzbericht 7/2014*.

Bruckmeier, K. and J. Wiemers (2016). Entwicklung der Zahl der Aufstocker nach Einführung des Mindestlohns im Jahr 2015. *IAB Aktuelle Berichte 10/2016*.

Bryan, M. L. (2007). Free to Choose? Differences in the Hours Determination of Constrained and Unconstrained Workers. *Oxford Economic Papers 59*(2), 226–252.

Burda, M. C. and J. Hunt (2011). What Explains the German Labor Market Miracle in the Great Recession? *Brookings Papers on Economic Activity 42*(1), 273–335.

Card, D. (1990). The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial & Labor Relations Review 43*(2), 245–257.

Card, D. (1992). Using Regional Variation in Wages to Measure the Effects of the Federal Minimum Wage. *Industrial & Labor Relations Review 46*(1), 22–37.

Card, D., J. Heining, and P. Kline (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics 128*(3), 967–1015.

Card, D. and A. B. Krueger (2015). *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton University Press.

Carlin, W. and D. Soskice (2009). German Economic Performance: Disentangling the Role of Supply-Side Reforms, Macroeconomic Policy and Coordinated Economy Institutions. *Socio-Economic Review 7*(1), 67–99.

Carroll, C. D., K. E. Dynan, and S. D. Krane (2003, August). Unemployment Risk and Precautionary Wealth: Evidence from Households' Balance Sheets. *Review of Economics and Statistics 85*(3), 586–604.

Carroll, C. D. and A. A. Samwick (1998, August). How Important Is Precautionary Saving? *Review of Economics and Statistics 80*(3), 410–419.

Clarke, E. H. (1971). Multipart Pricing of Public Goods. *Public Choice 11*(1), 17–33.

Concas, S. and A. Kolpakov (2009). *Synthesis of Research on Value of Time and Value of Reliability*. Center for Urban Transportation Research, National Center for Transit Research `http://ntl.bts.gov/lib/31000/31500/31535/FDOT_BD549_46_rpt.pdf` (Retrieved 22 July 2017).

Conrad, C., M. Lechner, and W. Werner (1996). East German Fertility after Unification: Crisis or Adaptation? *Population and Development Review 22*(2), 331–358.

Cullen, J. B. and J. Gruber (2000). Does Unemployment Insurance Crowd out Spousal Labor Supply? *Journal of Labor Economics 18*(3), 546–572.

Cunha, F. and J. Heckman (2008). A New Framework for the Analysis of Inequality. *Macroeconomic Dynamics 12*(S2), 315–354.

Cunha, F. and J. Heckman (2016). Decomposing Trends in Inequality in Earnings into Forecastable and Uncertain Components. *Journal of Labor Economics 34*(S2), S31–S65.

Cunha, F., J. Heckman, and S. Navarro (2005). The 2004 Hicks Lecture: Separating Uncertainty from Heterogeneity in Life Cycle Earnings. *Oxford Economic Papers 57*(2), 191–261.

de Lange, M., M. H. Wolbers, M. Gesthuizen, and W. C. Ultee (2014). The Impact of Macro- and Micro-Economic Uncertainty on Family Formation in the Netherlands. *European Journal of Population 30*(2), 161–185.

de Palma, A. and R. Lindsey (2011). Traffic Congestion Pricing Methodologies and Technologies. *Transportation Research Part C: Emerging Technologies 19*(6), 1377–1399.

Del Bono, E., A. Weber, and R. Winter-Ebmer (2012). Clash of Career and Family: Fertility Decisions after Job Displacement. *Journal of the European Economic Association 10*(4), 659–683.

Del Bono, E., A. Weber, and R. Winter-Ebmer (2015). Fertility and Economic Instability: The Role of Unemployment and Job Displacement. *Journal of Population Economics 28*(2), 463–478.

Deutsche Bundesbank (2017). Demografischer Wandel, Zuwanderung und das Produktionspotenzial der deutschen Wirtschaft. *Deutsche Bundesbank Monatsbericht 69*(4), 37–50.

Dolado, J., F. Kramarz, S. Machin, A. Manning, D. Margolis, and C. Teulings (1996). The Economic Impact of Minimum Wages in Europe. *Economic Policy 11*(23), 317–372.

Downs, A. (1993). Point of View: Implementing Peak-hour Road Pricing at Fullscale: Finding Solutions to Practical Problems. *TR News 167*, 7–9.

Dustmann, C., B. Fitzenberger, U. Schönberg, and A. Spitz-Oener (2014). From Sick Man of Europe to Economic Superstar: Germany's Resurgent Economy. *Journal of Economic Perspectives 28*(1), 167–188.

Eaton, J. and H. S. Rosen (1980a). Labor Supply, Uncertainty, and Efficient Taxation. *Journal of Public Economics 14*(3), 365–374.

Eaton, J. and H. S. Rosen (1980b). Optimal Redistributive Taxation and Uncertainty. *Quarterly Journal of Economics 95*(2), 357–364.

Eckey, H.-F., B. Schwengler, and M. Türck (2007). Vergleich von deutschen Arbeitsmarktregionen. *IAB-Discussion Paper 3/2007*.

Edelman, B., M. Ostrovsky, and M. Schwarz (2007). Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords. *American Economic Review 97*(1), 242–259.

Edlin, A. S. and P. Karaca-Mandic (2006). The Accident Externality from Driving. *Journal of Political Economy 114*(5), 931–955.

Ehrentraut, O., M. Matuschke, S. Schmutz, and R. Schüssler (2011). *Fiskalische Effekte eines gesetzlichen Mindestlohns*. WISO Diskurs, Friedrich-Ebert-Stiftung, Bonn.

Engen, E. M. and J. Gruber (2001). Unemployment Insurance and Precautionary Saving. *Journal of Monetary Economics 47*(3), 545–579.

Euwals, R. (2005). The Predictive Value of Subjective Labour Supply Data: A Partial-Adjustment Model with Measurement Error. *Empirical Economics 30*(2), 309–329.

Feenberg, D. and E. Coutts (1993). An Introduction to the TAXSIM Model. *Journal of Policy Analysis and Management 12*(1), 189–194.

Feigenbaum, J. and G. Li (2015). Household Income Uncertainties over Three Decades. *Oxford Economic Papers 67*(4), 963–986.

Feyrer, J., B. Sacerdote, and A. D. Stern (2008). Will the Stork Return to Europe and Japan? Understanding Fertility within Developed Nations. *Journal of Economic Perspectives 22*(3), 3.

Fiori, F., F. Rinesi, A. Pinnelli, and S. Prati (2013). Economic Insecurity and the Fertility Intentions of Italian Women with One Child. *Population Research and Policy Review 32*(3), 373–413.

