Essays on prices and price convergence

Dissertation

zur Erlangung des akademischen Grades
Doctor Rerum Politicarum

an der
Fakultät für Wirtschafts- und Sozialwissenschaften
der Ruprecht-Karls-Universität Heidelberg

vorgelegt von

Andreas Lindenblatt

Heidelberg, März 2015

Contents

List of Figures

List of Tables

Acknowledgments

During the five years that I have been working on this thesis, many people have supported me. I am very grateful to all of them!

First, I am especially grateful to my supervisor Switgard Feuerstein for the plentiful advise and support she gave me during my studies as well as for co-authoring an article with me and letting me pursue my own research ideas. I am also very grateful to Christian Conrad for supporting me and readily being a member of the thesis committee. Thank you to Peter Egger for being my co-author and supporting me with the project on sex work in Germany. Katharina Zweig taught me how to program in Java and Stephan Kassemeyer how to program in Python. Without him, there would not be a sex work data set. Additionally, I had excellent help by student researchers: Jan Mellert, Christoph Altmeppen, Lisa Seitz and Luisa Kling worked as student researchers at the working group for International trade and supported me with all organizational and research matters. Furthermore, Christoph Altmeppen, Johanna Bocklet, Michael Reiß, Roja Saidi, Yannjik Saßmann and Paula von Haaren helped me in creating the sex work data set.

I am also grateful to my (former) co-workers at the Alfred-Weber-Institute, Heidelberg University. During the last years I have found a very motivating and inspiring atmosphere. Interesting and intellectually challenging discussions have been held, not only in the seminars, but also during the lunch and coffee times and in the noodle group. Here, I am especially grateful to my (former) colleagues Isadora Kirchmaier, Karin Loch, Sarah Necker, Dennis Pohl, Benjamin Roth and Andrea Vosskort, who by now have become good friends.

Thank you also to all my friends and family for their constant support. My parents have supported me during my whole studies and thesis and were always there for me. My brothers have come by, occasionally helped me solve programming problems and looked after Felix. My friends Julian and Stephan motivated me when needed and I had a lot of fun with them during this time.

Finally, and most important, I want to thank my wife Suse. She has supported me whenever it was necessary and is always there for me when I need her.

Chapter 1

Introduction

3.5 to 1.5 million years ago, predecessors of mankind started to walk upright, eat meat and trade. In this era, food was traded in hope of getting back goods and commodities in the future or to win other mens' and womens' favor. Then, in the Upper Paleolithic (45,000-10,000 years B.C.), labor division started and goods were traded against goods (Alfred, 2011, p. 243ff.). Already then, the concept of a price emerged in the form of terms of trade in a barter, and later on with the introduction of money.

Today, a more or less consistent view of what role prices play has evolved. Friedman (1988) summarizes as follows: First, prices have an information function. Prices transmit information on preferences, production possibilities and availability of resources, thereby coordinating market participants and clearing the market. Second, they have an allocation function. Prices provide an incentive for individuals to search for the cheapest way of producing goods and using resources where their utility is highest. Third, they have a distribution function. When a good is produced and consumed, a rent is created. Prices define the share of this rent each individual involved in this process gets. However, in the history of economic thought, it has been a long road until the importance of prices for markets and economies has become that clear.

1.1 A short history of prices

The earliest surviving comments on prices originate from Platon and Aristotle (5th and 4th century B.C., respectively). While Platon only stated that one should not price discriminate between customers, Aristotle distinguished between the value in use and the value in exchange of a product. The latter could be higher than the value in use and is supposed to be determined by "proportionality" of the traders. Hence, a higher status trader is allowed to charge a higher price (Schefold, 1994, p. 113ff.). In the 13th century, theologian Thomas Aquinas developed the concept of a "just price", which compensates the seller for his individual cost, but not more.

This price is god-made and hence higher prices are unjust. In the 14th and 15th century, this concept was modified. Johannes Buridanus was the first to argue that not individual, but collective supply and demand determine the price of a good. Furthermore, Antonin of Florence argued that prices not only reflect utility, but also rarity and subjective value of a good. Hence, building on Aquinas, they argued that there could exist a range of prices from just to unjust ones (Schinzinger, 2002).

In the 16th and 17th century, not much new was added to price theory until in the 18th century classical economists such as Adam Smith and David Ricardo successively introduced their labor theory of value (prices). They argued that the "natural" (long-run) price of a good can be derived by adding up the amounts of labor directly and indirectly (through the accumulation of capital) used in its production. Prices can vary around this "natural price" due to market power of the sellers (e.g., monopolies) or due to a mismatch of supply and demand (Schefold & Carstensen, 2002).[1]

In the 19th century, with the *"Marginal Revolution"*, authors such as William Stanley Jevons, Carl Menger, Léon Walras and Alfred Marshall shifted the focus towards marginal utility (Niehans, 1994). In their theory, prices do not reflect the imputed labor, but the marginal utility of the last unit traded. Producers will stop offering more units of a good, as soon as marginal cost of production are higher than the marginal consumer is willing to pay. Consequently, Jevons (1871) argued that in a market with perfect competition and no transaction cost, taxes, etc., there can only exist one price.[2] Using this concept of marginal utility, since then, contributions of price theory and industrial economics have been concerned with explaining price formation under more general assumptions and looked at situations of market power, market failure, incomplete information or state interventions (see, e.g., Weber, 2012 for a detailed overview).

1.2 Outline of the thesis

This thesis looks at prices in two different markets. The first one is the market for food products in Europe. With the introduction of the common market in 1992, most European markets have been integrated. When 10 more countries joined the EU in 2004, another round of integration took place and the common market was extended to these countries as well. Chapter 2 analyses if retail prices for food products have converged in the time after this "shock" of the EU enlargement.[3] While there exists

[1] Adam Smith only acknowledges that the demand depends on the price, but not the other way round.

[2] This idea has further been developed to the purchasing power parity by Cassel (1918).

[3] Chapter 2 is based on a joint project with Switgard Feuerstein (Lindenblatt & Feuerstein, 2015). Both authors contributed to these papers in equal shares to the design, execution, data analysis and writing. In this paper I have an above average stake in the programming work to create and to clean the dataset as well as to implement the statistical analysis. The Copyright of this

an extensive literature on convergence in general, this chapter is the first to be able to look at retail price convergence within the European Union at a micro-data level. By decomposing price convergence into within sub-groups and between sub-groups of countries convergence, we add further insight to the literature on what causes the strong price convergence within the enlarged EU. In future research it might be interesting to look at a longer time horizon and to include products of other sectors into the analysis.

The second market this thesis looks at is the one for internet facilitated sexual services in Germany. Sex work and the advertisement thereof is legal in Germany, which has led to a range of internet platforms concerned with selling sexual services. While many platforms only contain advertisements, one of these platforms – *www.gesext.de* – gives sex workers the opportunity to sell their services either as an auction or at a fixed price. This has allowed us to create a dataset on sex work with information based on actual concluded contracts, which is a very unique feature in this kind of literature. Furthermore, each data point is geo- and time-referenced. This dataset is used to show that 1) offering unprotected sexual services is endogenous, 2) local events influence the supply, demand and price of sexual services, and 3) regional effects influence local prices and habits.[4]

In writing these three chapters we have focused on basic research questions that had still not been sufficiently answered in the given literature, but the richness of this dataset allows for several future research projects. First, one might look at how the decision to bid on a sex worker depends on customers' preferences. Second, sex workers are repeatedly observed and it will be interesting to look at the dynamics of how sex workers react to competition and information gathered from their (or others') earlier contracts. Third, there are hardly any datasets on goods or services whatsoever that contain repeated observations both on the seller and on the buyer side of both sold and unsold offers (supply) as well as winning and loosing bids (demand). Most datasets only include market prices of sold goods. With this kind of information demand curves can be created and price elasticities calculated. Furthermore, the concept of a hedonic price regression, which in its normal form only includes market prices (where supply and demand are matched), can be improved by including unmatched demand and supply.

The rest of this thesis is organized as follows. The next paragraphs give a short overview of the five main chapters, which are followed by the chapters themselves. These five chapters can be read independently from each other, but share a common

article lies with the Oxford Journals, Oxford University Press, Oxford, *DOI: 10.1093/erae/jbu038*

[4]The three articles that the chapters 4–6 are based on are co-authored by Peter Egger (Egger & Lindenblatt, 2015a,b,c). Both authors contributed to these papers in equal shares to the design, execution, data analysis and writing. In these papers I have an above average stake in the programming work to create and to clean the dataset as well as to implement the statistical analysis. The Copyright of Egger & Lindenblatt (2015a) lies with the Springer publishing group Berlin-Heidelberg, *DOI: 10.1007/s10198-014-0640-2*

bibliography and appendix.

Chapter 2: Price convergence after the Eastern enlargement of the EU: Evidence from retail food prices

In chapter 2 findings on food price convergence in Europe after the enlargement of 2004 are presented. Since the single market was introduced in 1992, there has always been the expectation for prices to converge. However, two decades and many studies later, it has become clear that prices continue to differ. On the one hand, higher competition, converging labor cost and the lack of trade barriers should close the gap between prices of identical goods in European countries. On the other hand, if retailers provide value adding services to customers or if their preferences differ from country to country, price differences may persist in competitive markets.

From a statistical point of view, these two opposing forces translate into mean reversion decreasing all price differences by a certain percentage and random shocks on individual price differences increasing them again. In the earlier convergence literature, price convergence was measured by looking at the size of mean reversion (β-convergence). However, as random shocks can increase prices again, mean reversion is only a necessary, but not a sufficient condition for price convergence. In contrast, the concept of σ-convergence looks at the change of the variance and hence also includes random shocks increasing prices again. In the long-run, the two countervailing effects balance out and price dispersion reaches a steady state. Only if this state has not yet been reached, or if an external shock changes it, price convergence will be observed. The enlargement round of the EU in 2004 might constitute such a shock. At that time, eight Central and Eastern European member states (as well as Malta and Cyprus) have joined the European Union, markets have been liberalized and trade barriers have fallen.

To analyse if prices converged after the enlargement in 2004, a unique dataset on European retail food prices is used - both at the level of individual products as well as at a semi-aggregate level for EU15 and all eight Central and Eastern European accession countries (EU8). In line with the new convergence literature, we test for σ-convergence and find that from 2001-2003, the period before the accession, prices diverged within EU23. In contrast, there was strong price convergence from 2003-2006 and 2006-2009.[5] When price convergence is decomposed into within and between sub-groups convergence, we find that it is dominated by convergence between EU8's and EU15's prices and not prices within each of these two subgroups.

[5]Micro-level data is not available for every year, but only for the years mentioned here.

Chapter 3: Sex work: An introduction and a new dataset

Chapter 3 gives some background on sex work in Germany and introduces the full *www.gesext.de* dataset as well as its descriptive statistics. The chapter first gives a historical overview on sex work in ancient Europe and especially modern times Germany. Until recently, sex work in Germany was in a legal gray area and only with the introduction of a new legal framework in 2002, the legal status has been completely clarified. In comparison to other European countries, the German laws are very progressive. Sex workers' rights have been put on par with workers' rights in other sectors, while still keeping forced prostitution and human trafficking illegal (Kavemann & Steffan, 2013).

Against this background of sex work being legal in some countries and illegal in other ones, existing theories on sex work are discussed. These theories basically explain three features of the market. First, the existence of a price-premium between sex workers and non-sex workers, is explained by the opportunity cost for encountering higher risk of violence and social stigma. Second, price premia between sex workers for certain types of sexual services can again be explained as compensation for risk-taking or higher opportunity cost (e.g., offering unprotected sex and hence encountering a higher risk of catching a sexually transmitted disease). Third, the decision of a sex worker to enter the market is discussed.

Furthermore, existing datasets for the market for sexual services are presented. These are mainly survey based and for developing countries. None of them contains micro-level data for Germany. In this way, our German dataset, which is extracted from *www.gesext.de*, is unique: it consists of contract data for a developed country.

As in chapters 4 – 6 sub-samples of this dataset are used, we explain in detail how this dataset was generated and present the descriptive statistics of its variables. On this platform, sellers can advertise sexual services and either offer them as an auction or at a fixed price. Using a self-written script, we downloaded all sexual service offerings posted on *"www.gesext.de"* from January, 16 in 2011 to September, 9 in 2012. In total female suppliers posted 102,096 offerings, of which 33% were sold. All sellers together made 6.2 million Euro or on average 103 Euro per hour during this time period. During the following year, the information contained in this dataset was extracted partly by pattern recognition software and partly by students categorizing the information. Additionally to the information on offered services as well as seller- and buyer-identifiers, the key features of the dataset are regional- and time-identifiers. These allow us to merge the dataset with information on the socio-demographics of the population living in a given region and events that took place in the same region at a given time. After describing the variables and their descriptive statistics in detail, the chapter compares some selected statistics over regions.

Chapter 4: Endogenous risk taking and physical appearance of sex workers

A key issue in sex work is health protection. Chapter 4 looks at risk (price) premia due to offering unprotected sexual services. In our dataset, a considerable share of sexual services are provided without a condom (6%), which is consistent with previous work (Rao et al., 2003; Gertler et al., 2005; Robinson et al., 2011; Arunachalam & Shah, 2013). Thereby both the sex worker and its client are endangered of catching a sexually transmitted disease (STD). As Cunningham & Kendall (2011b) argue, this share is likely to be even higher in the offline part of the market, as sex workers there tend to be less educated and engage in more risk-taking.

In studies that do not control for endogeneity of risk-taking, the estimates of price premia thereof range between 9% in Kenya (Robinson et al., 2011) and 30% in Mexico (de la Torre et al., 2010). The only study that does control for endogeneity estimates the price premium to be in India in 1993 as high as 194% − 376% (Rao et al., 2003). Several studies show that the likelihood that sex workers take this risk is endogenous and depends on many factors such as the price premium or the sex worker's attractiveness (Gertler et al., 2005; Chang & Weng, 2012). However, these studies do not control for the resulting endogeneity when estimating price premia.

Based on an instrumental variable approach, our results suggest that the incentive (price premium) for risk-taking is about twice as high when accounting for endogeneity than when assuming random assignment of risk-taking. When a service is offered without protection, the seller earns a premium of 91%. The estimate is lower than the estimate in Rao et al. (2003), but this might be due to the fact that Rao et al. look at a developing country and an earlier time span (1993), with less access to HIV and other medication than in Germany today. Within regions, we find some spill overs, resulting in a higher likelihood to offer unprotected services if other sellers offer them as well. In contrast to previous work, we do not find any evidence for overweight sex workers receiving a lower price premium than non-overweight sex workers. These results are relevant from a policy perspective as they clearly show that 1) there is an incentive to offer unprotected sexual services and this incentive is even higher than when not controlling for endogeneity, and 2) this incentive depends on practices of other local sex workers.

Chapter 5: Fun and games: How soccer games and conventions raise the demand for paid sex

In chapter 5 we look at one specific aspect of large events: how shocks on the demand side influences the market for sexual services. The literature on the economic effects of events in general comes to mixed conclusions whether hosting these events pays off in economic terms. While large sport events tend to have no or a negative impact on economic well-being (e.g., Allmers & Maennig, 2009; Baade & Matheson, 2004; Hagn & Maennig, 2009; Feddersen & Maennig, 2012), there is some evidence that trade fairs increase economic growth (e.g., Tourism Vancouver, 2007; Penzkofer, 2008; Bathelt & Spigel, 2012). What has however become clear in the literature is that the price as well as the supply of sexual services react to market forces such as the size of supply and demand (Moffatt & Peters, 2004; Della Giusta et al., 2008, 2009a; Cunningham, 2011).

Our newly created dataset (see above) includes geo- and time-references for every individual transaction, which can be used to match the services offered and purchased with event information. Namely, we merge our dataset with information on national league soccer games, European Championship games and trade shows. This allows us to analyze how these three types of events influence supply, demand, and prices in the region at the time an event takes place, as well as in its neighboring regions.

Employing Poisson quasi-maximum-likelihood estimators, we provide evidence that events are relatively important. Ongoing soccer game (German league or European Championship) and trade fair events raise the number of sellers and buyers (extensive margin) as well as contracted services in the respective region by a high two-digit percentage. The same is true for neighboring regions. Furthermore, there is also an effect on the intensive margin: on event days, sellers offer more services per day and buyers bid more often. If all events are combined and both direct (within region) and indirect effects (on neighboring regions) are included, this rise is in the high two-digit percentage range. Furthermore, there are also effects on prices, which are smaller and, for soccer events, depend on whether events happen locally (home games versus away games), and how the local team fares (looses or wins). Overall, we estimate moderate price effects of, e.g., $3\% - 5\%$ for auctions.

Chapter 6: The long shadow of the iron curtain for female sex workers in German cities: Border effects and regional differences

The last chapter of this thesis provides evidence on regional differences within Germany in the market for sexual services. In the social sciences, it has become clear that regional economic and social conditions influence sex workers' decisions. First, local authorities and citizens determine moral and ethical norms and define what in their region is accepted and what it forbidden (Hubbard, 2001; Hubbard & Sanders, 2003; Hubbard & Whowell, 2008). Second, regions differ in their social organization, i.e., in how these norms can be enforced (Kubrin & Weitzer, 2003). Third, in the sexual fields theory it is argued that individual desires aggregate to a field of geographical scope with gravitational forces (Green, 2008, 2011). All three arguments lead to individuals' desires being influenced by the attitudes of the region they live in and to individuals moving to regions that fit their desires.

In Germany, regional differences in the market for sexual services originate from several sources. Before the reunification in 1990, offering sexual services was forbidden in Eastern Germany and in a legal gray area in Western Germany. It was existent in both parts of Germany (the State security service, Stasi, even used sex workers to acquire secrets from the West), but while in the East sex workers saw it as a way to earn Deutschmarks and get access to a life in luxury, in the West it was seen as a last resort to make money. With the reunification, the West-German code of law was also introduced in the former Eastern states. Until 2002, offering sexual services remained in a legal gray area (see section 3.1) and since then offering sexual services is completely legal. However, regional differences still exist as sex work is still policed and taxed on a local level. Furthermore, moral attitudes only change gradually and differences between regions and especially between the former East and the former West still exist (e.g., see Alesina & Fuchs-Schündeln, 2007).

The empirical part of chapter 6 alludes to a regional pattern, in particular, in pricing. This pattern varies with the size of cities and across as well as along the former East-West German border. In particular, the evidence from a regression discontinuity design (RDD) suggests that there is a long shadow of the former Iron Curtain which leads to higher conditional prices in the former East than in the West, and especially, in larger agglomerations such as Berlin. Moreover, there is evidence of habit formation and spillovers within regions which lead to regionally clustered prices as well as unprotected services offerings by sexual workers. This is in favor of the theories mentioned above, arguing regions have an effect on individuals' behavior due to the existence of certain norms and their enforcement through, e.g., social fields.

Chapter 2

Price convergence after the Eastern enlargement of the EU: evidence from retail food prices

2.1 Introduction

This chapter studies the effect of the EU enlargement in 2004 on food price convergence. Using comprehensive Eurostat data[1], it is the first study looking at price convergence of retail prices at a disaggregate level, both for the old member states (EU15) and the eight new Central and Eastern European member states (EU8). The data allows us to look at price convergence for a time period that includes more than just the initial effects right after the EU enlargement.

Within the single market of the European Union, large price differences for identical products are well documented, which can only partly be explained by cost differences or by differing overall price levels. In general, it was expected that these price differences decrease over time, as ongoing market integration amounts to convergence of prices. There are, however, also reasons why price differences may persist and need not disappear. Especially in the retail sector, price differences for individual products are feasible even in competitive markets. If retailing gives a value-adding service to consumers, arbitrage will not take place. After some adjustment, a constant level of price dispersion may be reached. Price convergence will only be observed if either this steady state has not been reached yet, or if an external shock has lead to a new steady state with lower price dispersion.

The Eastern enlargement of the EU in 2004 is likely to be such an external shock, but it remains an empirical question, if this shock has actually induced price convergence. The results are of political relevance, as (Dreger et al., 2008, p.66) point out that, "price convergence facilitates the working of common economic policies". In this context, the food market is particularly interesting. While in most sectors,

[1]We thank Paul Konijn of Eurostat for providing us with the data.

trade liberalization took place gradually over the whole decade before the accession, the agricultural sector was only liberalized in 2004.

Until the 1990s, price convergence was thought of as *β-convergence* (mean reversion), which is found in virtually all studies on European prices. However, in the recent literature it has been widely discussed that β-convergence is not sufficient for price convergence in the sense of a decreasing price dispersion between countries. It is possible to observe significant mean reversion, while the variance of prices increases due to ongoing random shocks on prices. Hence, it is necessary to examine *σ-convergence*, which is defined as a decrease of the variance over time (Barro & Sala-i-Martin, 1995; Carree & Klomp, 1997). This point is further discussed in section 2.4. Applying a test for σ-convergence by Egger & Pfaffermayr (2009), we find *price divergence* within the European Union as a whole in the lead-up to accession, but *price convergence* thereafter. Furthermore, we decompose the variance of prices and conclude that price convergence within the EU as a whole is partly driven by the dynamics within EU8, but the main effect is convergence between EU8's and EU15's prices.

The chapter is organized as follows. Section 2.2 gives a short overview of the literature. In sections 2.3 and 2.4, we present the data and the method used. Section 2.5 looks at σ-convergence as well as its decomposition and section 2.6 concludes.

2.2 Retailing, the EU and price convergence

According to the law of one price (LOOP), on a perfectly integrated market a good will only have one price. When the European Single Market was introduced in 1992, it was therefore widely expected that prices would converge within the European Union and that price differences between member countries would merely reflect remaining cost differences, such as transport costs or differing value added taxes (Cecchini et al., 1988). Similar expectations emerged when the European Monetary Union (EMU) started (European Commission, 1990).

There are, however, also reasons why price differences are observed and may persist over time. Retail prices do not only include production costs or import prices, but also retailing costs, which to a large extent reflect labour costs, and profit margins. Price differences for identical products are feasible as long as they do not exceed arbitrage costs (which may well be higher than mere transportation costs). Taking into account that retailing gives a value-adding service to consumers, arbitrage will hardly take place at the level of retailing and price differences may persist over time even when all trade barriers have been removed.

Price levels are in general lower in countries with a low per capita income. The accession of the Central and Eastern European transition countries, which have low price levels, has increased price dispersion within the EU. Dreger et al. (2008)

emphasize two countervailing effects on the price level indices of these countries. The integration into the internal market will increase competition and thereby lower prices, while the process of catching up increases prices due to the Balassa-Samuelson effect, i.e., due to rising wages. When the second effect dominates, rising labour costs lead to rising prices. Price differences within the EU become smaller and thus price convergence is observed.

However, price patterns are much more complex than just some countries having high prices and other countries having low prices. Relative prices of pairs of goods also vary a lot among the member states of the EU, with individual products being relatively expensive in some countries and relatively cheap in others (see section 2.3). Such price differences may reflect different mark-ups for the same product in different countries, thus they might emerge from price setting behaviour in the retail sector. Bliss (1988) develops a theory of retail pricing and suggests that mark-ups to cover overhead-costs will be set according to Ramsey-taxation rules, leading to higher mark-ups for goods with a lower elasticity of demand. Consequently, retail prices for individual products will vary within a group of countries as long as demand patterns – and therefore elasticities of demand – differ, even in competitive markets. Such differences in product-specific profit margins can explain long-run price differences for individual products. MacDonald & Ricci (2005) discuss the effect of the distribution sector in the context of the Purchasing Power Parity (PPP), albeit at a highly aggregate level. There are thus forces reducing price differences in integrated markets, but also reasons for long-run price differences.

This idea translates into a statistical process of convergence with two countervailing effects: First, mean reverting forces generally lower all price differences between two points in time. Second, random shocks on individual prices increase price dispersion.[2] These countervailing effects eventually balance, and a constant level of price dispersion will be reached. Hence, price convergence is observed if the initial variance of prices is larger than in this long-run steady state. This might be the case if a given long-run steady state has not yet been reached or if an external shock has changed the long-run steady state. Whether prices converge in a given situation is thus an empirical question (see Barro & Sala-i-Martin, 1991; Carree & Klomp, 1997; Egger & Pfaffermayr, 2009).

Empirical studies show that the law of one price does not hold within the European Union. Large price differences for identical products are observed, which are not just due to cost differences (see, e.g., Cumby, 1996 for Big Mac Hamburgers, Haskel & Wolf, 2001 for prices of IKEA, a furniture retailer, and Goldberg & Verboven, 2004 for cars). In view of these price differences, several studies looked at σ-convergence of retail prices within the European Union. Egger et al. (2009)

[2] If one only looks at two points in time (as in the test by Egger & Pfaffermayr, 2009, which we use in this chapter), these random shocks also include reasons for permanent price differences of individual products (emerging, e.g., from varying demand structures as in Bliss, 1988; see above).

find some significant σ-convergence for tradeable but not for non-tradeable goods for EU12 in the time period when the internal market was introduced (1990 - 1996). Wolszczak-Derlacz & De Blander (2009) reject price convergence within EU15 and three selected new member-states in the following ten years from 1995 - 2005, and Parsley & Wei (2008) do not find price convergence for Big Macs during the time of the Euro changeover within the countries of the European Monetary Union (EMU). For the same time period (1995 - 2005), Fischer (2012) finds large variation of the prices of washing machines and even price divergence within EU15. Moreover, EMU seems to have had no effect on price convergence. Only two studies have data available for all new member states of the European Union, both using data until 2005. While Funke & Koske (2008) only look at mean reversion and not at σ-convergence, Dreger et al. (2008) find small (initial) effects of the EU enlargement on convergence. On the whole, the literature provides hardly any evidence of price convergence on a disaggregate level within the EU. However, the Eastern European enlargement in 2004 constitutes an especially large shock, shifting the long run equilibrium of price dispersion and hence creating potential for price convergence. This particularly applies for food prices, as in contrast to many other markets liberalization of this market indeed took place mainly at the date of accession.

Our chapter expands the existing literature in several dimensions. By covering the EU15 and *all* Central and Eastern European countries, which acceded the EU in 2004, as well as extending the period considered until 2009, it accounts for the full magnitude of the enlargement shock. It includes not only initial, but also medium-term effects on prices. Moreover, the analysis uses disaggregate data. We can thus study comprehensively, whether or not after the external shock of Eastern enlargement, retail food price convergence has taken place within the new EU. Moreover, we decompose convergence into within and between subgroups convergence.

2.3 The data

In the framework of the Eurostat-OECD Purchasing Power Parity Program, retail prices for more than 400 narrowly defined food products are collected every three years (e.g., "fresh milk, unskimmed, 1 liter, well known brand"). For every product, employees of the respective national statistical bureau collect prices at 15-20 different outlets ignoring short term price reductions such as promotional offers.[3] Each national statistical bureaus then adjusts these prices for spatial and seasonal effects (e.g., for fruit, fish, etc.) and weights them such that for the given product the average price reflects the national shopping pattern (with regard to the different types of outlets). As prices are reported in their national currencies, they are converted into euro with the official Eurostat exchange rate for the given year. The

[3] For specified brands, the price is collected in 5-10 outlets.

data is available for the three waves 2003, 2006 and 2009 and shows that there are large price differences for identical products within the EU, and also within EU15 (Eurostat, 2010). Unfortunately, not all products can be traced over time, as the set of products has somewhat changed. Moreover, countries only collect prices that they consider relevant, meaning that not all countries report all prices, which results in missing data. Matching the products of these three waves, we have compiled a data set of 147 products. For a more detailed description of the data see Eurostat (2007).

Eurostat also provides the data at a semi-aggregate level of 34 product group price level indices (also called "basic headings", e.g., fresh milk, pasta), and there are good reasons to analyse data at this level as well. In contrast to the individual products, there are observations for all countries for every product group and problems of missing data do not occur. Moreover, the price level indices are available for a longer time period (1995-2009 for EU15 and 2001-2009 for EU25), allowing us to consider price dispersion also in the period before accession. Hence, in section 2.5 the analysis of price convergence is conducted both at the level of individual products and at the level of the product groups.[4]

In the following analysis, we distinguish between three groups of countries. (i) the old member countries EU15, (ii) the Central and Eastern European countries EU8, which accessed the EU in 2004, and (iii) these two groups together as EU23.[5] Since Bulgaria and Romania only joined in 2007 and data is not available before 2006, these two countries are left out. Moreover, we omit Malta and Cyprus. Their situation as small island states is hardly comparable to the EU8.

Starting at the highest level of aggregation – the comparative price level index food and non-alcoholic beverages (CPLI food) – large price level index differences within EU23 are observed (Figure 2.1). With Poland as the cheapest country within EU23 (64% of the EU27 average) and Denmark as the most expensive one (145%), the difference in price level indices is more than 100%. It may be no surprise that price level indices vary substantially within the European Union of 23, as the new members just recently joined the EU. But even when we only consider EU15 countries, the price level indices differ from 96% of the EU27 average in Portugal to 145% in Denmark.

Although Denmark on the whole has the highest prices for food, there are also some food products that cost less than the EU27 average in Denmark and the Danish price level indices of the 34 product groups range from 81% for rice to 216% for mineral waters (EU27=100%). In Poland, the cheapest country, rice costs as much as in Denmark, i.e., 81% of the EU27 average. In contrast, rice costs 125

[4]Other papers on price convergence, which are based on data from the Eurostat OECD PPP program, look at the product group level (e.g., Egger et al., 2009) or an even more aggregate level (e.g., Dreger et al., 2008). They do not use individual product prices.

[5]One other group that comes directly to mind are the EMU countries. We do not report results for this subgroup as these do not differ from EU15.

Figure 2.1: *Comparative price level index food and non alcoholic beverages 2009.* The graph presents the CPLI food with the EU27 average being normalized to 100%. *Source: Eurostat data, own representation.*

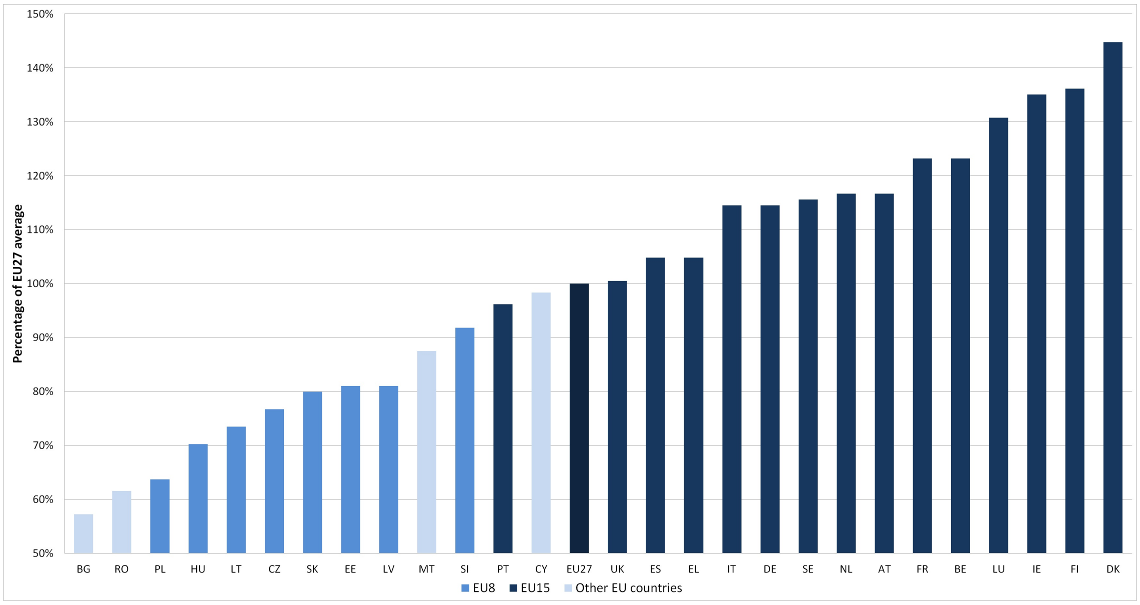
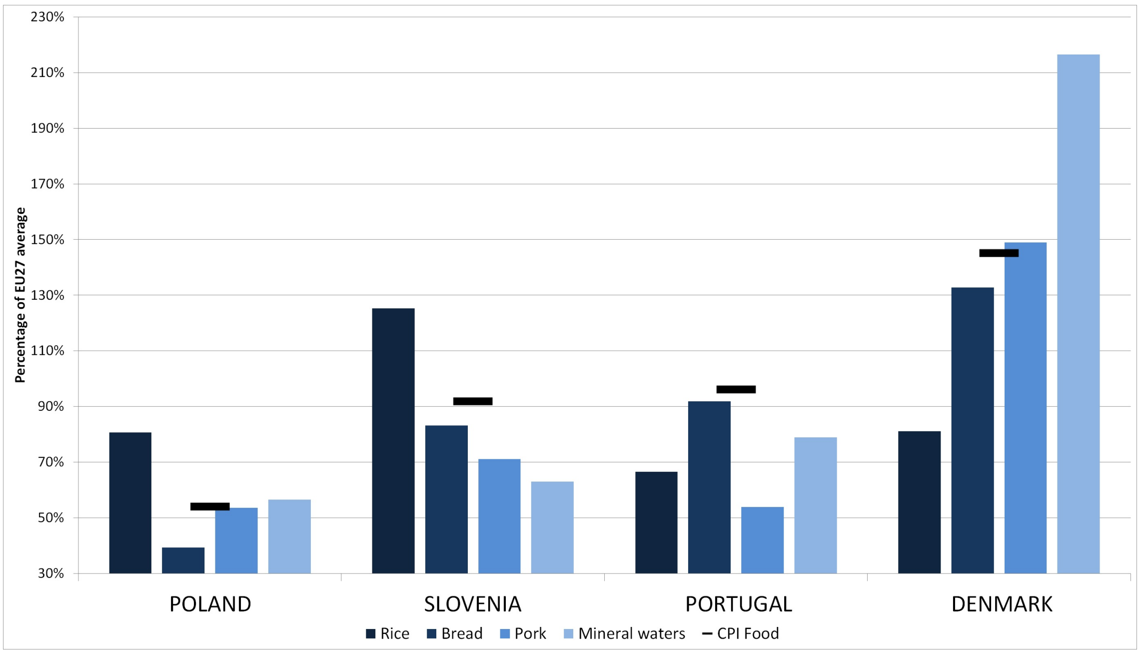


Figure 2.2: *Price level indices of selected product groups and countries, 2009.* The graph presents the PLI of some selected product groups and countries. The respective EU27 average has been normalized to 100%. The black bar depicts the CPLI food of the respective country. *Source: Eurostat data, own representation.*



% of the EU27 average in Slovenia, where the CPLI food is below 100%. Figure 2.2 illustrates this variety showing the price level indices of some selected product groups and countries as well as their CPLI food.