Fitzenberger, B. (2009). Anmerkungen zur Mindestlohndebatte: Elastizitäten, Strukturparameter und Topfschlagen. *Journal for Labour Market Research (Zeitschrift für ArbeitsmarktForschung) 42*(1), 85–92.

Flodén, M. (2006). Labour Supply and Saving Under Uncertainty. *Economic Journal 116*(513), 721–737.

Fossen, F. M. and D. Rostam-Afschar (2013). Precautionary and Entrepreneurial Savings: New Evidence from German Households. *Oxford Bulletin of Economics and Statistics 75*(4), 528–555.

Friedman, M. (1957). The Permanent Income Hypothesis. In *A Theory of the Consumption Function*, pp. 20–37. Princeton University Press.

Frings, H. (2013). The Employment Effect of Industry-Specific, Collectively Bargained Minimum Wages. *German Economic Review 14*(3), 258–281.

Fuchs-Schündeln, N. and M. Schündeln (2005). Precautionary Savings and Self-Selection: Evidence from the German Reunification "Experiment". *Quarterly Journal of Economics 120*(3), 1085–1120.

Furuhata, M., M. Dessouky, F. Ordóñez, M.-E. Brunet, X. Wang, and S. Koenig (2013). Ridesharing: The state-of-the-art and future directions. *Transportation Research Part B: Methodological 57*, 28–46.

Galor, O. and D. Weil (1996). The Gender Gap, Fertility and Growth. *American Economic Review 86*(3), 374–387.

Galor, O. and D. Weil (2000). Population, Technology and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review 90*(4), 806–826.

Garloff, A. (2010). Minimum Wages, Wage Dispersion and Unemployment in Search Models. A Review. *Journal for Labour Market Research (Zeitschrift für Arbeitsmarkt-Forschung) 43*(2), 145–167.

Garloff, A. (2016). Side Effects of the New German Minimum Wage on (Un-)employment: First Evidence from Regional Data. *IAB-Discussion Paper 31/2016*.

Gartner, H. and S. Klinger (2010). Verbesserte Institutionen für den Arbeitsmarkt in der Wirtschaftskrise. *Wirtschaftsdienst 90*(11), 728–734.

Gebel, M. and J. Giesecke (2009). Ökonomische Unsicherheit und Fertilität. Die Wirkung von Beschäftigungsunsicherheit und Arbeitslosigkeit auf die Familiengründung in Ost- und Westdeutschland [Economic Insecurity and Fertility. The Effects of Job Insecurity and Unemployment on Family Formation in East and West Germany]. *Zeitschrift für Soziologie 38*(5), 399–417.

Gernandt, J. and F. Pfeiffer (2007). Rising Wage Inequality in Germany. *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik) 227*(4), 358–380.

Gibbard, A. (1973). Manipulation of Voting Schemes: A General Result. *Econometrica 41*(4), 587–601.

Goldin, C. and L. F. Katz (2000). Career and Marriage in the Age of the Pill. *American Economic Review 90*(2), 461–465.

Goldin, C. and L. F. Katz (2002). The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. *Journal of Political Economy 110*(4), 730–770.

Goldstein, J. R. and M. Kreyenfeld (2011). Has East Germany Overtaken West Germany? Recent Trends in Order-Specific Fertility. *Population and Development Review 37*(3), 453–472.

Goldstein, J. R., T. Sobotka, and A. Jasilioniene (2009). The End of "Lowest-Low" Fertility? *Population and Development Review 35*(4), 663–699.

Google Trends (2017). *Search Queries for "Mindestlohn", originating from Germany, 2013-2015.* Google Inc. `https://trends.google.de/trends/explore?date=2013-01-01%202015-12-31&geo=DE&q=Mindestlohn` (Retrieved 22 July 2017).

Grabka, M. M. and J. Goebel (2014). Reduction in Income Inequality Faltering. *DIW Economic Bulletin 4*(1), 16–25.

Greenwood, J., A. Seshadri, and G. Vandenbroucke (2005). The Baby Boom and Baby Bust. *American Economic Review 95*(1), 183–207.

Gregory, T. (2014). When the Minimum Wage Bites Back: Quantile Treatment Effects of a Sectoral Minimum Wage in Germany. *ZEW Discussion Paper 14-133.*

Groll, D. (2015). Mindestlohn: Erste Anzeichen für Jobverluste. *Wirtschaftsdienst 95*(6), 439–440.

Groves, T. (1973). Incentives in Teams. *Econometrica 41*(4), 617–631.

Haghpanah, N., N. Immorlica, V. Mirrokni, and K. Munagala (2013). Optimal Auctions with Positive Network Externalities. *ACM Transactions on Economics and Computation 1*(2), 13–22.

Hartwick, J. M. (2000). Labor Supply under Wage Uncertainty. *Economics Letters 68*(3), 319–325.

Harvey, A. C. (1976). Estimating Regression Models with Multiplicative Heteroscedasticity. *Econometrica 44*(3), 461–465.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47*(1), 153–161.

Heckman, J. J. and B. Singer (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica 52*(2), 271–320.

Heckman, J. J. and J. R. Walker (1990a). Estimating Fecundability from Data on Waiting Times to First Conception. *Journal of the American Statistical Association 85*(410), 283–294.

Heckman, J. J. and J. R. Walker (1990b). The Relationship between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data. *Econometrica 58*(6), 1411–1441.

Heckman, J. J. and J. R. Walker (1990c). The Third Birth in Sweden. *Journal of Population Economics 3*(4), 235–275.

Henzel, S. and K. Engelhardt (2014). Arbeitsmarkteffekte des flächendeckenden Mindestlohns in Deutschland–eine Sensitivitätsanalyse. *Ifo Schnelldienst 67*(10), 23–29.

Hetschko, C., R. Schöb, and T. Wolf (2016). Income Support, (Un-)Employment and Well-Being. *CESifo Working Paper No. 6016*.

Hofmann, B. and K. Hohmeyer (2013). Perceived Economic Uncertainty and Fertility: Evidence from a Labor Market Reform. *Journal of Marriage and Family 75*(2), 503–521.

Hofmann, B., M. Kreyenfeld, and A. Uhlendorff (2017). Job Displacement and First Birth Over the Business Cycle. *Demography 54*(3), 933–959.

Holtz-Eakin, D., W. Newey, and H. S. Rosen (1988). Estimating Vector Autoregressions with Panel Data. *Econometrica 56*(6), 1371–1395.

Hondroyiannis, G. (2010). Fertility Determinants and Economic Uncertainty: An Assessment using European Panel Data. *Journal of Family and Economic Issues 31*(1), 33–50.