Similar patterns arise when individual products within a product group are considered. Even when taking country-product group-specific effects into account, considerable unexplained price differences remain at the individual product level. For many individual products, prices (measured relatively to the European average of this product) are well above or well below the price level index of the respective product group in that country. Within the product group "rice" individual prices vary for example from 74% - 84% of the EU27 average in Poland, 99% - 147% in Slovenia and 71 - 121 % in Denmark. Thus retail price differences are much more complex than just countries being expensive or cheap on the whole, and price convergence must be considered at a disaggregate level. Moreover, these price patterns imply that price convergence may depend on the level of aggregation. It is possible that prices converge at the level of individual products but diverge at the aggregate level of product groups (and vice versa). As mentioned above, the Danish prices for the different kinds of rice vary from 71% to 121% of the EU27 average, with a product group price level index for rice of 81%. If the 121%-price falls, this contributes to convergence at the product level but to divergence for the product group rice.

2.4 Methodology

2.4.1 Measuring convergence

In the literature, two different approaches to measure price convergence over time are used. The concept of *β-convergence* refers to mean reversion. It measures the forces that make deviations from a uniform price – caused by a one-time initial shock – fade out over time. In contrast, *σ-convergence* measures if the variance of prices within a group of countries becomes smaller over time (Barro & Sala-i-Martin, 1995). The latter concept takes into account, that there are two opposing effects: mean reversion generally reduces all price deviations, but ongoing random shocks on individual prices increase price dispersion again. Eventually a long-run steady state is reached, where these two forces balance out and the variance of prices remains constant over time. Price convergence (divergence) will be observed if the initial price dispersion exceeds (is below) the long run steady state value. In this case, mean reversion is strong enough (not strong enough) to dominate the dispersion induced by these ongoing additional random shocks. Thus, mean reversion is only a necessary, but not a sufficient condition for price convergence.

To analyse σ-convergence, we apply a Wald-test proposed by Egger & Pfaffermayr (2009), which tests whether the variance has declined between two points in time, t_0 and t_1.[6] The advantage of this test is that it neither needs long time-spans or high frequency data nor it is biased by pooling over several points in time. The test is based on the β-convergence regression[7]:

$$p_k^i(t_1) = \alpha_k + \pi \cdot p_k^i(t_0) + u_{kt_1}^i \tag{2.1}$$

$p_k^i(t_1) = ln(P_k^i(t_1))$ is the natural logarithm of price P of product (or product group) k in country i in period t_1. Using OLS, it is regressed on $p_k^i(t_0)$, the log price in period t_0. As the mean – i.e., the average of the log prices – towards which reversion is measured might be product specific, a product fixed effects α_k is employed. Mean reversion measures the decrease of deviations of prices from the (product k specific) country groups average. The coefficient π can then be interpreted as the remainder of period t_0's price deviation in period t_1 and $\pi < 1$ implies mean reversion. The smaller is π, the stronger is the mean reversion.

In addition to the mean reversion, there are random shocks captured in the error terms u_{kt}^i. Hence β-convergence does not necessarily imply that price dispersion actually decreases.

For the Wald-test whether the variance σ_t^2 declined over time against the null hypothesis that the variance remained constant ($H_0 : \sigma_{t_0}^2 = \sigma_{t_1}^2$ vs. $H_1 : \sigma_{t_0}^2 > \sigma_{t_1}^2$), Egger & Pfaffermayr (2009) derive the following test statistic, which is $\chi^2(1)$ distributed[8] :

$$W_0 = \frac{N(\hat{\sigma}_{t_0}^2(\hat{\pi}^2 - 1) + \hat{\sigma}_u^2)^2}{4\hat{\sigma}_{t_0}^4 \hat{\sigma}_u^2} \sim \chi^2(1) \tag{2.2}$$

To calculate this one-sided test, the estimated variance of the regression's residuals $\hat{\sigma}_u^2$ and the corresponding coefficient $\hat{\pi}$ as well as the number of observations N are taken from the β-convergence regression above. In addition, we need $\hat{\sigma}_{t_0}^2$, the estimated variance of prices at the initial date t_0, which is defined as:

$$\hat{\sigma}_{t_0}^2 = \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{I_k} (p_k^i(t_0) - \bar{p}_k(t_0))^2 \tag{2.3}$$

Here K is the total number of products (or product groups), I_k is the number of reporting countries for product (or product group) k and $N = \sum_{k=1}^{K} I_k$ is the total

[6] Another approach would be to look at non-parametric methods as in Quah (1996), who shows that even without σ-convergence, "convergence clubs" may evolve. In our research we follow the parametric σ-convergence literature, as we look at how exogenously given groups evolve.

[7] In the literature, an equivalent formulation of the regression above exists (hence the name β-convergence):

$$\Delta p_k^i(t_1) = \alpha + \beta \cdot p_k^i(t_0) + u_{kt_1}^i$$

Using this formulation, mean reversion is observed if $\beta < 0$.

[8] The same test can be used to test for divergence:

$$H_0 : \sigma_{t_0}^2 = \sigma_{t_1}^2 \text{ vs. } H_1 : \sigma_{t_0}^2 < \sigma_{t_1}^2$$

number of observations. For the 34 product groups, I_k always equals 8, 15 or 23 respectively. For the individual products, I_k varies.[9] Note that price convergence is conditional on the product (but not on the country), since when calculating the variance in equation (2.3), the deviation of each price p_k^i from $\bar{p}_k$ is used for each product (or product group) k. Alternatively, the price data can be demeaned, i.e., normalized such that $\bar{p}_k(t) = \frac{1}{I_k}\sum_{i=1}^{I_k} p_k^i(t) = 0$ for $t = t_0, t_1$ and each k (Egger & Pfaffermayr, 2009, p. 460, footnote 5). $p_k^i(t)$ is then actually the deviation of country i's price from the average price across countries (e.g., across EU15 countries). Price level indices in logs are generally demeaned by definition, however, for considering the different groups of countries (EU8, EU15, EU23), also price level indices have to be renormalized such that the mean of the respective country group is equal to zero (e.g., $\bar{p}_{k,EU8} = 0$). With properly demeaned data, the product fixed effects α_k in equation (2.1) will equal zero.

The considered time period covers the years of the hike and the subsequent fall in world market food prices in the years 2007 and 2008 and also the recession year 2009. Insofar as these shocks affect prices in the countries considered equally, the effects are captured in the averages $\bar{p}_k$ and price convergence is unchanged. However, these shocks may affect prices differently or at different speed in different countries, which could then be interpreted as shocks u_k^i counteracting price convergence.

2.4.2 Convergence decomposition

For the interpretation of the results on price convergence or divergence, we decompose the variance of a single product group k into within-groups and between-groups variances:[10]

$$\sigma^2_{k,EU23} = \frac{15}{23} \cdot \sigma^2_{k,EU15} + \frac{8}{23} \cdot \sigma^2_{k,EU8} + \frac{15}{23} \cdot \bar{p}^2_{k,EU15} + \frac{8}{23} \cdot \bar{p}^2_{k,EU8}$$

where $\sigma^2_{k,EU23}$ ($\sigma^2_{k,E15}$, $\sigma^2_{k,E8}$) is the variance of product k's prices within EU23 (EU15, EU8) and $\bar{p}^2_{k,EU15}$ ($\bar{p}^2_{k,EU8}$) is the squared mean of product k's prices within EU15 (EU8). In Appendix A.1.1, we show that for the special case of price level indices, where $\bar{p}_{EU23} = 0$, the decomposition can be further simplified, aggregated over products groups k and extended to changes (e.g., $\Delta\sigma^2_{EU23} = \sigma^2_{EU23,t} - \sigma^2_{EU23,t-1}$):

$$\Delta\sigma^2_{EU23} = \frac{15}{23}\Delta\sigma^2_{EU15} + \frac{8}{23}\Delta\sigma^2_{EU8} + \frac{8}{15}\frac{1}{34}\sum_{k=1}^{34}\Delta\bar{p}^2_{k,EU8} \tag{2.4}$$

[9] This expression is more general than in Egger & Pfaffermayr (2009), as it does not restrict its use to cases with the same number of reported observations for every observed entity (e.g., I firms in each of K industries).

[10] The basis of this decomposition is the same as in the one-way analysis of variance (ANOVA). While ANOVA then uses this decomposition to test if the subgroups' population means are the same, we want to know to what extent the variance between or within subgroups has changed. For the special case of only two subgroups, ANOVA is actually identical to the t-test (Rabe-Hesketh & Skrondal, 2008).

The estimated change of the variance within EU23 equals the weighted sum of variance changes within the subgroups EU15 and EU8, plus a term measuring the change of the dispersion between the two subgroups (between-groups variance). To estimate the significances of the first two terms, the Wald-Test proposed above is used. The latter term's significance can be estimated with a simple paired t-test or a Wilcoxon matched-pairs signed-ranks test. Note however, that this equation only refers to the semi-aggregate data at the product group level, as due to the missing data at the level of individual products, the weights would differ for the individual products.

2.5 σ-convergence

In this section, we analyse σ-convergence both at the level of individual products and of the product groups. Significant β-convergence is found in all cases, i.e., both for products and for product groups and for all time periods considered.

2.5.1 Convergence within country groups at a disaggregate level - products

As mentioned in section 2.3, the set of prices observed has somewhat changed over time and not all prices are reported in all countries. We only include those products, for which prices are reported at all three dates (2003, 2006 and 2009) and for which more than 50% of the observations are available in both EU8 and EU15, which leaves us with 147 products[11]. This ensures that for every product we indeed measure convergence towards EU8/15/23 averages and not yet another subgroup of these.

Table 2.1: *σ-convergence results at the product level.*

	K	N	**2003-2006**	**2006-2009**	**2003-2009**
$\Delta\hat{\sigma}^2_{EU23}$	147	2817	-0.0282***	-0.0168***	-0.0450***
			(585.50)	(373.71)	(1370.23)
$\Delta\hat{\sigma}^2_{EU15}$	147	1824	-0.0025***	-0.0033***	-0.0058***
			(12.00)	(28.52)	(56.25)
$\Delta\hat{\sigma}^2_{EU8}$	147	993	-0.0189***	-0.0038***	-0.0227***
			(386.40)	(28.55)	(553.16)

Notes: The table reports the change of the variance of pooled product prices in the given period. Wald-test statistics are given in brackets, *** denotes significance at the 1% level.

[11] Varying the cut-off level between 25% and 80% does not change the qualitative results.

Table 2.1 shows the changes of the estimated variances and the results of the Wald-test[12]. Prices have significantly converged in all country groups EU8/15/23 in both subperiods 2003-2006 and 2006-2009, and thus in the entire period 2003-2009. The decrease in the variance of prices within EU15, is rather small. Within EU8 the effect is much stronger in the sub-period 2003-2006 right after accession than in 2006-2009, where the change in variance is roughly of the same size as in EU15.

As described in section 4, these results are obtained when the estimation is run with pooled data and thus refer to the question, if overall, food prices have converged when considered at the disaggregate level of prices for narrowly defined individual products. A further question that might be of interest is how price dispersion for individual products has evolved over time. Convergence results for the single products, that can be found in Figure 2.3, support the results for the pooled data: Between 2003 and 2009 prices have converged for a majority of products within EU23, EU15 and EU8, and this effect is least pronounced in EU15.

2.5.2 Convergence within country groups at a semi-aggregate level - product groups

In this subsection, price convergence at the level of product group price level indices is considered. These price level indices are available since 1995 for EU15 and since 2001 for EU8 (whereas prices for individual products are only available since 2003). The σ-convergence results are shown in Table 2.2.

Table 2.2: *σ-convergence results at the product group level.*

	N	**1995-1998**	**1998-2001**	**2001-2003**	**2003-2006**	**2006-2009**	**1995-2009**	**2003-2009**
$\Delta\hat{\sigma}^2_{EU23}$	782			0.0115***	-0.0249***	-0.0162***		-0.0411***
				(136.59)	(609.01)	(538.64)		(1309.90)
$\Delta\hat{\sigma}^2_{EU15}$	510	-0.0092***	0.0027***	-0.0004	0.0042***	-0.0018***	-0.0045***	0.0024***
		(334.68)	(39.76)	(0.52)	(87.16)	(15.72)	(42.95)	(24.28)
$\Delta\hat{\sigma}^2_{EU8}$	272			0.0007	-0.0143***	0.0046***		-0.0097***
				(1.56)	(748.82)	(168.14)		(224.07)

Notes: The table reports the change of the variance of pooled product group price level indices in the given period. Wald-test statistics are given in brackets and *** denotes significance at the 1% level.

Within EU15, we only find significant σ-convergence in the period from 1995 to 1998 (the period right after Finland, Sweden and Austria had joined the EU) and in 2006-2009, but price dispersion is still higher in 2009 than in 1998. In the other periods, the food price level indices diverge or the results are not significant. Hence since 1998, within the old member countries EU15, prices do not converge any more. Looking at EU23 in the period 2001-2003 right before the EU enlargement, the product groups' price level indices significantly diverge, but they converge

[12]For this one-sided test, the critical values of the $\chi^2(1)$ test statistic at the 10%/5%/1% significance level are 2.71/3.84/6.63.

strongly in the periods after the accession. In our last subgroup (EU8), price dispersion fell after these countries joined the EU and slightly rose again in the following period 2006-2009. Figure 2.4, shows the results for individual product groups for the period 2003-2009. In EU23, only two product groups diverge, whereas price dispersion falls for all other product groups. Price level indices also converged for most product groups in EU8. In EU15, the picture is mixed, and on the whole, price divergence dominates.

Comparing results for individual products and for product groups

Comparing table 1 and table 2 shows, that the results for the disaggregate level of individual products and the semi-aggregate level of product groups differ in some cases. Results are sensitive to the level of aggregation. While at the level of invidual products, prices converge in all cases considered, this does not hold for the product group price level indices. For EU15, product group price level indices diverge in the period 2003-2006, and this effect is so strong, that there is also divergence for the longer period of 2003-2009. In the EU8, price level indices diverged between 2006 and 2009, but in the longer period of 2003-2009, this divergence is dominated by the strong convergence in the earlier period of 2003-2006.

However, the differing results are not inconsistent with each other. There are several explanations why convergence at the level of individual products may differ from convergence for the product groups. First, we restrict our dataset to products that are sufficiently common (50% of EU8 as well as 50% of EU15 countries reporting prices), which most likely differs from the products used in each country for product group aggregation. These most common products might converge faster than the other ones. Second, the products in a given basket might enter with different weights. Thirdly, when aggregating the prices from products to product groups, relatively cheap and relatively expensive products average out. As discussed in section 2.3, it is possible that a price change contributes to convergence at the product level but to divergence at the level of product groups (and vice versa).

Note however that the result of price convergence in EU23 both in the periods 2003-2006 and 2006-2009 is robust. When considering the long period 2003-2009 after accession, the result of price convergence within the EU8 also holds both at the disaggregate level of products and of product groups. In contrast, within the group of the old member countries EU15, prices converged between 2003 and 2009 when considered for individual products, but diverged at the product group level.

Thus it may indeed be relevant whether price convergence is considered based on disaggregate or on semi-aggregate data. Prices for narrowly defined products, that are widespread in Europe, converged in EU15 on the whole, but this is not reflected at the level of price level indices for the product groups.

Figure 2.3: *σ-convergence ($\Delta\sigma^2$) by products, 2003 - 2009.*
The graphs present the change of the variance of every product in the dataset for EU23, EU15 and EU8. The first bar depicts the pooled convergence result of the respective country group. Light blue bars are not statistically significant. *Source: Eurostat data, own representation.*

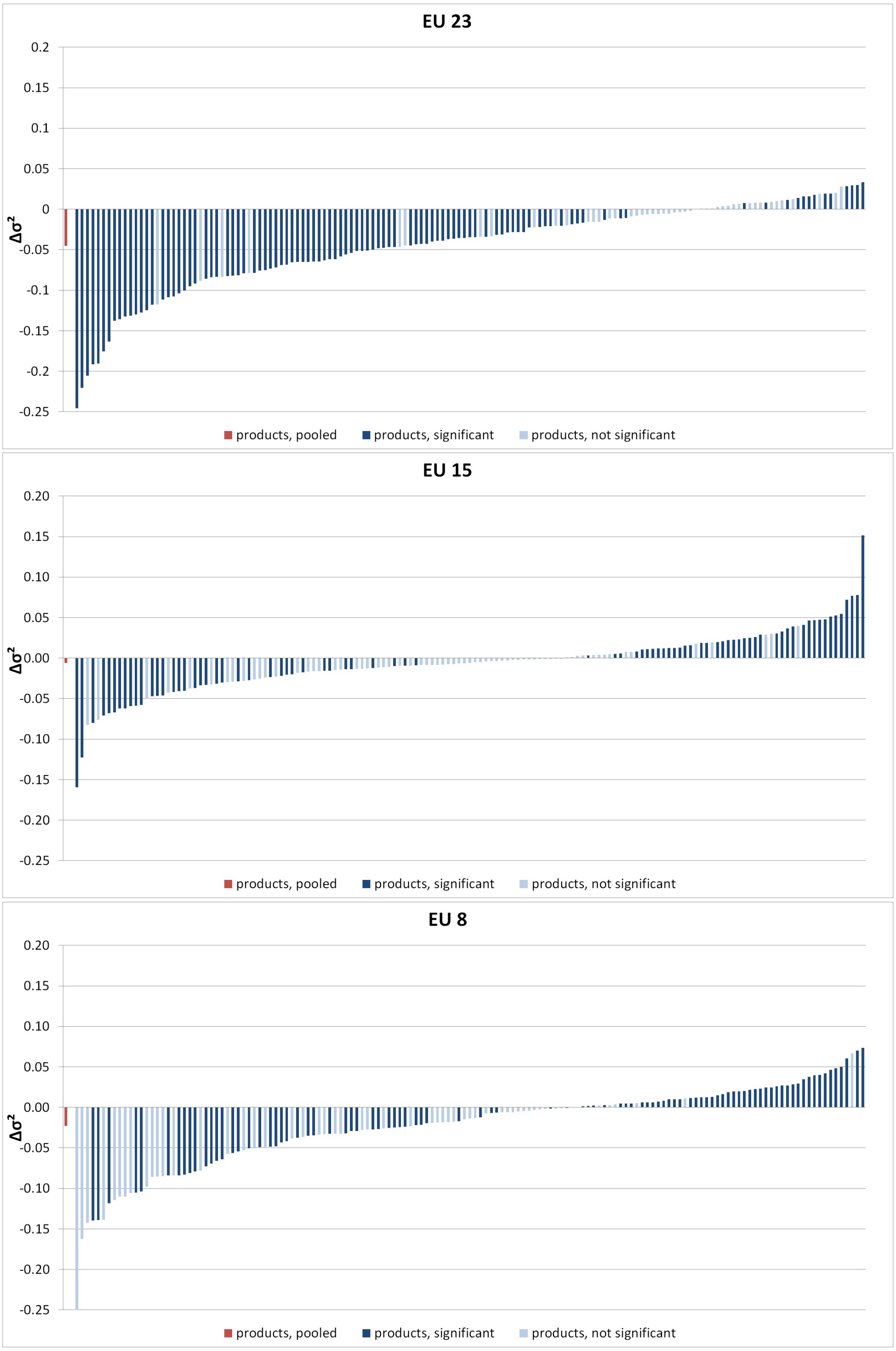
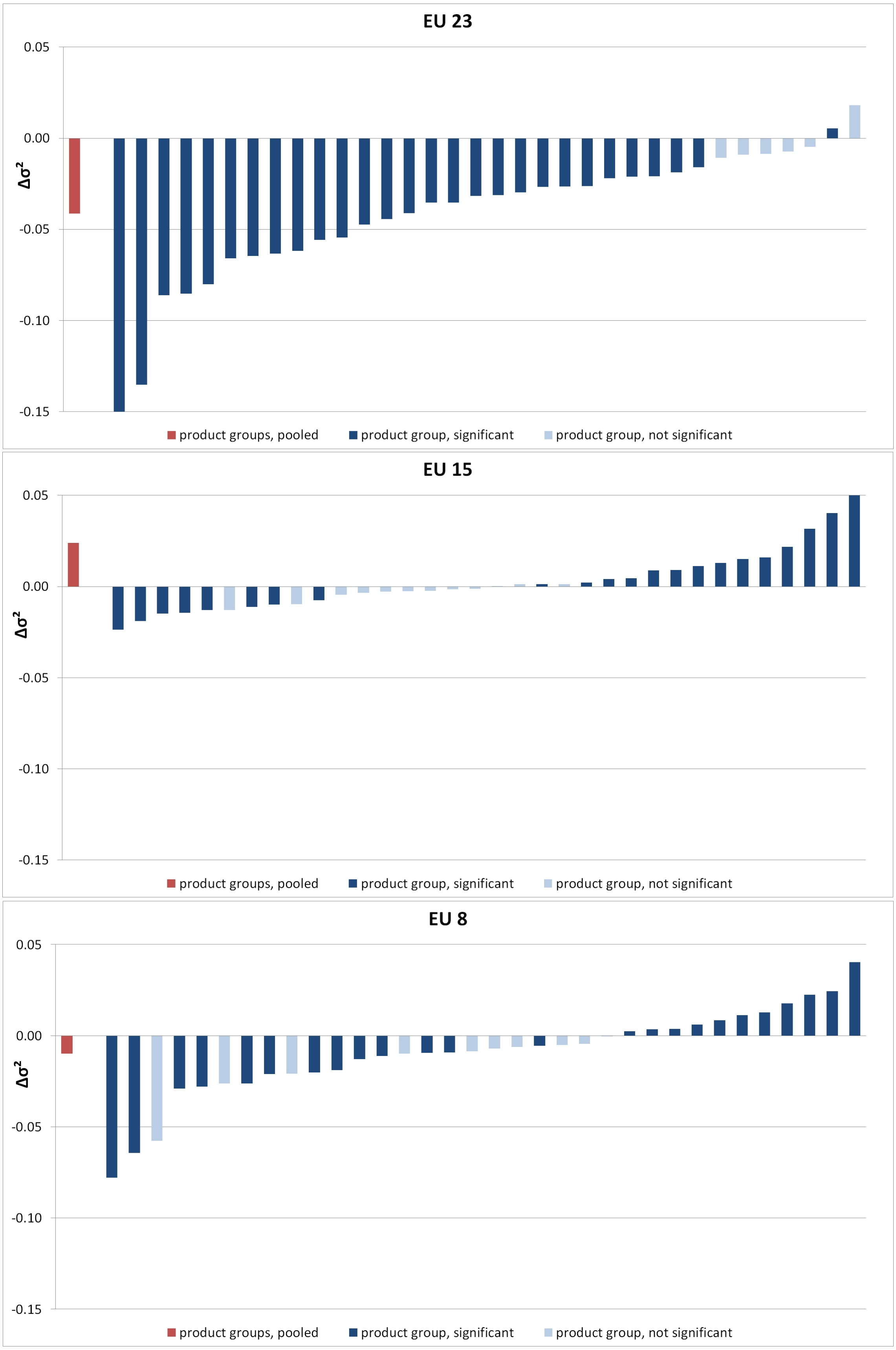


Figure 2.4: *σ-convergence ($\Delta\sigma^2$) by product groups, 2003 - 2009.*
The graphs present the change of the variance of every product group in the dataset for EU23, EU15 and EU8. The first bar depicts the pooled convergence result of the respective country group. Light blue bars are not statistically significant. *Source: Eurostat data, own representation.*



2.5.3 Convergence between country groups

The σ-convergence results within country groups suggest, that the strong convergence within EU23 could be caused by convergence between EU15 and EU8 and not mainly by convergence within these two subgroups. To further investigate this point, we look at the variance decomposition at the product group level (see section 2.4.2). Table 2.3 summarizes the results.[13]

Table 2.3: *σ-convergence decomposition at the product groups level.*

	N	**2001-2003**	**2003-2006**	**2006-2009**	**2003-2009**
$\Delta\hat{\sigma}^2_{EU23}$	782	0.0115***	-0.0249***	-0.0162***	-0.0411***
$\frac{15}{23}\Delta\hat{\sigma}^2_{EU15}$	510	-0.0002	0.0027***	-0.0012***	0.0015***
$\frac{8}{23}\Delta\hat{\sigma}^2_{EU8}$	272	0.0002	-0.0050***	0.0016***	-0.0034***
$\frac{8}{15}\frac{1}{K}\sum_{k=1}^{K}\Delta\bar{p}^2_{k,EU8}$	272	0.0115***	-0.0227***	-0.0166***	-0.0394***

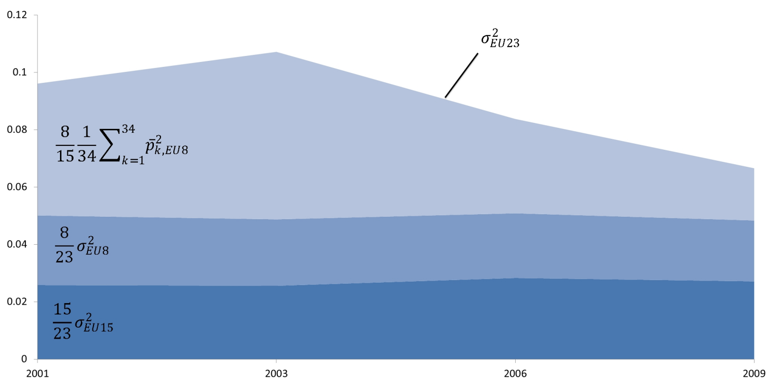
Notes: The table presents the results of the convergence decomposition on the product group level. The change of variance in the whole EU23 (first line) can be decomposed into the weighted change of EU15's variance (second line), the weighted change of EU8's variance (third line) and the weighted change of the squared mean of the EU8 price level indices (forth line). *** denotes significance at the 1%-level.

In the period before the accession (2001-2003), price dispersion both within EU8 and EU15 remained virtually unchanged. Price divergence within EU23 can be explained by the fact that price differences between EU15 and EU8 increased. In the period 2003-2009, i.e., after the enlargement shock, the decline in between-groups variance amounts to more than 90% of the total convergence. In the last sub-period considered (2006-2009), there would not be any convergence in EU23 without convergence between the old and new member states. Hence the driving force behind the strong convergence within EU23 after the EU enlargement in 2004 is not convergence within these two subgroups, but between them. In 2003, $\bar{p}_{EU8}$ (the average of price level indices of all product groups in all EU8 countries) was 33% lower than the EU23 average while it was only 18% lower in 2009. Figure 2.5 illustrates the variance decomposition graphically.

One question that comes to mind is whether only the EU8's prices rose or if prices actually moved towards each other. This cannot be distinguished when using price level indices. By construction, if one subgroup becomes relatively more expensive, the mean of the whole sample changes and hence the other subgroup has to become relatively less expensive. However, if we look at food prices at the individual product level, we find that the average price increase per year between 2003 and 2009 within EU15 was only 0.9%. In real terms prices then have actually moved towards each other.

[13] The results of the paired t-test and the Wilcoxon test, whether $\frac{8}{15}\frac{1}{34}\sum_{k=1}^{34}\bar{p}^2_{k,EU8}$ has changed between two given points in time, can be found in appendix A.1.2 .

Figure 2.5: *Variance decomposition on the product groups level over time.*
The graph presents the composition of the variance of EU23's product group price level indices between 2001 and 2009. The weighted EU15 variance ($\frac{15}{23}\hat{\sigma}^2_{EU15}$), the weighted EU8 variance ($\frac{8}{23}\hat{\sigma}^2_{EU8}$) and the weighted, squared mean of EU8 price level indices ($\frac{8}{15}\frac{1}{34}\sum_{k=1}^{34}\bar{p}^2_{k,EU8}$) add up to the variance in the whole EU23 ($\hat{\sigma}^2_{EU23}$). *Source: Eurostat data, own representation.*



2.6 Conclusion

As on a perfectly integrated market, a good can only have one price, it is widely expected that ongoing market integration leads to price convergence, thereby enhancing efficiency. However, from a theoretical point of view, prices need not necessarily converge, as for instance retailing services are non-tradable and arbitrage will be reluctant to react to retail price differences. Nevertheless, if price convergence occurs, this can be interpreted as an indicator of market integration, showing that the internal market project, which aims at increasing welfare, works well. This idea can be applied to the new Central and Eastern European member states: price convergence may indicate that these countries successfully integrate into the markets of the European Union.

Using a unique micro level data set on retail food prices from the Eurostat-OECD PPP Program, we find that in the lead-up to accession (2001-2003), food prices diverged within EU23. In contrast, price convergence has occured after 2003, indicating that the enlargement had a vital effect on food price dispersion. Given that price convergence is found, the underlying dynamics and subgroup behavior is of interest. The driving force behind price convergence in the European Union is convergence between the two groups of countries EU8 and EU15, with the difference of average prices for food products of these two groups becoming smaller. A subsequent research question could be whether convergence clubs as described in Quah (1996) will endogeneously emerge. The possibly evolving clubs would not necessarily have to coincide with the exogeneously determined country groups EU8 and EU15. We leave this question to further research.

Chapter 3

Sex work: An introduction and a new data set

Prostitution or sex work is often called the "oldest profession" in the world. While the "meat for sex" theory argues that early predecessors of mankind traded food for sex as early as 3.5 million years B.C. (Alfred, 2011, p. 238), the first recordings only date back as far as 600 B.C., when state brothels were introduced in ancient Greece (Glazebrook & Henry, 2011, p.31). Although from an economic point of view there has been a market for quite some time, from a legal perspective (and perhaps in a literal meaning as well) the claim of the "oldest profession" does not hold.

Over the last three millennia, selling sexual services has always been part of the black or gray market, but has not been a legitimate business (and in many countries still is not). (Semi-) Illegal markets differ from legal markets in that, for example, basic legal rights granted in a legal market – like being payed after having provided a service – cannot be enforced. Hence, price formation and market entrance follow different rules than in a legal market. Coping mechanisms to solve these legal problems might result in additional cost and criminal activities (for a further discussion see Becker et al., 2004; Feige, 1990). Although the German market for sexual services has been legalized in 2001, it is still in a transition process and, for example, costs due to social stigmatization still exist.

In this chapter, we first give some background on sex work in general and then present a newly created dataset on sex work in Germany. To build this dataset, we wrote a script to download all information on services offered between January, 16 in 2011 and September, 9 in 2012 on *www.gesext.de*, an online platform where sexual services can be sold. As chapters 4 to 6 are based on sub-samples of this dataset, we present the creation and descriptive statistics of the full sample in this chapter in detail.[1]

[1] Although both men and women act as sellers in this market, in this work we focus on female sellers. This is especially due to the lack of quality data on male sellers in our dataset, but also because the female part of the market is by far larger than the rest.

The remainder of the chapter is structured as follows: The next sections gives a short historical overview on the legal status of sex work in Europe and especially in Germany(Section 3.1) and discuss existing theories on price formation in the market for sexual services (Section 3.2). Section 3.3 looks at modes of data generation as well as existing datasets. Section 3.4 then describes how the *www.gesext.de* dataset was generated and presents its basic descriptive statistics. In Section 3.5, regional aspects of selected variables of the dataset are compared and Section 3.6 concludes.

3.1 A historical overview of sex work in Germany (and Europe)

As mentioned above, the first recordings of sex work in Europe are from ancient Greece and Rome, where sexual services were often provided by slaves without any personal rights. Already there, but also during the following centuries up to the 20th century, sex work has been deemed immoral and was often illegal (Schmitter, 2013). This led to a constant struggle between ethical (religious) doubts about the morality of women selling (or being forced to sell) these services and men nevertheless keeping on demanding them. Since the early days of Christianity, the woman was deemed to be the cause of all evil, seducing men into buying sexual services, but this was still considered better than permanently violating the bible's sixth commandment: "Thou shalt not commit adultery" or commit a sexually motivated crime. This view has stayed alive up until the 20th century and explains why the church did not use its influence to abolish prostitution (Schmitter, 2013, p.23).

From the 13th to the 15th century, city-run brothels (so called womens' houses) were introduced. For easier identification, women had to wear special dresses and live in these houses, hitherto stigmatizing them (Sanger, 2006). These womens' houses were abolished with the reformation in the 16th century, only to be reintroduced in the 18th century.[2] In the mid of the 19th century and during the beginning of the 20th century, the ambiguous statutory state of the business became particularly clear. On the one hand, it was illegal to advertise or rent apartments to sell sexual services, ban zones were installed by local authorities and claims for payment after having delivered these services could not be enforced. On the other hand, buying and selling sexual services was largely tolerated (during the second world war, the military even operated their own brothels in regions occupied by German forces) and sex workers had to register with local health authorities and even to pay taxes (Schmitter, 2013, p.23).

In Germany, this legal nirvana was ended by several court rulings in the 1990s and by a new legal setting in 2002 (Kavemann & Steffan, 2013). Sex work still

[2]Martin Luther famously compared state run brothels to having state run houses for murderers (Schmitter, 2013, p.23)

cannot be regarded as a job like every other one: the job centers do not suggest sex work as a possible source of income to the unemployed and buyers cannot enforce the fulfillment of a concluded contract as in other markets. Yet, with the introduction of this new law, sex workers were accredited many rights and duties normal workers have. These include, among others, the access to the governmental health insurance system, to rent apartments for selling sexual services, to sue for payment thereof, but also the duty to pay business tax.[3] In most cities, the so called "Duesseldorf procedure" has been adopted (and had already been in use in several cities before the new law was introduced) requiring women to pay a daily flat tax when offering services. In addition, some cities have introduced a tax on renting apartments for offering sexual services (e.g., 10 Euro per square meter and year in Stuttgart; Preiss, 2012).

From a European perspective, countries can be divided into four categories: 1) selling sexual services is illegal (e.g., Croatia, Serbia), 2) selling sexual services is legal, but buying them is illegal (Sweden, Norway), 3) selling sexual services is legal, but brothels, pimping, and sometimes advertising and street work are illegal (e.g., France, Spain, United Kingdom) and 4) selling sexual services is legal (e.g., Austria, Germany). The German laws are very progressive as sex workers have extensive legal and social rights, but what all European countries have in common is that forced prostitution and human trafficking are illegal (for a further discussion see Cho et al., 2012). To sum up, in Germany sex work has always been in a legal gray (and moral black) area, but with changing social and moral attitudes of society and law makers, it has been decriminalized within the last 15 years. Furthermore, as sex work remains illegal in other European countries, Germany has become a destination for sex tourism from neighboring countries (Barnett, 2013). However, the exact market size and the number of sex workers remain unknown. Estimates range from 64,000 to 200,000 sex workers, earning up to 14.5 billion Euro (Reichel & Topper, 2003; Kavemann & Steffan, 2013).