Hotz, V. J., J. A. Klerman, and R. J. Willis (1997). The Economics of Fertility in Developed Countries. *Handbook of Population and Family Economics 1*(Part 1), 275–347.

Hryshko, D. (2012, 07). Labor Income Profiles are not Heterogeneous: Evidence from Income Growth Rates. *Quantitative Economics 3*(2), 177–209.

Hubbard, R. G., J. Skinner, and S. P. Zeldes (1995, April). Precautionary Saving and Social Insurance. *Journal of Political Economy 103*(2), 360–399.

Huggett, M. (1993). The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies. *Journal of Economic Dynamics and Control 17*(5), 953–969.

Jehiel, P., B. Moldovanu, and E. Stacchetti (1996). How (Not) to Sell Nuclear Weapons. *American Economic Review 86*(4), 814–829.

Jehiel, P., B. Moldovanu, and E. Stacchetti (1999). Multidimensional Mechanism Design for Auctions with Externalities. *Journal of Economic Theory 85*(2), 258–293.

Jenkins, S. P. (1995). Easy Estimation Methods for Discrete-time Duration Models. *Oxford Bulletin of Economics and Statistics 57*(1), 129–136.

Jessen, R., D. Rostam-Afschar, and V. Steiner (2017). Getting the Poor to Work: Three Welfare Increasing Reforms for a Busy Germany. *Public Finance Analysis (FinanzArchiv) 73*(1), 1–41.

Jeziorski, P. and I. R. Segal (2015). What Makes Them Click: Empirical Analysis of Consumer Demand for Search Advertising. *American Economic Journal: Microeconomics 7*(3), 24–53.

Keane, M. P. (2011). Labor Supply and Taxes: A Survey. *Journal of Economic Literature 49*(4), 961–1075.

Kimball, M. S. (1990). Precautionary Saving in the Small and in the Large. *Econometrica 58*(1), 53–73.

Kind, M. and J. Kleibrink (2013). Sooner or Later–Economic Insecurity and the Timing of First Birth. *Ruhr Economic Paper 422*.

Kleiner, A., B. Nebel, and V. A. Ziparo (2011). A Mechanism for Dynamic Ride Sharing based on Parallel Auctions. In *IJCAI*, Volume 11, pp. 266–272.

Knabe, A. and R. Schöb (2009). Minimum Wage Incidence: The Case for Germany. *Public Finance Analysis (FinanzArchiv) 65*(4), 403–441.

Knabe, A., R. Schöb, and M. Thum (2014). Der flächendeckende Mindestlohn. *Perspektiven der Wirtschaftspolitik 15*(2), 133–157.

Knabe, A., R. Schöb, M. Thum, and M. Weber (2016). *Mindestlohn ohne Nebenwirkungen?* Ökonomenstimme.org. `http://www.oekonomenstimme.org/artikel/2016/07/mindestlohn-ohne-nebenwirkungen/` (Retrieved 22 July 2017).

Knight, F. H. (1921). Risk, Uncertainty and Profit.

Knight, F. H. (1924). Some Fallacies in the Interpretation of Social Cost. *Quarterly Journal of Economics 38*(4), 582–606.

Kreyenfeld, M. (2010). Uncertainties in Female Employment Careers and the Postponement of Parenthood in Germany. *European Sociological Review 26*(3), 351–366.

Kuhn, P. and F. Lozano (2008, 04). The Expanding Workweek? Understanding Trends in Long Work Hours among U.S. Men, 1979-2006. *Journal of Labor Economics 26*(2), 311–343.

Land Transport Authority (2016). *LTA to Launch Autonoumous Mobility-on-Demand Trials*. `https://www.lta.gov.sg/apps/news/page.aspx?c=2&id=73057d63-d07a-4229-87af-f957c7f89a27` (Retrieved 22 July 2017).

Land Transport Authority (2017). *Electronic Road Pricing (ERP).* `https://www.lta.gov.sg/content/ltaweb/en/roads-and-motoring/managing-traffic-and-congestion/electronic-road-pricing-erp.html` (Retrieved 22 July 2017).

Leape, J. (2006). The London Congestion Charge. *Journal of Economic Perspectives 20*(4), 157–176.

Low, H. (2005, October). Self-Insurance in a Life-Cycle Model of Labor Supply and Savings. *Review of Economic Dynamics 8*(4), 945–975.

Lundborg, P., E. Plug, and A. W. Rasmussen (2017). Can Women Have Children and a Career? IV Evidence from IVF Treatments. *American Economic Review 107*(6), 1611–1637.

Lusardi, A. (1998). On the Importance of the Precautionary Saving Motive. *American Economic Review 88*(2), 449–453.

MaCurdy, T. E. (1981). An Empirical Model of Labor Supply in a Life-Cycle Setting. *Journal of Political Economy 89*(6), 1059–1085.

Manning, A. (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton University Press.

Marcet, A., F. Obiols-Homs, and P. Weil (2007). Incomplete Markets, Labor Supply and Capital Accumulation. *Journal of Monetary Economics 54*(8), 2621 – 2635.

Mastrogiacomo, M. and R. Alessie (2014). The Precautionary Savings Motive and Household Savings. *Oxford Economic Papers 66*(1), 164–187.

Mayet, J. and M. Hansen (2000). Congestion Pricing with Continuously Distributed Values of Time. *Journal of Transport Economics and Policy 34*(3), 359–369.

Meghir, C. and L. Pistaferri (2011). Earnings, Consumption and Life Cycle Choices. *Handbook of Labor Economics 4*, 773–854.

Menezes, C. F. and X. H. Wang (2005). Duality and the Slutsky Income and Substitution Effects of Increases in Wage Rate Uncertainty. *Oxford Economic Papers 57*(3), 545–557.

Möller, J. (2016). Lohnungleichheit–Gibt es eine Trendwende? *IAB-Discussion Paper 9/2016.*

Möller, J. and M. König (2008). Mindestlohneffekte des Entsendegesetzes?: Eine Mikrodatenanalyse für die deutsche Bauwirtschaft. *Journal for Labour Market Research (Zeitschrift für ArbeitsmarktForschung) 41*(2/3), 327–346.

Montero, J.-P. (2008). A Simple Auction Mechanism for the Optimal Allocation of the Commons. *American Economic Review 98*(1), 496–518.

Morrison, S. A. (1986). A Survey of Road Pricing. *Transportation Research Part A: General 20*(2), 87–97.

Müller, K.-U. (2009). How Robust Are Simulated Employment Effects of a Legal Minimum Wage in Germany? *DIW Discussion Papers 900*.