In general, this market can be divided into three sections: Street level sex services, brothels and escort services (private offerings) – with average income and personal safety of the sex worker increasing in this order. While the exchange and contracting of the first two sectors takes place to a large extent in the "offline" world by simply visiting certain locations, in the third part appointments are made by telephone or via the internet and the meeting takes place at an arranged location. The upcoming of the internet has had a strong impact on this market. For sellers it has become easier and safer to advertise and find customers. Hence, the supply side of this market has increased. Furthermore (although only for online sellers), the risk of becoming a victim of violence and catching a sexually transmitted dis-

[3] As a side-effect, also the renting out of rooms to sex workers has been decriminalized, leading to stronger competition in the market and the emergence of very large brothels. Europe's largest brothel – situated in Stuttgart – spans an area of $5800m^2$.

ease, as well as getting caught by the police (in countries where sex work is illegal), has decreased (Cunningham & Kendall, 2011b). Today, in Germany exists a large number of websites on which sex workers can advertise or sell their services.

3.2 Theory on the market for sex

The market for sexual services differs from normal markets in many ways. There is a moral stigma to selling sexual services, the risk of physical violence exists and in some countries, selling sexual services is even illegal (see above). This has several consequences on the price formation in the market and the willingness to enter it. First of all, moral stigma drives up the cost of providing sexual services, as sex workers have to be compensated. Second, in illegal markets the lack of access to the legal institutions does not allow sellers to enforce contracts. This risk has either to be factored in or sellers pay for protection by, e.g., organized crime. Third, participants might risk punishment when their participation is discovered, which again can be factored in or be minimized by paying for protection (for a further discussion on illegal markets see Ruggiero et al., 2002). All three points lead to higher prices and lower entrance than in legal markets. On the other hand, cost might be lower, as no taxes have to be payed and no state regulations have to be complied with.

Compensation for opportunity cost

The existence of a wage premium for sex workers compared to non sex workers is generally acknowledged by the theoretical (and especially the empirical) literature on price formation and market entrance in the market for sexual services. The literature also acknowledges that this wage premium must be due to some sort of market power and / or compensation for opportunity cost. Selling sexual services is probably best modeled as monopolistic competition, an idea first introduced by Chamberlin (1933). Although most articles in the sex work literature use the general idea of sexual services being differentiated goods, the only adaptation of this model to sex work is by Cunningham & Kendall (2010b). Each sex worker sells one or several varieties of a given product (sex) and has some market power for her variety. In the standard model of monopolistic competition, sellers and demand are symmetric, which in the long run results in a unique market clearing price at zero profits.[4] However, if sellers and demand are heterogeneous, there need not be just one price. Sellers can charge a higher price if they have higher (opportunity) cost to provide their variety or higher market power due to offering a more unique or more demanded variety only they can provide. For example, the demand for high class – beautiful and smart – escort services can only be met by very few sellers, who need to pay for expensive dresses

[4] As long as profits can be made, other sellers will enter the market and buyers substitute to these new varieties, hence lowering demand for existing varieties. Market power is gone and prices are lower.

and manicure, and could work in a good paying non sex work job as well.[5]

This idea is also acknowledged in the other models concerning price formation in the sex work market. While Arunachalam & Shah (2012) as well as Cunningham (2011) focus on higher market power of a given sex worker compared to other sex workers – due to being more beautiful, offering rare services, etc. – other articles focus on wage premia compared to non sex workers. Here, Rao et al. (2003), Gertler et al. (2005), Robinson et al. (2011) and Arunachalam & Shah (2013) mainly focus on opportunity costs. They argue that sex workers have to be compensated for risks such as being arrested, encountering violence or catching a sexually transmitted disease.[6] A substantially broader interpretation of opportunity cost (e.g., social exclusion, boredom while waiting for customers, loss of recreational sex pleasure, highly skewed lifetime earnings due to early retirement, etc.) and excess demand for certain services are the explanation for wage premia in Cameron (2002). In addition, there are some alternative models worth mentioning that rather discuss a seller's decision to enter the market and have brought different new aspects into this literature.

Market entrance

In a theoretical paper Edlund & Korn (2002) model the market for sex work such that women are confronted with a binary choice: they can either choose to get married or to become a sex worker. Both groups provide sex to men.[7] While the first group has a normal job and additionally benefits from their husbands earnings, in an equilibrium the latter group has to be compensated for their losses on the marriage market. Apart from introducing a new form of opportunity cost (foregone earnings from getting married), an interesting application of this model is that migrants should be able to offer sexual services at lower prices, as their potential spouses at home are less likely to find out about their job as a sex worker and hence marriage opportunities are less decreased.[8]

[5]Empirically, the wage of a given sex worker could then be determined by adding up the price premium for 1) every single attribute of this sex worker and 2) the circumstances under which he or she offers services, both compared to the average sex worker or the average worker.

[6]For more on a risk-premium on unprotected sex see chapter 4 of this thesis

[7]An interesting assumption about the female utility function in this model is that in both cases they *"do not care for sex, but derive utility from their children [...] and consumption* (Edlund & Korn, 2002, p192).

[8]In an answer to this article Arunachalam & Shah (2008) strongly reject Edlund and Korn's theory and defend the idea of wage premia just being risk compensations. Using a dataset on sex work in Ecuador and Mexico, their argument is two-fold. First, there are many sex workers that get married and at a young age, when wage premia are highest, sex workers are more likely to be married than non sex workers. Second, they also find a wage premium for male sex workers, which should not exist if it was a compensation for forgone marriage opportunities. However, their study does not allow to completely reject the Edlund and Korn model. In their wage premium regression, they do not distinguish between married and unmarried sex workers. Additionally, the fact that at younger ages sex workers are more likely to be married than non-sex workers does not reveal anything about the likelihood of the other sex workers to get married later on. In fact, they state that in Ecuador the average sex worker – unconditional on age – is 16% less likely to be married.

In Della Giusta et al. (2009a) and Della Giusta et al. (2009b), the central argument is based on social stigma – the cost of society finding out that an individual is a (part-time) sex worker. Individuals have a normal job and if income from providing sexual services exceeds reputation losses they will start to participate in the market of sexual services. In general, their model assumes that men or women could both be sellers, but with income differentials, the group with the higher income (in most societies men) will demand and the other group will provide sexual services. In their model the threshold to become a seller rises, the more "free pre-marital sex" is provided and the higher income of the sellers and / or reputation losses are. Furthermore, price premia for selling sexual services increase the higher income from normal jobs on the seller side is and the higher potential reputation losses of both the buyer and the seller are. This has three interesting implications: First, in their model rich women can also be consumers of sexual services. Second, with women being able to control pregnancy, the cost of offering "free sex" decreases and less sexual services for pay will be provided (and if they are provided the price premium is lower). Third, as in the Edlund & Korn (2002) model, for migrant sellers reputation losses are lower and they can provide sexual services at lower prices. The same is true if social stigmatization decreases.

The idea of the Della Giusta et al. (2009b) model is extended in Cunningham & Kendall (2010b), distinguishing between fixed and variable cost of supplying sexual services (stigma and correspondingly lower marriage opportunities vs. risk of diseases, arrest, etc.). With lower fixed (and also variable) cost, workers will enter the market more frequently. Given that they entered, they will provide more services due to their lower variable cost. This result again leads to three interesting new implications. First, with a negative income shock an individual is more likely to become a seller. Second, if one assumes that more educated individuals have higher income in normal jobs, which they could loose by entering the market, and shorter spans of unemployment, their fixed cost are higher. With better access to credit markets, negative income shocks can more easily be absorbed. Third, given that they already have entered the market, these sellers should have lower variable cost, as they have less violent, higher-status clients, a lower risk of getting arrested and can bundle sexual and escort services (nice dinners or opera visits).

Last but not least, Cameron & Collins (2003) as well as Collins & Judge (2008) present a model of male participation in the market. Similarly to the models above, opportunity costs such as search time for sexual encounters, gaining additional pleasures from variety, etc. are used to explain if a man enters the market or not. Their models imply that by changing the regulatory setting, policy makers can influence male participation and risk-taking.

In part, the propositions made by the models presented here have been tested. The most researched part of the literature are the wage premia and incentives to offer risk-taking, but there is generally a lack of good data concerning this market. The next section gives an overview of the existing datasets.

3.3 Data availability and quality

Although sex work has always been a more or less open part of societies, economists only recently have started to pay attention to this market. There are some earlier studies and books on the economics of prostitution (e.g., Reynolds, 1986), but nearly all existing empirical studies have been conducted during the last 10 to 15 years. This is mainly due to the data availability being just recently enhanced by government funded programs (e.g., Arunachalam & Shah, 2013) and the existence of internet platforms (e.g., Cunningham & Kendall, 2010b).

Whereas there exist at least some macro level (aggregate on the national level) estimates on the volume of the sex work sector in Germany (e.g., Nagel, 2013), there is no micro level data (i.e., information on single transactions) available in Germany. Until recently, this kind of data could only be acquired by studying police reports or directly surveying sex workers. In gathering information from match maker platforms in the internet, a third way has emerged. While in this thesis we use the third strategy to create a dataset, we shortly discuss the advantages and disadvantages of all three information gathering strategies and present existing datasets along the way. An overview of most of these datasets' results can be found in chapters 4.1, 5.1 and 6.1, where we present the empirical results of the literature corresponding to the respective chapter's research question.

Police reports are only a source of information in regions where the provision and consumption of sex work is illegal. Obviously these are hard to get hands on, and if one can look into this kind of data (e.g., as in the US; Della Giusta et al., 2009a), the subject pool is highly biased (arrested sex workers or clients) and might be unwilling to answer questions truthfully, trying to minimize legal consequences.

Surveys have the advantage that, at least in theory, they can capture all parts of the market, but they as well might be subject to biases. First, they rely on the possibility to create a representative sample and the sex workers' willingness to cooperate. This might introduce a bias due to self-selection into the study of sex workers willing to answer at all and unavailability of individuals from certain sectors of the market. For example, in most studies using survey data, the online market is not represented (e.g., Rao et al., 2003; Gertler et al., 2005; Arunachalam & Shah, 2013). Second, they rely on the sex workers' ability to remember all details correctly and willingness to answer truthfully. Questions into sex work are a very private and sensitive matter. For example, information on transaction prices might be biased by fear that the surveyor is related to tax authorities in some way or the rate of condom use might be exaggerated if the survey is carried out as a part of an anti AIDS campaign. In general, survey data can be used and biases corrected, but survey data might have their limits (Bertrand & Mullainathan, 2001).

Finally, extracting data from an (online) match-maker platform has the same disadvantage as survey data have, in that the observed individuals only represent a part of the market. Only individuals willing to make an offer on the specific platform under observation will be included in the dataset (which for example excludes street sex workers). But one advantage is that at best actual contracts can be observed, including the information on appearances, services offered, etc., which are provided to the other party before a contract is made and the price actually paid. Even if these were wrongly stated, this is the information a potential buyer has and hence in a market economy the price he pays should correspond to his willingness to pay for the specific service offered.

While for Germany, to the best of our knowledge, there is no other dataset on sex work with information on single transactions, internationally there exist some datasets. Most of these datasets are survey based and located in developing or emerging countries. Usually surveys are conducted by people trusted by the sex workers (medical staff or former sex workers) and ask for personal information as well as details on the last three or more contracts with clients. This is true for all of the following datasets: In the framework of an HIV/AIDS intervention in 1993 Rao et al. (2003) surveyed 608 sex workers in the brothels of one nightlife district in Calcutta, India. In 2001, Gertler et al. (2005) selected the two regions in Mexico with the highest HIV rates and surveyed roughly $1,000$ sex workers identified through the local AIDS program. Gertler & Shah (2011) and Arunachalam & Shah (2008, 2012, 2013) use Ecuadorian data collected in 2003 by a national Ecuadorian STI prevention program, aiming at surveying a representative sample of $3,000$ sex workers all over the country. In 2004 - 2006, de la Torre et al. (2010) interviewed sex workers in Ciudad Juarez, Mexico, who had engaged in unprotected sex prior to the survey. Robinson et al. (2011) had 192 sex workers in a Kenyan border town write journals about their sexual encounters and health status and created a very rich dataset of $19,041$ transactions over the course of three month. Finally Chang & Weng (2012) interviewed 140 randomly selected street sex workers in Taipei city, Taiwan's largest sex district. The only dataset surveying sex workers online is Cunningham & Kendall (2010b), who in 2008 reached out to roughly $13,000$ sex workers in the United States of America and received back 685 questionnaires.

Online platform based datasets include Moffatt & Peters (2004), who in 1998 created a dataset of reviews on male sex workers in the UK and Edlund et al. (2009), as well as Cunningham & Kendall (2010a,b, 2011b) who look at a review website for female sex workers in the United States of America. Both studies utilize downloaded information from *www.theeroticreview.com* on more than 100,000 sex workers with more than 500,000 reviews , who were offering sexual services between 1998 and 2008.

A third dataset is by Cunningham & Kendall (2011a), who downloaded classified advertisement of sex workers from *craiglist.com*, offering services during a three day period of the two big political parties national conventions in 2008. However, these three datasets utilize offer or reviewer data and not actual contracts. In this way, the *www.gesext.de* data we used to create our dataset is unique.

3.4 The *www.gesext.de* dataset

As we are especially interested in micro-level data on concluded contracts, we follow the strategy of downloading data from an online match maker site. In the German market, there exist several of these match-making platforms, of which the largest and most prominent one is www.gesext.de. Individuals can offer all kinds of sex related products and sexual services on this website, which basically works like ebay. Sellers can offer their products and services as an auction with a given end date or at a fixed price or both at the same time. When buyers enter the site, they can filter offers for their region and see all active offers. They can then, depending on the type of offer, either bid on an auction according to their willingness to pay or buy a fixed price offer directly at the fixed price. While an auction will be won by the highest bidder (who has to pay the second highest bid), the fixed price offer will go to the first bidder. A registration is needed both for selling and buying and both parties have to be at least 18 years old.[9] The platform keeps full names and addresses of both parties and exchanges contact information after a contract was made (although the real name will not be provided). Sellers have to pay 15% of the end price to the platform owners as a match making fee, buyers normally pay cash at the beginning of the meeting.

In general there seems to be a market for male as well as female sex workers. However, on this platform men mostly sell services at very low prices, sometimes even for 1 euro, and pay the hotel cost themselves. Due to these mostly negative earnings for male offerings, we restricted our analysis to women offering sexual services.

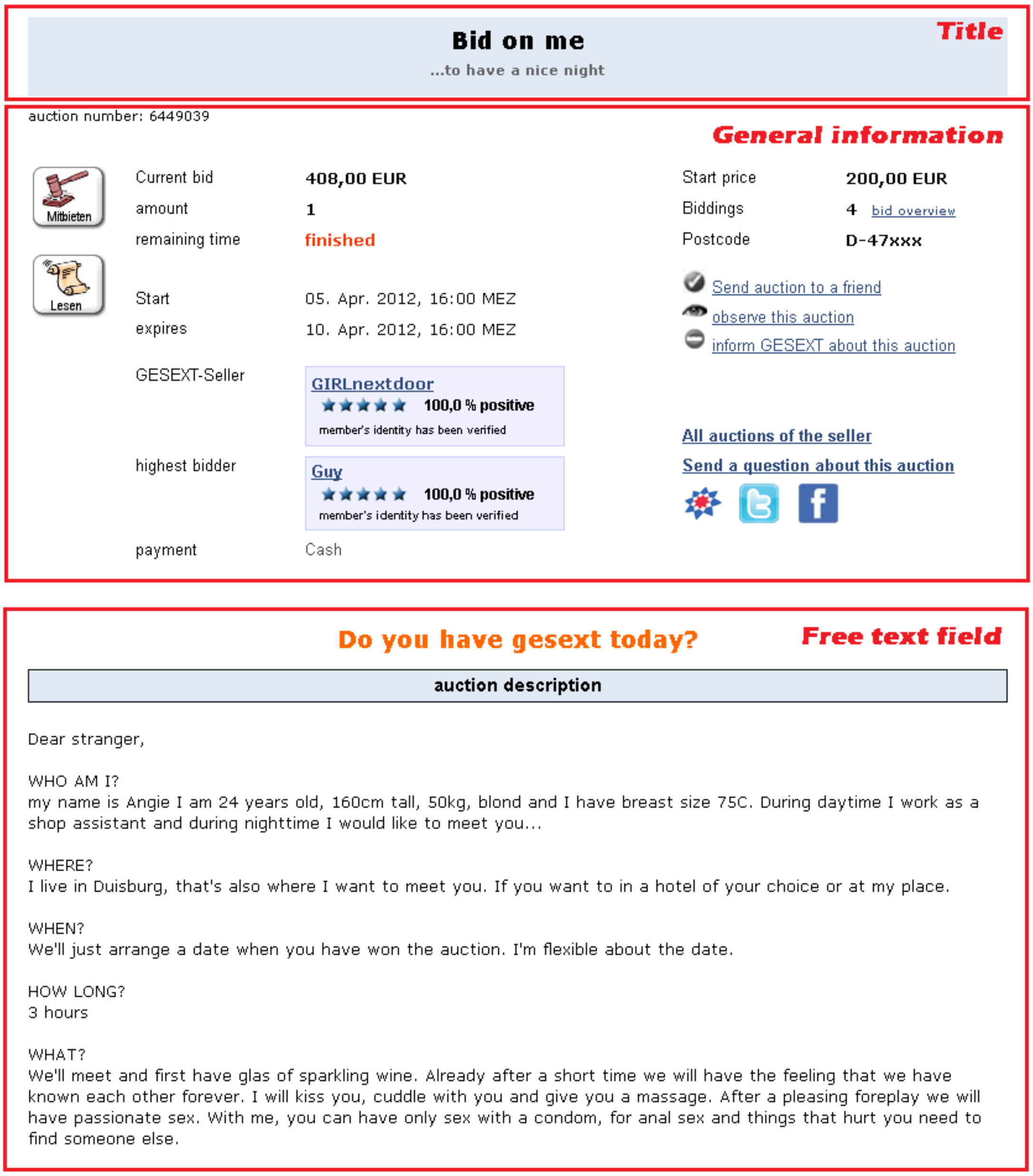
To gather the data we wrote a python[10] script that runs on a server and checks twice a day if new offers have been added to the platform. A second script checks daily if an offer has been finished and adds the URL links of these offers to a list. This list can then be used to download the information on all concluded contracts, which consists of three parts: general information, a free text (including a free text title) and pictures. The offers contain rich information on the female sellers, but unfortunately hardly any information on the sellers (except for their nickname, bidder history and rating).

[9]The age is checked using the individuals identity card number.
[10]Python is a programming language.

General information

Figure 3.1 shows an example of an offer. As the general information part of the data is well structured, we were able to write a script using pattern recognition that can automatically extract this information. Table 3.1 presents all variables we were able to extract automatically from the raw data.

Figure 3.1: *Example offer*
The figure presents an example of an offer at *www.gesext.de*. It consists of three parts: The title, general (structured) information and a free text field. Additionally sellers can post pictures of themselves. *Source: www.gesext.de, translated.*



The script ran from January 16 in 2011 to September 9 in 2012, downloading a total of 102,096 offers posted on this website.[11] Of the 102,096 downloaded offers, 23% of the services were offered as an auction, 44% at a fixed price and 33% as

[11] Descriptive statistics in this chapter deviate from the following chapters, as these look at different sub samples of the full sample presented here.

Table 3.1: *Descriptive statistics - General information*

Variable name	**Type**	**All offers**		**Auctions only**		**Fixed price only**		**Auction & fixed price**	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Number of observations	metric	102,096.00		24,334.00		44,487.00		33,275.00	
Only Auction offered	binary	0.23	0.43	1.00	0.00	0.00	0.00	0.00	0.00
Only Fixed price offered	binary	0.44	0.46	0.00	0.00	1.00	0.00	0.00	0.00
Both auction and fixed price offered	binary	0.33	0.47	0.00	0.00	0.00	0.00	1.00	0.00
Offer sold	binary	0.33	0.47	0.52	0.50	0.30	0.46	0.23	0.42
Start price	metric	179.18	177.14	152.11	195.98	185.76	153.18	190.19	189.97
End price if sold	metric	185.76	207.07	210.29	284.20	167.78	122.84	176.15	162.45
Hourly wage	metric	102.96	56.70	92.74	60.40	114.43	54.46	99.80	50.15
Start date	date								
End date	date								
Number of days offer is online	metric	5.19	5.97	7.13	6.72	3.24	4.36	5.38	5.98
Number of bids if sold	metric	3.78	6.21	8.57	7.46	1.0	0.0	2.38	3.43
Seller nickname	string								
Number of sellers	metric	3,444.00		2206.00		1668.00		1280.00	
Number of offered services per seller	metric	25.96	78.14	9.81	31.13	20.94	61.63	20.94	69.12
Number of sold services per seller	metric	9.65	22.96	5.70	18.78	7.88	15.56	5.87	13.13
Buyer nickname	string								
Number of buyers	metric	14,093.00		6906.00		8049.00		5245.00	
Number of bids per buyer	metric	7.24	579.33	3.52	141.44	5.51	348.25	6.34	355.68
Number of successful bids per buyer	metric	2.36	2.87	1.82	2.12	1.63	1.31	1.43	0.99
2 digit zip code	string								
Number of zipcodes	metric	95.00		95.00		95.00		95.00	
Number of services offered per zipcode	metric	1,074.66	1009.94	256.15	283.56	473.27	475.45	350.26	393.62
Number of services sold per zipcode	metric	349.79	331.84	132.40	169.01	139.78	139.40	80.77	78.56
Seller rating	metric	98.34	3.72	98.17	4.25	98.46	3.57	98.36	3.53
Buyer rating	metric	96.70	9.74	96.56	9.53	96.91	9.62	96.56	10.28

Notes: The table presents the descriptive statistics gathered from the structured part of an offer. Statistics are given for the whole dataset as well as for the subsamples of services offered as an auction only, offered at a fixed price only and services offered as both.

both.[12] The average starting price of an auction was 152 Euro and the average fixed price offer was 185 Euro. 33% of the offers were sold (17% as an auction and 16% at a fixed price). While 52% of the services offered exclusively as an auction were sold, only 30% of the fixed price offers and 23% of services offered in both ways were sold. The average sold price was 186 Euro (210 Euro for auctions, 168 Euro for fixed price offers and 176 if both was offered), which corresponds to an hourly wage of 103 Euro (93 Euro, 114 Euro and 100 Euro, respectively).[13] These hourly wages are quite high compared to the median hourly wage in Germany of 14 Euro (Brenke & Grabka, 2011), but do not account for standby time and expenses (estimated to be about 50% of transaction cost; Nagel, 2013). Sold offers were on average online for 5.2 days (7.1, 3.2 and 5.4 days, respectively) and sold auctions had an average of 8.6 bids. In the sample period $3,444$ women were active, posting on average 30.0 offers and selling 9.7 services. The $14,093$ men in our sample made on average 7.2 bids and bought 2.4 services.

Depending on which estimate of the whole population of sex workers and their clients in Germany is used (see Nagel, 2013), this corresponds to $0.2-0.5\%$ of female sex workers in Germany and $0.13-0.14\%$ of male clients. When offering a service, sellers specify the first two digits of the postal area they are offering their service in.[14] In each of these 95 two digit zip codes on average $1,074.7$ services were offered of which on average 349.8 were sold. Together with the seller and buyer nicknames, the regional identifier makes this dataset very unique. With the two digit zip-code it is possible to exploit geographical information and the nicknames allow to look at repeated observations of both the buyers and the sellers. In addition, sellers and buyers give each other a rating, which also could be automatically extracted from the data. The average rating of a seller (by her former customers) is 98.3% and of a buyer 96.7%. If a service is falsely described or a seller / buyer does not fulfill the contract, the counter party is rated rather low. Hence, misunderstandings between sellers and buyers or falsely described services seem to occur seldom.

Information in the free texts

In addition to the obligatory structured information, sellers can give more information in the title, subtitle and the free text field. Figure 3.1 shows a fairly representative example of such a description. In this part of the offer, sellers describe themselves, the kind of service they offer (and do not offer), and sometimes impose restrictions on who may bid (e.g., minimum age). Text length ranges from 2 to $1,614$ words with an average of 190 words of which on average 3% are misspelled.

[12] If an offer is made both as a buy now and an auction article, in the beginning it can be bought directly at the fixed price. As soon as one person bids on the auction, it cannot be bought at a fixed price anymore.

[13] In the free texts, sellers provide the duration of the offered service, which we used to calculate hourly wages.

[14] Germany uses a 5 digit zip code system to identify postal areas.

Table 3.2: *Descriptive statistics - Free texts*

Variable name	**Type**	**Mean**	**SD**
Seller specifies age	binary	0.64	0.48
Age in years if stated	metric	31.87	8.87
Seller specifies height	binary	0.53	0.50
Height in cm if stated	metric	166.85	6.19
Seller specifies weight	binary	0.31	0.47
Weight in kg if stated	metric	65.15	20.55
Seller specifies hair color	binary	0.84	0.37
Hair color blond if stated	binary	0.44	0.50
Hair color brown if stated	binary	0.20	0.40
Hair color black if stated	binary	0.32	0.47
Hair color red if stated	binary	0.04	0.20
Seller specifies hair length	binary	0.25	0.43
Hair long if stated	binary	0.89	0.31
Seller claims to have tattoo	binary	0.04	0.19
Seller claims to have piercing	binary	0.07	0.26
Seller claims to be intimately shaved	binary	0.24	0.43
Seller specifies cupsize	binary	0.88	0.33
Cupsize medium if stated	binary	0.63	0.48
Cupsize large if stated	binary	0.28	0.45
Seller claims to be in relationship	binary	0.05	0.22
Seller claims to be single	binary	0.05	0.21
Seller claims to be higher educated	binary	0.04	0.19
Seller claims to be no professional sex worker	binary	0.27	0.44
Duration of the service in hours	metric	2.87	7.75
Seller does not kiss	binary	0.03	0.29
Seller offers massage only	binary	0.01	0.10
Seller offers manual stimulation only	binary	0.02	0.15
Seller offers oral stimulation only	binary	0.10	0.30
Seller specifies something about vaginal sex	binary	0.94	0.24
Seller offers vaginal sex, no anal sex	binary	0.72	0.45
Seller specifies something about anal sex	binary	0.52	0.50
Seller offers vaginal and anal sex	binary	0.16	0.37
Seller specifies something about condom use	binary	0.78	0.41
Seller offers sex without condom	binary	0.04	0.20
Seller specifies something about the location	binary	0.88	0.32
Seller offers service at hotel	binary	0.48	0.50
Seller offers service at seller's home	binary	0.30	0.46
Seller offers service at buyer's home	binary	0.54	0.50
Seller offers service outdoors	binary	0.06	0.24
Seller offers service in car	binary	0.10	0.29
Seller offers to go for dinner / drink	binary	0.11	0.31
Seller claims to offer for the first time	binary	0.03	0.16
Seller imposes minimum age on potential buyers	binary	0.13	0.34

Notes: The table presents the descriptive statistics of all variables generated from the free text fields. The underlying sample consists of all 102, 096 offered services.

To extract information from these texts, we first wrote a script that compares the texts of the 102,096 offers to each other and looks for similarities finding identical or nearly identical free texts. This reduces the number of free texts to analyze to roughly 35,000 texts. Again using pattern recognition the script looks for keywords and automatically finds out the information we are interested in. As the meaning of a keyword as well as more complex pattern recognition sometimes is ambiguous, depending on the context, the results had to be checked manually. From January 2012 to October 2012, six students looked at the texts as well as the extracted information and corrected them where necessary. The procedure was done a second time by a different student to make sure that no mistakes remained. Table 3.2 presents a list of all variables generated from free texts.

Regarding the personal details posted in the free texts, we find that 64% of the sellers state their age, with an average of 31.9 years. The reported average height and weight is 167cm and 65kg (stated by 53% and 31% of the sellers respectively). 84% of the sellers state their hair color (44% blond, 20% brown, 32% black, 4% red) and 25% their hair length (89% long). 4% of these sellers claim to have a tattoo, 7% to have a piercing and 24% to be intimately shaved. Of the 88% sellers having stated their cup size 9% claim to have small breast (A, AA), 63% to have medium sized breast (B,C) and 28% to have large breast (D and larger). 5% of the sellers claim to be in a relationship, 5% to be single and 3.9% to be university level students or to have higher education.

In the free texts also the service offered itself is being described. The average duration of an offered session is 2.9 hours. Interestingly the average sold auction includes 3.1 hours of sexual services, while the average fixed price offer only includes 2.1 hours of services. 3% of the sellers decline to kiss, 0.01% of the offers only include a massage, 2% only manual stimulation, 10% only oral stimulation, 68% offer vaginal sex, but no anal sex, and 16% offer both vaginal and anal sex. The remaining 4.7% do not disclose which services are included and which not. Nevertheless, the platform explicitly sells sexual services and most of these services' free texts can be interpreted such that the customers can actually assume that vaginal sex is also offered in these unspecified cases. For condom use, from reading the free texts, one gets the impression that the standard is protected sex. If nothing is disclosed in the text (in 22% of the offers), customers must assume that they might have to use a condom. 4.3% of the sellers explicitly offer unprotected sex. The rate is quite low, but a bit higher among sold services (6.4%).[15] This rate might be even higher as customers could pay an extra fee to get additional services when meeting the sex worker.

Concerning the meeting place, 88.0% of the offers specify possible locations: 48% at a hotel, 30% at the buyers residence, 54% at the sellers residence, 6% outdoors

[15]For all other variables in this paragraph, the numbers for sold services are nearly the same as for the offers.

(parks, etc.) and 10% in a car (sometimes even while driving). Note that offering multiple locations is possible. Additionally, 11% of the offers include that the seller will go to a bar or restaurant with the buyer (which the customer has to pay for). When deciding where to offer a service, there are two opposing arguments. On the one hand, a hotel or the seller's own residency is safer for the seller, as she can control the environment. On the other hand, hotel cost drive up the transaction's complete price and the sellers' neighbors might start to wonder if every other day a different man drops by. This renders the buyers home or an outdoor place an alternative.

Finally, 3% of the sellers claim to offer for the first time, 13% specify a minimum age for the customers of 30 to 45 years and 27% claim to offer sexual services only privately, i.e., they claim to be no professional sex worker. This last feature is interesting, because in other studies this part of the sex work business is not represented at all. These sellers provide on average longer free texts, but their personal attributes do not differ that much. However, there are some differences. 14% of the non-professionals claim to be students, but none of the other sex workers. They earn the same hourly wages, but offer slightly different services. Compared to the sub-sample of other sex workers, the sub-sample of non-professionals offers longer sessions (3.1 vs. 2.7 hours) and rather vaginal sex, than other non-intercourse services. Furthermore, they offer less often unprotected sex (3.8% vs. 4.6% or for sold services 5.7% vs. 6.7%) as well as anal sex (14.9% vs. 16.9%). Customers seem not to favor one or the other group. Only four percent of the men who bought more than one service only bought these from non-professional sex workers.[16]

Picture information

The third and final part of the information is embodied in the up to three pictures sellers can post with their offerings. Again we wrote a script to find duplicates, which reduces the number of pictures necessary to examine from nearly 200,000 to 46,000. The same six students then looked at the unique pictures to extract information (see table 3.3). To correct mistakes every picture was examined by a different student for a second time.

94% of the offers included at least one, on average 1.9 pictures. Most sellers post a picture of their full body (67%), and only wearing underwear (64%). While 27% of the offers include a picture of the naked breasts, only 8% (9% if sold) include a picture of the seller completely naked (in 1.4% of the offers even the vagina can be seen). In 17% of the offers most parts of the face and on 25% the whole face can be seen. The last variable the students extracted was on the sellers weight, which we could then use to update weight information of sellers that did not specify their weight in the free texts. Of the 96,293 offers with a posted picture, 1.2% included an

[16] For more on the differences between professionals and non-professionals see section A.2.1 in the appendix.

Table 3.3: *Descriptive statistics - Picture information*

Variable name	Type	Mean	SD
Seller posts one or more pictures	binary	0.94	0.23
Number of pictures posted	metric	1.93	0.95
Seller posts picture of full body	binary	0.67	0.47
Seller posts picture in underwear	binary	0.64	0.48
Seller posts picture completely naked	binary	0.08	0.27
Seller posts picture with naked breasts	binary	0.27	0.44
Seller posts picture of vagina	binary	0.01	0.12
Seller posts picture showing her face	binary	0.25	0.43
Seller posts picture showing half her face	binary	0.17	0.37
Seller posts picture and appears to be underweight	binary	0.01	0.11
Seller posts picture and appears to be of normal weight	binary	0.73	0.45
Seller posts picture and appears to be slightly overweight	binary	0.21	0.41
Seller posts picture and appears to be obese	binary	0.07	0.24

Notes: The table presents the descriptive statistics of all variables extracted from pictures posted in an offer. 94% of the 102, 096 offers included at least one picture and on average an offer included 1.9 pictures.

underweight seller ($BMI < 18.5$), 73% a normal weight seller ($18.5 < BMI \leq 25$), 21% a slight overweight seller ($25 < BMI \leq 30$) and 7% an obese seller ($BMI > 30$). As attributes such as good looks are subject to the viewer, we did not code these.

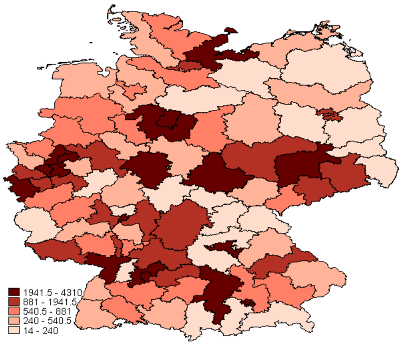
To sum up, the dataset we created has some unique features as it is based on actual contracts and for every contract includes a regional, seller and buyer identifier, but it also has some disadvantages. Much information were only posted in a free text and categorized by students. In some cases sellers did not post all necessary information resulting in missing values. These gaps could partly be filled with information from pictures (e.g., on the sellers weight in the pictures), but in other cases information remains missing. Although this process was done twice, there might remain some mistakes.