Müller, K.-U. (2010). Employment Effects of a Sectoral Minimum Wage in Germany: Semi-Parametric Estimations from Cross-Sectional Data. *DIW Discussion Papers 1061*.

Müller, K.-U. and V. Steiner (2009). Would a Legal Minimum Wage Reduce Poverty? A Micro-simulation Study for Germany. *Journal of Income Distribution 18*(3), 131–151.

Müller, K.-U. and V. Steiner (2011). Beschäftigungswirkungen von Lohnsubventionen und Mindestlöhnen–Zur Reform des Niedriglohnsektors in Deutschland. *Journal for Labour Market Research (Zeitschrift für ArbeitsmarktForschung) 44*(1-2), 181–195.

Myerson, R. B. (1981). Optimal Auction Design. *Mathematics of Operations Research 6*(1), 58–73.

Navarro, S. (2011). Using Observed Choices to Infer Agent's Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance. *CIBC Working Paper 2011-8*.

Navarro, S. and J. Zhou (2017). Identifying Agent's Information sets: An Application to a Lifecycle Model of Schooling, Consumption and Labor Supply. *Review of Economic Dynamics 25*, 58–92.

Neumark, D. and W. L. Wascher (2008). *Minimum Wages*. MIT Press, Cambridge.

Ochel, W. (2005). Hartz IV-Welfare to Work in Germany. *CESifo DICE Report 2*, 18–25.

Parker, J. A. and B. Preston (2005). Precautionary Saving and Consumption Fluctuations. *American Economic Review 95*(4), 1119–1143.

Parker, S. C., Y. Belghitar, and T. Barmby (2005). Wage Uncertainty and the Labour Supply of Self-Employed Workers. *Economic Journal 115*(502), C190–C207.

Parry, I. W., M. Walls, and W. Harrington (2007). Automobile Externalities and Policies. *Journal of Economic Literature 45*(2), 373–399.

Pigou, A. C. (1920). *The Economics of Welfare*. Macmillan.

Pijoan-Mas, J. (2006). Precautionary Savings or Working Longer Hours? *Review of Economic Dynamics 9*(2), 326–352.

Pistaferri, L. (2003). Anticipated and Unanticipated Wage Changes, Wage Risk, and Intertemporal Labor Supply. *Journal of Labor Economics 21*(3), 729–754.

Prud'homme, R. and J. P. Bocarejo (2005). The London Congestion Charge: A Tentative Economic Appraisal. *Transport Policy 12*(3), 279–287.

Rabe-Hesketh, S., A. Skrondal, and A. Pickles (2005). Maximum Likelihood Estimation of Limited and Discrete Dependent Variable Models with Nested Random Effects. *Journal of Econometrics 128*(2), 301–323.

Rattenhuber, P. (2014). Building the Minimum Wage: The Distributional Impact of Germany's First Sectoral Minimum Wage on Wages and Hours Across Different Wage Bargaining Regimes. *Empirical Economics 46*(4), 1429–1446.

Robins, P. K. and R. W. West (1980). Labor Supply Response Over Time. *Journal of Human Resources 15*(4), 524–544.

Roodman, D. (2009). How to do xtabond2: An Introduction to Difference and System GMM in Stata. *Stata Journal 9*(1), 86–136.

Rostam-Afschar, D., R. Jessen, and S. Schmitz (2016). How Important is Precuationary Labor Supply? *FU Berlin - School of Business & Economics Discussion Paper 2016/10*.

Rürup, B. and D. Heilmann (2012). Arbeitsmarktreformen: Was noch zu tun bleibt. *Wirtschaftsdienst 92*(5), 339–344.

Saks, R. E. and S. H. Shore (2005). Risk and Career Choice. *B.E. Journal of Economic Analysis & Policy 5*(1), 1–45.

Schäfer, H. and J. Schmidt (2012). *Der Niedriglohnsektor in Deutschland: Entwicklung, Struktur und individuelle Erwerbsverläufe*. Institut der deutschen Wirtschaft Köln Medien GmbH.

Schmitt, C. (2012). Labour Market Integration, Occupational Uncertainty, and Fertility Choices in Germany and the UK. *Demographic Research 26*(12), 253–292.

Schrank, D., B. Eisele, T. Lomax, and J. Bak (2015). *2015 Urban Mobility Scorecard*. Texas A&M Transportation Institute and INRIX.

Shephard, A. (2009). FORTAX: Reference Manual. Technical report, University College London.

Small, K. A. (2012). Valuation of Travel Time. *Economics of Transportation 1*(1), 2–14.

Small, K. A., C. Winston, and J. Yan (2005). Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability. *Econometrica 73*(4), 1367–1382.

Small, K. A. and J. Yan (2001). The Value of "Value Pricing" of Roads: Second-best Pricing and Product Differentiation. *Journal of Urban Economics 49*(2), 310–336.

Sommer, K. (2016). Fertility Choice in a Life Cycle Model with Idiosyncratic Uninsurable Earnings Risk. *Journal of Monetary Economics 83*, 27–38.

SPD (2013). *Das Wir entscheidet. Das Regierungsprogramm 2013-2017.* Sozialdemokratische Partei Deutschlands, Berlin. `https://www.spd.de/fileadmin/Dokumente/Beschluesse/Bundesparteitag/20130415_regierungsprogramm_2013_2017.pdf` (Retrieved 22 July 2017).

Statistik der Bundesagentur für Arbeit (2016a). *Arbeitsmarkt in Zahlen, Erwerbsfähige Leistungsberechtigte (ELB).* Bundesagentur für Arbeit, Nürnberg [Federal Employment Agency, Nuremberg].

Statistik der Bundesagentur für Arbeit (2016b). *Arbeitsmarkt in Zahlen, Sozialversicherungspflichtig und geringfügig entlohnte Beschäftigte.* Bundesagentur für Arbeit, Nürnberg [Federal Employment Agency, Nuremberg].

Statistik der Bundesagentur für Arbeit (2016c). *Arbeitsmarkt in Zahlen, Verteilung der sozialversicherungspflichtig Vollzeitbeschäftigten der Kerngruppe nach Klassen monatlicher Bruttoarbeitsentgelte.* Bundesagentur für Arbeit, Nürnberg [Federal Employment Agency, Nuremberg].

Statistisches Bundesamt (2016). Geburtenziffer 2015: Erstmals seit 33 Jahren bei 1,50 Kindern je Frau. In *Pressemitteilung vom 17. Oktober 2016 - 373/16*. Wiesbaden: Statistisches Bundesamt.

Steimetz, S. S. and D. Brownstone (2005). Estimating Commuters Value of Time with Noisy Data: A Multiple Imputation Approach. *Transportation Research Part B: Methodological 39*(10), 865–889.