3.5 The geography of sex work in Germany

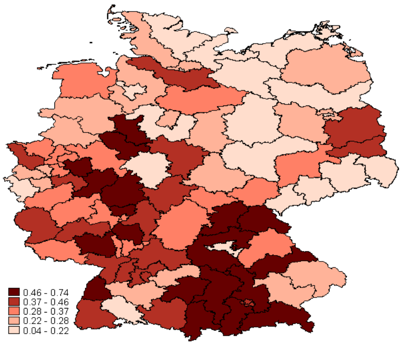
One valuable feature for research purposes of *www.gesext.de* is that sellers have to specify the first two digits of the postal area in which they intend to offer their service. This allows us to look at each area independently as well as to compare the areas to each other. Furthermore, we can merge our dataset with regional disaggregated data. One important dataset to mention here is the Genesis-online data base by Destatis (the German national statistical bureau, Destatis, 2014) offering regional statistics on the level of local administrative units (LAU-2). These can be aggregated to the level of 2 digit zip code areas and include amongst others the socio-demographic information of the population living in these regions. In this section we look at regional differences of some selected variables concerning the market

Figure 3.2: *Maps: Market participation and contract terms*
Sex work in Germany at the 2 digit zip code level. Darker colors represent higher values of the variable in the given region. *Source: Own representation.*

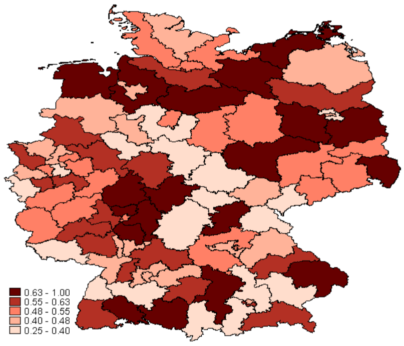
(a) Total number of offers



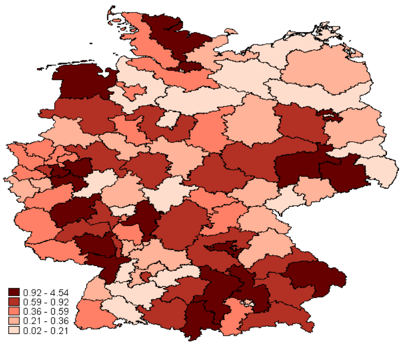
(b) Sold services relative to offers



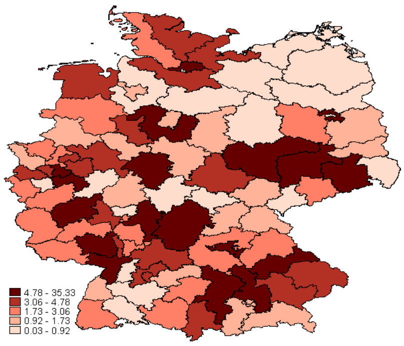
(c) Auctions relative to all offers



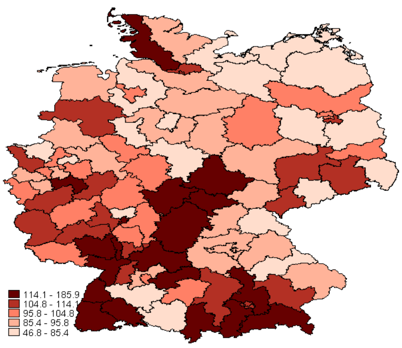
(d) Number of sellers per 1000 women



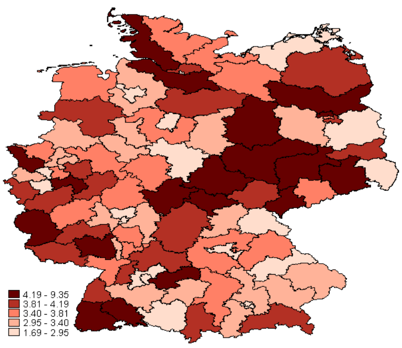
(e) Number of buyers per 1000 men



(f) Average hourly wages for vaginal sex



(g) Average hourly wages for vaginal sex relative to GDP per capita.



of sex work, and start with market participation.

Figure 3.2a shows the absolute number of observations (sexual service offers) in our dataset. With an average of $1,075$ offers per region, the lowest number of offers is found in D-02 Bautzen (14) and the highest in D-40 Duesseldorf (4310). In general, the areas that are known to have a large market for sex workers (e.g., see Eigendorf et al., 2013), are also in our dataset represented with a high number of offers (e.g., Berlin, Dresden, Frankfurt, Hanover, Munich, Nuremberg, Stuttgart and the Ruhr area). Not surprisingly, rural and less populated areas tend to have a lower absolute number of offers. Geographically – with some exceptions – in Southern and Western Germany more than 30% of these offers are sold, while in Northern and Eastern Germany less than 30% are actually sold (Figure 3.2b). Especially in Bavaria supply seems to match demand or vice versa.

In Figures 3.2d and 3.2e, the number of sellers and buyers per 1000 female and male inhabitants respectively is presented. Again the most known areas also have the highest relative number of sellers (e.g., Berlin, Dresden, Leipzig, Nuremberg, Ruhr area), but in some areas it is actually the neighboring zip codes that have the above average number of sellers (e.g., the areas around Frankfurt and Munich). Concerning the relative number of buyers the picture is roughly the same and unsurprising.[17] Large cities and their direct surroundings have the highest relative number of bidders.

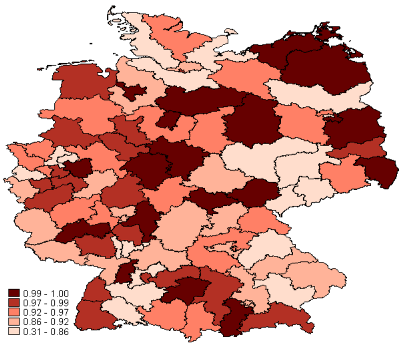
Figure 3.2c shows that the use of auctions to sell sexual services is quite heterogeneous among regions. On average, half of the offers are made as auctions and half of them at a fixed price (see above). In contrast, there is a clear pattern for average hourly wages (Figure 3.2f). Hourly wages range from 47 Euro (D-02 Bautzen) to 198 Euro (D-98 Suhl). With the exception of D-20 Elmshorn (near Hamburg), the most expensive regions are found in the South and the West of Germany. Interestingly, prices are higher in the regions surrounding Munich than in Munich itself. The geographical pattern changes if the hourly wage is corrected for the regional GDP (Figure 3.2g). On the one hand, some regions even stay relatively expensive compared to living cost (e.g., Cologne, Freiburg, etc.), but – again with some exceptions – sexual services in East German regions are relatively expensive compared to the local GDP. We study these regional wage differences in more detail in chapter 6.

Another dimension are contract and seller attributes. Figure 3.3 shows the share of offers that contain certain types of sexual services. As in the whole sample, 92% of the offers include vaginal sex (see above) meaning that in most areas the share of offers including sex is very high. In only 10 regions it is below 75%. The total share of sellers offering anal sex is 16%, but in a fifth of the regions higher than 27%. The overall rate of services offered without protection (condom use) is quite low in

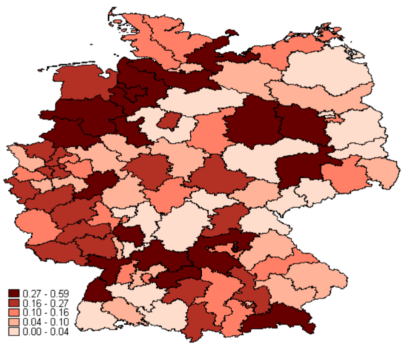
[17]Note that we do not know where the buyer lives, but only where the seller offers the service he bids on. Hence, the relative number of buyers is actually the number of buyer bidding in a certain area, relative to the men actually living there.

Figure 3.3: *Maps: Contract and seller attributes*
Sex work in Germany at the 2 digit zip code level. Darker colors represent higher values of the variable in the given region. *Source: Own representation.*

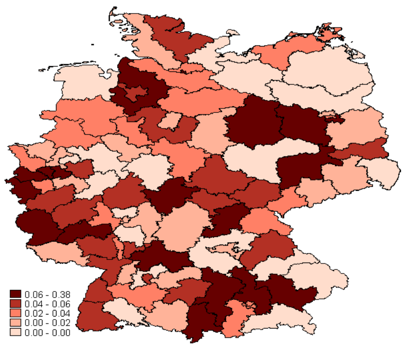
(a) Share of sellers offering vaginal sex



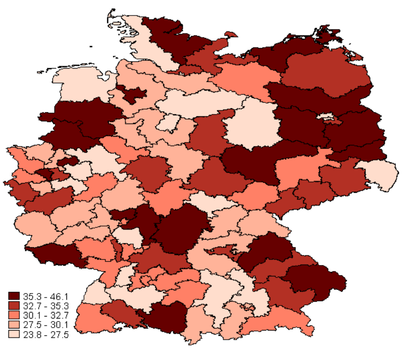
(b) Share of sellers offering anal sex



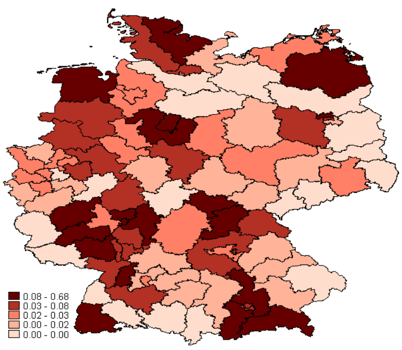
(c) Share of sellers offering unprotected sex



(d) Average age of a seller



(e) Students relative to all sellers



the sample (4%, see above), but in a fifth of the regions it actually exceeds 6%. In D-20 Hamburg it is 20% and in D-39 Magdeburg it is even 38%. In the 18 months, for which data is available, sexual services were offered without protection 108 and 198 times in these two regions, respectively. Figure 3.3d shows that the sellers' age is very heterogeneously distributed. There is a slight tendency of a higher average age in the East German regions (e.g., 46 years in Potsdam), but also in some West German regions the average age is relatively high (e.g., 40 years in Bremen). One last attribute that is of interest in a geographical sense is the share of sellers that claim to be students (Figure 3.3e). Interestingly, this share is not systematically higher in University cities, in their surroundings or in regions with more young women. For example, in Berlin, Frankfurt, Freiburg, Hannover, Mainz and Munich, the share of students is higher than on average, but other famous university towns do not stand out at all. To sum up, concerning sex work, there is much heterogeneity between the 95 regions in Germany.

3.6 Conclusions

In this chapter we have given a short overview over the legal status of sex work in Germany and existing theories and datasets. In the presentation of the descriptive statistics, it has become clear, that the *"www.gesext.de"* dataset is very unique. It is the only dataset based on actual contract data in a developed country not created by surveying sex workers, but by extracting data directly from a match-maker platform. According to the information contained in the dataset, sex work in Germany is heterogeneous between sex workers and regions. In the next chapters we investigate different aspects of this otherwise hidden market.

Chapter 4

Endogenous risk taking and physical appearance of sex workers

4.1 Introduction

While some countries ban the supply (e.g., Ireland, United States) or the demand (e.g., Finland, France, Ireland, Norway, Sweden, United States) of sexual services for pay, such contracts are legal elsewhere (e.g., Austria, Belgium, Brasil, Germany, Hungary, Netherlands, Switzerland, United Kingdom; see Cho et al., 2012). Independent of legality, there is abundant evidence of existing sex work – mainly supplied by female workers to male customers – all over the world. Over the last decade supply has shifted towards the internet. Street prostitution and brothels still exist and make up a large part of the market, but with the anonymity of the internet, platforms mediating between providers and customers of sexual services offer a cheap way to advertise sexual services and have become more and more important (Cunningham & Kendall, 2011b). It is hard to get reliable data for this market (for the "online" and, especially, the "offline" segment), but estimates for the whole market speak of up to 400,000 women working full or part-time in the sector, earning per year up to 14.5 billion Euro or 0.4% of the GDP in Germany alone (Reichel & Topper, 2003).

A key issue in professional sex work is health protection for all involved parties. Recent research provides evidence that a significant fraction of sexual services are provided without protection (e.g., Rao et al., 2003; Gertler et al., 2005; Cunningham & Kendall, 2010b; de la Torre et al., 2010; Robinson et al., 2011; Arunachalam & Shah, 2013), thereby encountering a risk of genital and other diseases such as HIV.[1] Most of the literature acknowledges (directly or indirectly) that risk-taking through the offering of unprotected sex is endogenous. Sex workers had been found to be willing to offer sex without condom more likely, if a price premium was paid (see Gertler et al., 2005; de la Torre et al., 2010) or the opportunity costs were high.

[1] Throughout this chapter, we use the terms "risk-taking", "unprotected sex", and "sex without condom" interchangeably.

E.g., de la Torre et al. (2010), Robinson et al. (2011), and Wilson (2012) showed that sex workers offered such services more frequently after spells of bad financial shape, negative income shocks, or illness (and an associated loss of income). This evidence is consistent with the larger prevalence of unprotected sex offerings in the "offline" (e.g., street prostitution) relative to the "online" (internet) market, since suppliers in the former segment of the market tend to be less educated, to charge lower prices, and to depend more on the income generated from their services than ones in the latter segment (see Cunningham & Kendall, 2011b).

The average price premia reported in the aforementioned work range from about 9% for Kenyan sex workers in Robinson et al. (2011) and Ecuadorian sex workers in Arunachalam & Shah (2013) over 23% in Gertler et al. (2005) to 30% in de la Torre et al. (2010) for Mexican sex workers to 194-376% for Indian sex workers in Rao et al. (2003). It is worth mentioning that, among those studies, the one by Rao et al. (2003) was the only one to be able to randomize their subject pool about condom usage, and it found the largest price premium among all of them.

Some recent research on the size and the determinants of the price premium for unprotected sex argues that the opportunity costs of denying such services (i.e., the economic incentives of supplying them) are larger for less attractive sex workers. Gertler et al. (2005) provide a theoretical rationale for this line of work. Their model suggests that attractiveness raises the bargaining power of sex workers, whereby attractive sex workers are able to charge a higher price for sex services in general and also a higher premium for unprotected sex than less attractive ones.[2] This shows in higher price premia for unprotected sex offerings of more attractive sex workers (see Gertler et al., 2005) and in lower price premia for obese sex workers (see Chang & Weng, 2012). However, the associated evidence assumes that sellers' choice to offer unprotected sex practice is random.[3]

This chapter utilizes data on the offering and contracting of sexual services via the internet through *www.gesext.de*.[4] We downloaded all contracted services of female suppliers between January 16 of 2011 and September 9 of 2012 from this database. In this study, we focus on all contracts where the weight and offered services of the worker are known and a contract has been concluded at a known price.[5]

[2] To some extent, the arguments in this line of research are consistent with the evidence that economic success and physical appearance are related to each other (see Hamermesh & Biddle, 1994; Fletcher, 2009; Mobius & Rosenblat, 2006; Johnston (2010).

[3] Gertler et al. (2005) demonstrate in their Table 5 that the practice of unprotected sex can be explained by sex worker and client characteristics, but their price comparisons of protected and unprotected sex practices only account for endogenous selection based on time-invariant sex worker characteristics. (Chang & Weng, 2012, p.482) acknowledge that their "analysis may suffer from endogeneity bias because condom use and prostitute price may be correlated due to some unobserved common factors."

[4] While this is only *one* such platform in Germany, it accounts for an annual revenue of about 4 million Euros for sexual services offered by females only.

[5] When rescinding a concluded contract, the contracted price is still due. This can be enforced since customers must fully reveal their identity to the platform owner. 15% of the contracted price is generally due as a fee for the platform services.

We estimate the price premia in contracts about unprotected sex, emphasizing the importance of controlling for endogeneity. We propose an identification strategy which relies on two alternative types of instruments – one related to the frequency of prior unprotected sex offerings by other sex workers in the same region where a given contract was made and one related to personal characteristics of other sex workers offering prior unprotected sex in the same region where a given contract was made. In line with earlier work, we find that offering unprotected sexual services raises hourly wages of sex workers. Consistent with the findings of Rao et al. (2003), we find that non-random selection (of sex workers and clients) into unprotected sex leads to a downward bias of the risk premium. Considering self-selection into risk-taking suggests that the risk premium on non-protection is more than twice as large as when assuming random risk-taking. We estimate the risk premium at about 91% of the average hourly wage. However, it is interesting to see that the magnitude of the premium is smaller than in Rao et al. (2003), since we analyse a sample of sex workers from a developed country and in the "online" segment, where at least the purely economic opportunity costs should be relatively high. Apart from the mentioned ones, reasons for the quantitative difference between the results in this chapter and the one of Rao et al. (2003) may be that medication to treat HIV is more easily available and affordable in Germany than in India and that the time periods are rather different between the two studies (pertaining to 1993 in Rao et al., 2003, and to 2011-2012 in this study). Interestingly and in contrast to the results in Gertler et al. (2005) and Chang & Weng (2012), the results in this chapter suggest that appearance affects the hourly wage but not the risk premium on unprotected sex.

The remainder of the chapter is organized as follows. The next section discusses the novel database on online sex contracts used in this chapter. Section 4.3 describes the identification strategy and the econometric framework. Section 4.4 summarizes the estimation results, and the last section concludes.