Steiner, V., K. Wrohlich, P. Haan, and J. Geyer (2012). Documentation of the Tax-Benefit Microsimulation Model STSM: Version 2012. *Deutsches Institut für Wirtschaftsforschung, Data Documentation 63*.

Stewart, M. B. and J. K. Swaffield (1997). Constraints on the Desired Hours of Work of British Men. *Economic Journal 107*(441), 520–535.

Sutela, H. (2012). Temporary Jobs and First Child Fertility in Finland. *Community, Work & Family 15*(4), 425–450.

The Swedish Transport Agency (2017). *Congestion Taxes in Stockholm and Gothenburg.* `http://www.transportstyrelsen.se/en/road/Congestion-taxes-in-Stockholm-and-Goteborg/` (Retrieved 22 July 2017).

Transport for London (2017). *Congestion Charge*. `http://www.tfl.gov.uk/modes/driving/congestion-charge` (Retrieved 22 July 2017).

Tsekeris, T. and S. Voß (2009). Design and Evaluation of Road Pricing: State-of-the-art and Methodological Advances. *NETNOMICS: Economic Research and Electronic Networking 10*(1), 5–52.

Van de Kaa, D. J. (1997). Options and Sequences: Europe's Demographic Patterns. *Journal of the Australian Population Association 14*(1), 1–30.

Varian, H. R. (2007). Position Auctions. *International Journal of Industrial Organization 25*(6), 1163–1178.

Verhoef, E. T. (2000). The Implementation of Marginal External Cost Pricing in Road Transport. *Papers in Regional Science 79*(3), 307–332.

Verhoef, E. T., P. Nijkamp, and P. Rietveld (1996). Second-best Congestion Pricing: The Case of an Untolled Alternative. *Journal of Urban Economics 40*(3), 279–302.

Verhoef, E. T. and K. A. Small (2004). Product Differentiation on Roads: Constrained Congestion Pricing with Heterogeneous Users. *Journal of Transport Economics and Policy 38*(1), 127–156.

Vickrey, W. S. (1961). Counterspeculation, Auctions, and Competitive Sealed Tenders. *Journal of Finance 16*(1), 8–37.

Vickrey, W. S. (1967). Optimization of Traffic and Facilities. *Journal of Transport Economics and Policy 1*(2), 123–136.

Vickrey, W. S. (1969). Congestion Theory and Transport Investment. *American Economic Review 59*(2), 251–260.

Vickrey, W. S. (1993). Point of View: Principles and Applications of Congestion Pricing. *TR News 167*, 4–5.

Vickrey, W. S. and C. Sharp (1968). Congestion Charges and Welfare. *Journal of Transport Economics and Policy 2*(1), 107–125.

vom Berge, P., S. Kaimer, S. Copestake, J. Eberle, W. Klosterhuber, J. Krüger, S. Trenkle, and V. Zakrocki (2016). Arbeitsmarktspiegel: Entwicklungen nach Einführung des Mindestlohns (Ausgabe 1). Technical report, Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg].

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements. *Schmollers Jahrbuch 127*(1), 139–169.

Werding, M. (2014). Demographischer Wandel und öffentliche Finanzen. Langfrist-Projektionen 2014 - 2060 unter besonderer Berücksichtigung des Rentenpakets der Bundesregierung. *Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung. Arbeitspapier 1/2014.*

Windmeijer, F. (2005). A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators. *Journal of Econometrics 126*(1), 25–51.

Witte, J. C. and G. G. Wagner (1995). Declining Fertility in East Germany after Unification: A Demographic Response to Socioeconomic Change. *Population and Development Review 21*(2), 387–397.

Wood, J., K. Neels, and J. Vergauwen (2016). Economic and Institutional Context and Second Births in Seven European Countries. *Population Research and Policy Review 35*(3), 305–325.

Yang, H., W. Xu, and Q. Meng (2005). A Sequential Experimental Approach for Analyzing Second-best Road Pricing with Unknown Demand Functions. In H. Mahmassani (Ed.), *Transportation and Traffic Theory. Flow, Dynamics and Human Interaction. 16th International Symposium on Transportation and Traffic Theory*, pp. 23–42.

Summary

This dissertation consists in four essays about a wide range of topics and policies, ranging from the impact of wage and income risk on fertility or labor supply, over the evaluation of a new minimum wage to design features for optimal traffic regulation. While these studies differ in key aspects and methods, they have in common that uncertainty plays a crucial role in understanding the impact of the respective and related policies.

The first chapter, *Is Our Income Too Risky to Have A(nother) Baby? Evidence from German Micro Data*, studies women's fertility transitions and the timing of fertility, and how it is affected by household income, the female earnings potential, and the associated uninsured, idiosyncratic measures of risk. Founding a family or having another baby are rather fundamental decisions because they have an irreversible and permanent character. Thus, if households experience increases in the riskiness of their income sources, they might wish to postpone family formation, for instance in order to accumulate savings to insure themselves against negative shocks. Additionally, the prospective parents might want to work more to establish themselves on the labor market and resolve some of the uncertainty about the future income trajectory.

Income risk is an important concept in economics to describe, model and analyze inter-temporal decisions. It has a prominent role in the economic literature on the permanent income hypothesis with precautionary saving, but is also related to other economic areas such as labor supply, as will be shown in the second chapter of this dissertation. Empirically, income risk has been volatile and generally increasing since the 1970s in Germany and elsewhere (Bönke et al., 2015; Blundell et al., 2015). Thus, the question of the first chapter is whether there is a *causal* link between income risk and changes in fertility patterns.

The first chapter is an empirical investigation using data from the German Socio-Economic Panel for the years 1984 to 2014. Fertility is modeled as the time it takes to have the first (or another) child and estimated using discrete time duration models, taking into account unobserved heterogeneity as proposed by Heckman and Walker (1990a,b,c). The hazard model is augmented by the female wage rate, household income and measures of risk inherent in these two sources of income, as well as a rich set of sociodemographic covariates of both partners. The income and risk measures are based on the permanent income potential rather than the actually observed income and wages. This allows us to abstract from women's actual labor supply decisions. The potential wage rate captures the women's substitution decision between time spent working and time spent on childrearing activities. In order to study changing patterns of fertility, the results are differentiated with respect to two birth cohorts, born between 1960 and 1974 and between 1975 to 1989, and whether women hold a college degree or not.

The results do not suggest any significant effects on the transition to the first child for the cohort born between 1960 and 1974. For the younger cohort however, those born between 1975 and 1989, female wage risk leads to significantly longer time to the first child. Thus, the results indicate that the observed shift to family formation later in life and the increasing share of childless women could be explained by women's responses to income risk. For the transition to the second child, women from the older cohort with higher income opportunities tend to have shorter spacing between the first and the second child. This could be explained by difficulties returning to work after the first child. There is no significant effect for the younger cohort; however, for these women, the riskiness of the male income significantly reduces the probability of having a second child. We do not find any striking effects for the transition to the third child, which only rarely occurs in Germany. These results are mainly driven by low educated women.

The second chapter, *How Important is Precautionary Labor Supply?*, also studies the impact of wage risk, but this time on the labor supply decisions of German prime-age men. The question is, whether individuals chose to work longer hours in order to self-insure against their wage risk and the possibility of a negative wage shock. Thus, this chapter quantifies the importance of *precautionary labor supply* defined as the difference between hours supplied in the presence of wage risk and hours under perfect foresight.

Economic theory suggests that individuals might respond to higher wage uncertainty by increasing hours of work (Block and Heineke, 1973; Eaton and Rosen, 1980a,b; Hartwick, 2000) and that individuals facing uncertainty work more at the beginning of work life in order to accumulate savings (Pistaferri, 2003; Low, 2005; Flodén, 2006). However, the empirical relevance of precautionary labor supply is an open question. Thus, the second chapter quantifies the importance of wage risk to explain the hours of work of married men in order to assess the empirical relevance of precautionary labor supply. From a policy perspective a thorough understanding of labor supply incentives is key, e.g., for the design of the tax and transfer system or the unemployment insurance. Precautionary labor supply could also explain differences in hours worked across occupations or why self-employed work more hours than employees for a given wage.

Also the second chapter is an exploratory empirical analysis based on data from the German Socio-Economic Panel. The analysis is based on a dynamic labor supply model, allowing for partial adjustment of hours worked. The measure of wage risk is the standard deviation of past hourly individual marginal net wages, which are calculated using the microsimulation model STSM (see Steiner et al., 2012). The wage risk measure is similar to the one used in Parker et al. (2005), who only study the labor supply of the self-employed using data from the Panel Study of Income Dynamics (PSID). The second chapter extends this analysis to employees using data from the German Socio-economic Panel for 2001 to 2012. To overcome potential endogeneity issues, wages are instrumented with lagged labor income. The analysis controls for a rich set of variables including unemployment probability calculated similarly as in Carroll et al. (2003).

The results show that wage risk has a statistically significant positive impact on hours worked. Workers choose about 2.5% of their hours of work or one week per year to shield against wage shocks. Precautionary labor supply is particularly important for the self-employed, a group that faces average wage risks substantially above the sample mean. This group works 5.53% of their hours because of the precautionary motive. There are no significant effects for civil servants, which is in line with previous studies. If all workers faced the same risk as the median civil servant, hours worked would decrease on average by 1% in the long run. Precautionary labor supply is also economically important. Considering a person who works 42 hours per week and a typical net wage rate of € 13, precautionary labor supply amounts to about € 710 per year.

The third chapter, *The Effects of Germany's New Minimum Wage on Employment and Welfare Dependency*, deals with an actual public policy, Germany's new statutory minimum wage of € 8.50 per hour. The questions are whether the minimum wage caused job losses, or whether the introduction of the new wage floor reduced welfare dependency by lifting the income of the working poor. It was often argued that a minimum wage would be necessary to supplement earnings in the growing low-wage sector and to cushion the large-scale labor market reforms of the early 2000s, the so-called Hartz reforms. A key target group of the minimum wage are households that receive supplementary welfare benefits (unemployment benefit II, *UBII*) while working, the so-called *Aufstocker*. The proponents of the minimum wage argued that increasing the labor earnings via the minimum wage would reduce welfare spending and help households to end their welfare dependency.

This chapter is an ex-post evaluation of the minimum wage using the difference-in-differences technique. Estimation is based on county-level administrative data, exploiting regional variation in the *bite* of the minimum wage. The bite is the county-specific share of employees paid less than € 8.50 *before* the introduction of the new policy. The idea is to exploit the fact that a uniformly set minimum wage is felt very differently across the country. Hence, also the effects of such a policy should be larger in more heavily affected counties. The analysis considers the effects of the new minimum wage on regular and marginal employment and on welfare dependency, the so-called *Aufstocker*.

The results suggest that the minimum wage had a considerable negative effect on marginal employment. A back-off-the-envelope calculation indicates that in 2015 150,000-200,000 marginal jobs have been lost due to the minimum wage. Concerning regular employment, the results indicate a rather small (short-run) negative effect of the minimum wage. Concerning welfare dependency, the minimum wage reduced the number of working welfare recipients, with some indication that about one half of them left welfare receipt due to the minimum wage. The effect on welfare reduction in absolute terms is rather small. The analysis only considers the short-run effects in the first year after the introduction of the minimum wage. Thus, the results cannot give a proper indication for the long-run impact of the minimum wage, for instance during the next economic recession.

The fourth chapter, *Congestion Pricing: A Mechanism Design Approach*, deals with traffic congestion and the optimal design of congestion pricing. From an economist's point of view, congestion is an externality problem since an individual driver does not take the effect of her journey on other drivers' travel time into account. The textbook solution is a simple corrective Pigouvian tax, set at a level to ensure that each driver internalizes the marginal cost of the increased travel time of other drivers (Pigou, 1920; Knight, 1924). In order to find the correct level of such a tax or price, a regulator needs knowledge about the drivers' value of travel time. These values differ substantially across drivers, both based on observed and unobserved sources (Small et al., 2005; Steimetz and Brownstone, 2005). The fourth chapter uses mechanism design in order to study the optimal design of congestion pricing, when the regulator does not have this kind of information.

Using a simple model, in which a regulator assigns drivers between two congestible roads, this chapter shows that Pigouvian prices are not efficient, even if he knows the distribution, due to *aggregate uncertainty*, which arrises in case of a finite number of drivers. For instance, if a single driver has a very high value of travel time compared to the other users, it might be optimal that she drives alone on the faster road, while the Pigouvian price based on the expected distribution will also admit other drivers to the fast road. Thus, a single Pigouvian price does not always induce the optimal allocation.

Nevertheless, the regulator can implement the efficient allocation with a so-called Vickrey-Clarke-Groves payment rule: Drivers not only pay for road access but also for *faster travel*. The proposed mechanism ensures that all drivers truthfully report their values of time and can be assigned optimally between the two roads. In equilibrium, reporting your value of travel time is analogous to deciding out of a menu of desired travel time-price combinations. Thus, the suggested procedure opens the possibility of using second-degree price discrimination to extract drivers' value of time.

Even if the regulator does not have any prior knowledge about the distribution of the values of time, the proposed mechanism instantaneously determines the correct Pigouvian price. Thus, regulators can deal with aggregate uncertainty and can respond quickly to changes in demand. The solution suggested in this chapter might have seemed impractical in the past, since it involves direct communication between the drivers and the regulator. However, modern technology, such as smart phones, GPS, and the advent of self-driving cars imply that these practical problems may soon be overcome.

Zusammenfassung

Diese Dissertation besteht aus vier eigenständigen Arbeiten über unterschiedliche politische Maßnahmen. Die Themen reichen vom Einfluss des Einkommensrisikos auf Fertilitäts- und Arbeitsangebotsentscheidungen, über eine Evaluation des gesetzlichen Mindestlohns in Deutschland, bis hin zum optimalen Design von Mautgebühren. Zwar unterscheiden sich die vier Arbeiten sowohl methodisch als auch inhaltlich, allerdings haben sie gemeinsam, dass Unsicherheit eine entscheide Rolle zum Verständnis der zu Grunde liegenden politischen Fragestellung spielt.

Das erste Kapitel, *Is Our Income Too Risky to Have A(nother) Baby? Evidence from German Micro Data*, untersucht den Einfluss von Haushaltseinkommen, Verdienstmöglichkeiten der Frau, und den dazugehörigen Einkommensrisikomaßen auf Fertilitätsentscheidungen. Die Gründung einer Familie oder der Entschluss ein weiteres Kind zu bekommen haben einen permanenten Einfluss auf die Lebenssituation. Steigen die Einkommensrisiken eines Haushalts an, kann dies dazu führen, dass die Familiengründung verschoben wird, zum Beispiel um Ersparnisse aufzubauen, oder um sich weiter auf dem Arbeitsmarkt zu etablieren. Einkommensrisiko spielt in der Ökonomie eine wichtige Rolle zur Erklärung intertemporaler Entscheidungen. So können zum Beispiel Vorsichtssparen, aber auch Arbeitsangebotsentscheidungen (wie im zweiten Kapitel dieser Dissertation) damit in Verbindung gebracht werden. Empirisch zeigt sich, dass das Einkommensrisiko volatil ist und seit den 1970ern Jahren in Deutschland und anderen Ländern angestiegen ist (Bönke et al., 2015; Blundell et al., 2015). Daher möchte das erste Kapitel die Frage beantworten, ob es einen *kausalen* Zusammenhang zwischen Einkommensrisiko und veränderten Fertilitätsverhalten gibt.

Das erste Kapitel ist eine empirische Analyse auf Basis des sozio-oekonomischen Panels (SOEP) für die Jahre 1984 bis 2014. Fertilität wird als Zeit bis zum ersten (oder nächsten) Kind modelliert und mittels eines diskreten Verweildauermodells geschätzt, welches unbeobachtete Heterogenität berücksichtigt (Heckman and Walker, 1990a,b,c). Dieses Modell wird mit dem Stundenlohn der Frau, dem Haushaltseinkommen und den dazugehörigen Risikomaßen, sowie einer Vielzahl von sozio-demografischen Kontrollvariablen erweitert. Die Einkommens- und Risikomaße basieren nicht auf den beobachten Werten, sondern auf dem zu erwartenden sogenannten permanentem Einkommenspotenzial, um von den tatsächlich getroffenen Arbeitsangebotsentscheidungen zu abstrahieren. Der potenzielle Stundenlohn der Frau bestimmt die Substitutionsentscheidung zwischen Arbeitszeit und Zeit für Kinderbetreuung. Um den Effekt von Einkommensrisiko besser zu verstehen, wird die Analyse getrennt für zwei verschiedenen Kohorten (geboren zwischen 1960 und 1974, sowie zwischen 1975 to 1989) und nach Bildungsabschluss (Akademikerinnen) durchgeführt.

Die Resultate legen nahe, dass Einkommens- und Risikomaße bei der älteren Kohorte keinen signifikanten Einfluss auf das Timing der Fertilitätsentscheidungen hat. Bei der jüngeren Kohorte hingegen führt das Lohnrisiko der Frau zu signifikant späteren Übergängen zum ersten Kind und auch zu mehr Kinderlosigkeit. Für den Übergang zum zweiten Kind zeigt sich, dass bei der Älteren Kohorte ein hohes Lohnrisiko der Frau den Abstand zum zweiten Kind verkürzt. Dies könnte darauf zurück geführt werden, dass es für diese Frauen schwierig ist nach der Geburt des ersten Kindes zurück im Job einzusteigen, und deshalb das zweite Kind "vorgezogen"wird. Dieser Effekt ist bei der jüngeren Kohorte nicht mehr zu beobachten. Allerdings zeigt sich bei diesen Frauen, dass das Risiko des Haushaltseinkommens einen negativen Einfluss auf die Übergangswahrscheinlichkeit hat. Für den Übergang zum dritten Kind können keine signifikanten Effekte gefunden werden. Es zeigt sich, dass die Effekte vor allem bei Nicht-Akademikerinnen auftreten.

Auch das zweite Kapitel, *How Important is Precautionary Labor Supply?*, befasst sich mit dem Einfluss von Lohnrisiko, aber diesmal auf das Arbeitsangebot von Männern. Hierzu soll mittels einer empirischen, explorativen Analyse die Frage beantwortet werden, ob Personen mehr Stunden arbeiten um sich gegen das Einkommensrisiko zu versichern. In diesem Kapitel wird also die Rolle von *Vorsichtsarbeiten* quantifiziert, definiert als der Unterschied zwischen den tatsächlich gearbeiteten Stunden und den hypothetischen gearbeiteten Stunden, wenn es vollständiger Sicherheit über den Einkommensverlauf gäbe.

Die Schätzungen im zweiten Kapitel beruhen auf einer dynamischen Arbeitsangebotsgleichung, ebenfalls auf Basis der Daten des sozio-oekonomischen Panels. Das Lohnrisiko wird in diesem Kapitel als Standardabweichung der persönlichen Stundenlöhne in der Vergangenheit operationalisiert und basiert auf den sogenannten marginalen Nettostundenlöhnen, also dem effektiven Nettostundenlohn, der sich aus dem gewählten Arbeitsangebot ergibt. Diese marginalen Nettostundenlöhne werden mit dem Mikrosimulationsmodell STSM (siehe Steiner et al., 2012) berechnet. Die Stundenlöhne werden mit vergangenen Arbeitseinkommen instrumentiert, um Endogenitätsprobleme zu vermeiden. Zudem wird auch die personenspezifische Wahrscheinlichkeit von Arbeitslosigkeit (Analog zu Carroll et al., 2003) mit in der Schätzung berücksichtigt.

Die Ergebnisse zeigen, dass das Lohnrisiko einen signifikanten positiven Einfluss auf die geleisteten Stunden hat. Angestellte arbeiten 2,5% ihrer Stunden, also in etwa eine Arbeitswoche pro Jahr, um sich selbst gegen ihr Lohnrisiko zu versichern. Vorsichtsarbeiten ist bei Selbstständigen besonders wichtig und beträgt 5,53% der geleisteten Stunden. Bei Beamten kann kein signifikantes Vorsichtsarbeiten nachgewiesen werden. Wenn alle Angestellten das Lohnrisiko des Medianbeamte hätten, würden die geleisteten Stunden um etwa 1% zurückgehen. Vorsichtsarbeiten ist auch wirtschaftlich relevant. Bei einer durchschnittlichen Arbeitszeit von 42 Stunden und einem Grenzstundenlohn von 13 €, ergibt sich ein monetärer Nettogegenwert des Vorsichtsarbeitens von etwa 710 € pro Jahr.

Das dritte Kapitel, *The Effects of Germany's New Minimum Wage on Employment and Welfare Dependency*, analysiert die Effekte des allgemeinen gesetzlichen Mindestlohns von 8,50 €. Es soll beantwortet werden, ob der Mindestlohn zu Jobverlusten geführt oder dazu beigetragen hat die Anzahl der sogenannten *Aufstocker*, also arbeitende Arbeitslosengeld-II-Bezieher, zu reduzieren. Im Vorfeld der Einführung des Mindestlohns wurde häufig argumentiert, dass dieser notwendig sei, um die Einkommen im Niedriglohnsektor zu erhöhen und die sogenannten Hartz-Reformen abzufedern.

Die Evaluation des Mindestlohns wird auf Basis von regionalen aggregierten Daten der Arbeitsagenturbezirke durchgeführt. Der Anteil der vom einheitlichen Mindestlohn betroffenen Arbeitnehmer unterscheidet sich stark zwischen den Bezirken. Die Effekte des Mindestlohns sollen durch den zeitlichen Vergleich zwischen unterschiedlich stark betroffenen Regionen, der sogenannten Differenz-in-Differenzen Methode identifiziert werden. Betrachtet werden die Effekte auf sozialversicherungspflichtige und geringfügige Beschäftigung, sowie auf die *Aufstocker* und nicht-arbeitende Arbeitslosengeld-II-Bezieher.

Die Resultate legen nahe, dass der Mindestlohn zu einem Rückgang der geringfügigen Beschäftigung von etwa 150 000 - 200 000 Personen geführt hat. Bezüglich regulärer, sozialversicherungspflichtiger Beschäftigung zeigen die Resultate ebenfalls einen kleinen Rückgang auf, dieser ist aber nicht sonderlich robust gegenüber Änderungen der Spezifikation. Auch die Anzahl der *Aufstocker* ist durch den Mindestlohn geringfügig zurückgegangen, allerdings kann nicht ausgeschlossen werden, dass dies nicht auch auf Jobverluste zurückzuführen ist. Zudem entsprechen die Resultate nur den kurzfristigen Effekten im Jahr nach der Einführung. Auf Basis der Ergebnisse können keine Aussage über die längerfristigen Effekte, zum Beispiel in der nächsten Rezession, gemacht werden.

Im vierten und letzten Kapitel, *Congestion Pricing: A Mechanism Design Approach*, geht es um Verkehrsüberlastung und das optimale Design von Mautgebühren. Aus Sicht der Ökonomie sind Staus und Verkehrsinfarkte ein Problem von externen Effekten, da jeder Fahrer seinen negativen Auswirkungen auf die anderen Fahrer ignoriert. Die Lehrbuchlösung für solche Probleme ist eine sogenannte Pigou-Steuer, die so gesetzt wird, dass jeder Fahrer seine eigenen negativen Auswirkungen mit berücksichtigt (Pigou, 1920; Knight, 1924). Um aber die korrekte Höhe der Pigou-Steuer zu ermitteln, braucht der Regulierer Informationen über die Verteilung der Opportunitätskosten der Fahrer. Diese können nicht direkt beobachtet werden und unterliegen zudem großen Schwankungen (Small et al., 2005; Steimetz and Brownstone, 2005). Im vierten Kapitel wird mittels der sogenannten *Mechanism Design* Theorie gezeigt, wie der Regulierer die korrekte Pigou-Steuer bestimmen kann, ohne direkt über die Verteilung der Opportunitätskosten informiert zu sein.

Der Regulierer verteilt in einem einfachen Modell Fahrer auf zwei möglicherweise von Verkehrsüberbelastung betroffene Straßen. Selbst wenn der Regulierer die Verteilung der Opportunitätskosten kennt, führt eine Pigou-Steuer nicht immer zur optimalen Verteilung, wenn es eine begrenzte Anzahl von Fahrern und damit *Verteilungsunsicherheit* (*aggregate uncertainty*) gibt. Beispielsweise kann es zu einer Situation kommen, in der ein Fahrer im Vergleich zu den anderen sehr hohe Opportunitätskosten hat, und es optimal wäre, dass dieser Fahrer eine Straße alleine benutzt. Die Pigou-Steuer auf Basis der Verteilung führt aber dazu, dass noch andere Fahrer auf der schnellen Straße fahren werden.

Allerdings kann der Regulierer mit einem sogenannten Vickrey-Clarke-Groves-Mechanismus dafür sorgen, dass immer die effiziente Verteilung erreicht wird. Fahrer zahlen hierbei nicht nur für den Zugang zur schnellen Straße, sondern auch für eine kürzere Reisezeit. Der Mechanismus stellt sicher, dass alle Fahrer dem Regulierer wahrheitsgemäß ihre Opportunitätskosten mitteilen und dieser dann die optimale Allokation vornehmen kann. Im Gleichgewicht ist die Mitteilung der Opportunitätskosten äquivalent zu einer Auswahl aus einer Reisezeit-Preis Kombination.

Selbst wenn der Regulierer keine Informationen über die Verteilung der Opportunitätskosten hat, erreicht der vorgeschlagene Mechanismus die optimale Allokation. Der Regulierer kann deshalb auch schnell auf Veränderungen in der Nachfrage und der Opportunitätskosten reagieren. Da permanent zwischen dem Regulierer und den Fahrern kommuniziert werden muss, konnte ein solcher Mechanismus in der Vergangenheit nicht umgesetzt werden. Mit dem Aufkommen moderner Technologien, wie Smartphones und GPS, sowie zukünftig mit selbstfahrenden Autos, könnte ein solcher Mechanismus aber eingeführt werden und die Verkehrssteuerung verbessern.