4.2 Data

Before introducing the data-set employed in detail, we should emphasize two features. First, as any other study on sex workers, data on this type of work are generally selected in the sense that census-type data on sex workers are not available. Data will not be available, unless sex workers are willing to participate in a field study (as in Rao et al., 2003; Gertler et al., 2005; de la Torre et al., 2010; Robinson et al., 2011; Chang & Weng, 2012; Arunachalam & Shah, 2013) or they come forward with an online posting and deliver information without knowing (as in this study and also in Moffatt & Peters, 2004; Cunningham & Kendall, 2011b). Second, as opposed to "offline" sex workers, "online" sex workers as used in this study tend to be better educated, financially less dependent, and part of the higher-price segment of the market (see Cunningham & Kendall, 2010b).

We employ digitally collected data on the offering and contracting of sexual services by females via the internet through *www.gesext.de*. When making an offer on this platform sellers describe themselves and their services (including, e.g., notes on condom use, duration of the meeting, the location, ...) and can offer sexual services either as a second price auction or at a fixed price. While in the first case the highest bidder has to pay the second highest bid, in the latter case the first bidder has to pay the fixed price. The meeting then takes place at a predefined place (hotel, flat, ...) and the payment is normally made in cash. It is unknown if the seller and buyer actually meet and engage in sexual services, but when rescinding a concluded contract, for the customer the full contracted price is due. For the seller 15% of the contracted price is generally due as a fee for the platform services. This can be enforced since customers must fully reveal their identity to the platform owner.

Table 4.1: *Descriptive statistics*

Variable	**Mean**	**SD**	**Binary**
Log price per hour	4.9970	0.7698	no
Risk-taking	0.0721	0.2587	yes
Overweight	0.2373	0.4254	yes
Height is provided	0.5815	0.4933	yes
Height in cm if height is provided	166.6254	6.0342	no
Height-squared in cm if height is provided	27,800.4400	1,994.4560	no
Age is provided	0.4098	0.4918	yes
Age in years if age is provided	29.7913	8.6382	no
Age-squared in years if age is provided	962.1291	583.7763	no
First-time supplier	0.0522	0.2224	yes
Single	0.0828	0.2756	yes

Notes: The table presents the descriptive statistics of the variables used in this chapter. The total number of observations (transactions) is 16,583 and the underlying number of sex workers is 2,517.

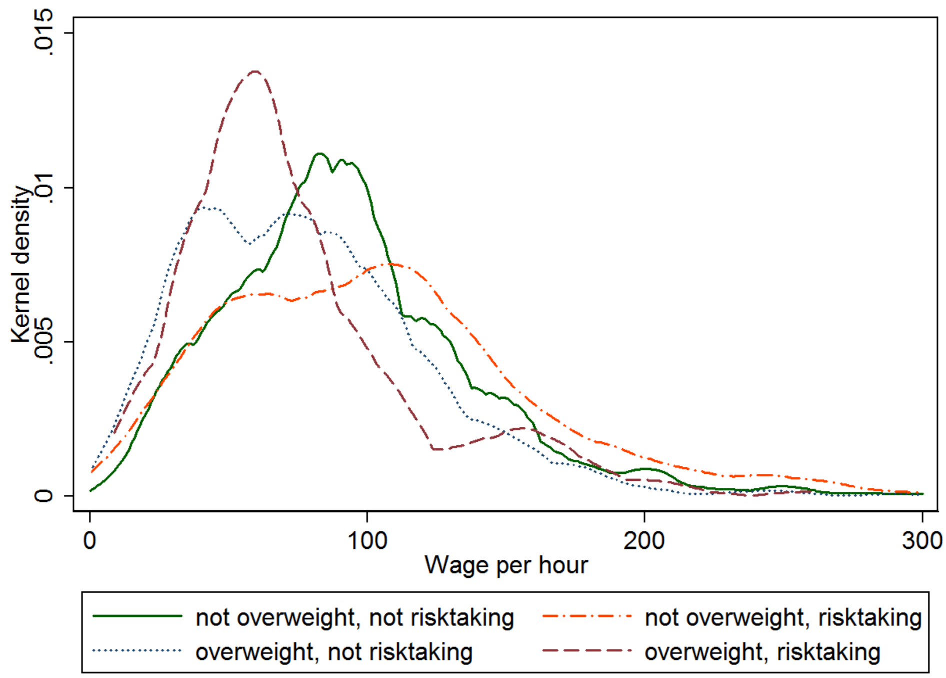
In general everybody can offer services on this platform, but the average price of about 148 Euro shows that this platform mainly covers a premium segment of the market.[6] We downloaded all contracted services of female suppliers between January 16 of 2011 and September 9 of 2012 from this database and focus on all offers where a contract has been concluded at a known price.

Table 4.1 and Figure 4.1 provide descriptive evidence of the hourly wage of 2,517 sex workers offering sex for pay at *www.gesext.de* between January 17 of 2012 and September 9 of 2012 through altogether 16,583 individual contracts. We classify workers with a body mass index of more than 25 as overweight. Some of the workers do not provide weight and height but only pictures. For those, we classify them

[6] Döring (2014) writes that the average price for an one hour session with a indoor sex worker in Germany is about 100 Euro.

Figure 4.1: *Kernel densities*

The figure presents the kernel densities for hourly wages of overweight/non-overweight by risk-taking/non-risk-taking female sex workers in Germany. *Source: own representation.*



as being overweight or not, depending on their physical appearance (this was done by six students between March 2012 and September 2012).[7] For all of the transactions we know whether workers offer unprotected services (sex without a condom) or not. Of the 16,583 contracts, 3,935 involve overweight sexual workers and 12,648 do not, and 1,196 involve unprotected service offerings while 15,387 do not. Interestingly, about 5.5% of the contracts offered by overweight workers involve unprotected service, while 7.7% are unprotected for non-overweight workers. Table 4.1 provides further information on characteristics of sex workers and on the contract environment in the database. For instance, about 58% of the concluded contracts involve workers who provide information about their height and, on average, those are about 167 centimeters tall. In about 41% of the contracts sex workers mention their age and, on average, those are about 30 years old. About 5% of the contracts involve workers who indicate that they are unexperienced (they are effectively first-time suppliers on that platform), and in 8% of the contracts workers indicate that they

[7]Of all 16,583 individual contracts covered, weight was imputed by students into the four categories "obese" (corresponding to a body-mass index, BMI, of higher than 30) an estimate BMI, "overweight" (corresponding to a BMI of higher than 25 up to 30), "normal weight" (corresponding to a BMI of higher than 18.5 and up to 25), "underweight" (corresponding to a BMI of up to 18.5) for 11,630 transactions. In order to avoid a measurement error to the largest possible extent, this categorization was done twice, by different students. Notice that subjective classification schemes of sex workers' appearances had also been used, e.g., by (Moffatt & Peters, 2004), (Gertler et al., 2005), (Cunningham & Kendall, 2011b).

are single. The average hourly wage across all contracts is about 148 Euros (about 5.00 in logs) with a standard deviation of about 2.16 Euros (0.77 in logs).[8]

Towards accounting for the endogeneity of risk-taking (contracting unprotected sexual services without condom), we propose two types of instruments both of which measure features of contracts that had been concluded by sex workers other than the one involved in and happening prior to transaction i but in the same two-digit district: one type of instrument is based on the prior average risk-taking of other sex workers, and one is based on the average height of other sex workers. To form two identifying instrumental variables, for either type of instruments, we use the value of the instrument, which is i-specific, as well as its squared value. The instruments are likely exogenous as long there is some inter-temporal habit formation, i.e., if sex workers adopt strategies of others in the same neighborhood in the past. The second type of instrument (based on others' height) is relevant and adequate when two things hold: first there is habit formation over time and height matters for risk-taking.

Figure 4.1 displays the distributions of log hourly wages for the four cells in the matrix overweight/non-overweight by risk-taking/non-risk-taking. The figure illustrates that the average hourly wages of non-overweight risk-taking suppliers is higher than that of non-risk-taking suppliers. Especially, the right tail of the former distribution is fatter than that of the latter distribution. These features are somewhat less clear-cut for overweight suppliers who clearly make up a smaller part of the overall distribution. However, we should bear in mind that the distributions are unconditional on explanatory variables and do not provide for direct inference about the relative importance of risk-taking on hourly wages of female sex workers.

4.3 Estimation methods

One particular goal of this section is to shed light on the differences between regressions which treat $Risk-taking_i$ as exogenous – i.e., assuming that sex workers do not self-select into unprotected sex – and ones which consider $Risk-taking_i$ to be endogenous upon self-selection. For the latter, we will generally pursue an instrumental-variable (IV) two-stage least-squares (2SLS) strategy which takes the binary nature of $Risk-taking_i$ into account. Procedures for such models had been suggested, among others, by Heckman (1978), Maddala (1983), Heckman & Vytlacil (1999), Vella & Verbeek (1999), and Wooldridge (2002). Specifically, Wooldridge (2002, pp. 621-633)) discusses the assumptions and suggests four alternative para-

[8]Clearly, all of the information contained in the online postings is self-reported. Hence, sex workers might strategically misreport (see Plankey et al., 1997; and Cawley, 2014; for a general discussion of misreporting). However, the contracts are concluded based on this information provided and they become legally effective. Moreover, most sex workers are rated by earlier customers, and misreporting would show in the ratings. The ratings do not indicate strategic misreporting on average.

metric procedures based on those (on pages 623, 626, 629f., and 631, respectively). We will rely on such approaches, where the average treatment effect – in our case, of $Risk - taking_i$ – may vary with some covariates – here, $Overweight_i$.

For the approaches considered here, we will maintain the following set of assumptions. (*Instrument exogeneity and Conditional mean independence.*) First, we will assume that at least one instrument $Instruments_i$ exists which can be used to model the probability that $Risk - taking_i = 1$,
$P(Risk - taking_i = 1) = \Phi(Overweight_i, Controls_i, Instruments_i) = \Phi_i$, where the instrument is uncorrelated with the unobservable determinants of – i.e., the stochastic term on – the dependent variable of interest, e.g., $\ln(HourlyWage)_i$. Denoting this stochastic term by ε_i, the formal requirement is $E(Instruments_i \varepsilon_i) = 0$ and assuming that Φ_i can be estimated, the prediction $\hat{\Phi}_i$ is a natural instrument for $Risk - taking_i$ (see Wooldridge (2002, p. 626)). Conditional on $Overweight_i$, a set of control variables entering the outcome equation, $Controls_i$, which include a constant, and the set of instruments, $Instruments_i$, $Risk - taking_i$ is then random. Formally, when using superscript 1 for $Risk - taking_i = 1$ and superscript 0 for $Risk - taking_i = 0$ (only one of which is observed for transaction i), the expectations $E(\varepsilon_i^s | Overweight_i, Controls_i, Instruments_i) = E(\varepsilon_i^s | Overweight_i, Controls_i)$ for $s \in \{0, 1\}$. (*Joint normality of the disturbances on $Risk - taking_i$ and outcome such as $\ln(HourlyWage)_i$.*) Under normality of the latent variable determining the desirability of $Risk - taking_i = 1$, the function Φ_i is the cumulative normal distribution function and $\hat{\Phi}_i$ is the prediction that $P(Risk - taking_i = 1)$. Under bivariate normality of the random component of $\ln(HourlyWage)_i$ and the random component in Φ_i, say, ν_i, the prediction of a probit model of $P(Risk - taking_i = 1)$ can be used as an instrument for $Risk - taking_i = 1$ (see Wooldridge (2002, p. 623), Procedure 18.1). Alternatively, one can use a control function using the probability density function, $\phi(Overweight_i, Controls_i, Instruments_i) = \phi_i$, based on the latent process for $P(Risk - taking_i = 1)$ in addition to instrumenting (see Wooldridge (2002, p. 629f), Procedure 18.3).

The same assumptions support models where the effect of $Risk - taking_i$ on outcome, e.g., $\ln(HourlyWage)_i$, varies with some or all observable elements in $(Overweight_i, Controls_i)$. In light of earlier work on the matter (see Chang & Weng, 2012), we will consider cases where the average treatment effect of $Risk - taking_i$ varies with $Overweight_i$ only. Suppose that the effect of $Risk - taking_i$ of interest on outcome is $\beta Risk - taking_i + \gamma Risk - taking_i \times Overweight_i$ (when omitting other effects for brevity). In that case, the average treatment effect of $Risk - taking_i$ is $\beta Risk - taking_i + \gamma \overline{Overweight}$, where $\overline{Overweight} = E(Overweight_i)$ can be estimated by the sample mean. Wooldridge suggests demeaning $Overweight_i$ in the interaction term and using $\widetilde{Overweight}_i \equiv Overweight_i - \overline{Overweight}$ in $\beta Risk - taking_i + \gamma Risk - taking_i \times \widetilde{Overweight}_i$, so that β measures the average treatment effect of $Risk - taking_i$, since $E(\widetilde{Overweight}_i) = 0$. Then $\hat{\Phi}(\cdot)$ and $\hat{\Phi}(\cdot) \times \widetilde{Overweight}_i$ are the suitable instruments for $Risk - taking_i$ and $Risk - taking_i \times \widetilde{Overweight}_i$, respectively (see Wooldridge (2002, pp. 626 and 629)).

Consequently, we run regressions per transaction $i = 1; ...; 16,583$ of the form

$$\begin{aligned}
\ln(HourlyWage)_i &= \beta Risk - taking_i + \gamma Risk - taking_i \times \widetilde{Overweight}_i \\
&+ \delta Overweight_i + \zeta Controls_i + (\eta\hat{\phi}_i) + \varepsilon_i. \qquad (4.1) \\
\overline{Overweight} &\equiv \frac{1}{16,583} \sum_{i=1}^{16,583} Overweight_i, \\
\widetilde{Overweight}_i &\equiv Overweight_i - \overline{Overweight}. \qquad (4.2)
\end{aligned}$$

Notice that the interaction term $Risk-taking_i \times \widetilde{Overweight}_i$ involves the demeaned value while the main effect involves the un-demeaned value of $Overweight_i$ as in Wooldridge (2002, pp. 626 and 629). The term $\eta\hat{\phi}_i$ is the control function based on the estimated density $\hat{\phi}_i$, which will not be included in all models. Models that exclude $\eta\hat{\phi}_i$ correspond to what Wooldridge (2002, p. 626) calls *Procedure 18.2*, whereas ones that include $\eta\hat{\phi}_i$ correspond to what Wooldridge (2002, p. 629) calls *Procedure 18.3*.

We run three pairs of versions of (4.1), three of them excluding $\gamma Risk-taking_i \times \widetilde{Overweight}_i$ and the other ones including it. Two models assume random assignment of sex workers into risk-taking whereby β could be estimated through ordinary least squares (OLS) on (4.1). Four models assume self-selection into risk-taking and aim at avoiding an associated bias by using instrumental variables in two-stage least squares (2SLS) regressions as described above. For the latter, we estimate a probability model of the form

$$P(Risk - taking_i = 1) = \Phi([Instruments_i, Overweight_i, Controls_i]\theta + \nu_i), \quad (4.3)$$

where we use two variants of $Instruments_i$ as described in Section 4.2. In the main part of the chapter, we employ $Instruments_i = (RegionHistory_i, RegionHistory_i^2)$, where $RegionHistory_i$ is the average value (probability) of risk-taking by other sex workers in the same two-digit zip code between January 16 of 2012 and September 8 of 2012 and at least one day prior to contract i and $RegionHistory_i^2$ is the squared value thereof. The vector θ is a conformable parameter vector on all the explanatory variables in the probit model (including fixed effects for 78 two-digit zip codes). $RegionHistory_i$ and $RegionHistory_i^2$ are valid instruments, if other sex workers than the one offering transaction i in the same district, where i is contracted, do not anticipate or are influenced by $Risk - taking_i$. Alternatively, we use $Instruments_i = (OthersHeight_i, OthersHeight_i^2)$, where $OthersHeight_i$ is the average value reported height by other sex workers in the same two-digit zip code between January 16 of 2011 and September 8 of 2012 and at least one day prior to contract i and $OthersHeight_i^2$ is the squared value thereof. Quite clearly, $OthersHeight_i$ and $OthersHeight_i^2$ cannot be influenced by $Risk-taking_i$ and, as long as there is some contagion in $Risk-taking_i$, we would expect characteristics of other sex workers – in particular, in a lagged fashion – to be suitable instruments for $Risk-taking_i$ (see Kuersteiner & Prucha, 2013 as well as Badinger & Egger, 2014). Usual standard errors are calculated using the observed information matrix (OIM).

4.4 Regression results

4.4.1 Main results

The results for the probit and six linear regression models are summarized in Table 4.2. The probit results indicate that overweight suppliers tend to offer unprotected services less likely than others. This is in some contrast with earlier work which pointed to a compensating role of risk-taking for personal appearance. However, the result is consistent with the frequencies of contracts in the 2×2 overweight/non-overweight and the risk-taking/non-risk-taking matrix as reported above. Moreover, there is some evidence of younger providers to be more likely to offer unprotected services. Finally, contagion is a strong factor in risk-taking: the more others offered unprotected services in the same region (at least one day) prior to the time contract i was concluded, the more likely will the worker involved in contract i offer unprotected services as well. The explanatory variables have a highly significant joint impact in the probit model and also the other models. Notice that all estimates reported pertain to regression coefficients (parameters) rather than marginal effects. Hence, the coefficients on continuous variables measure marginal effects on the *latent* desirability of or net benefit from risk-taking, a variable which has full support in real space between minus and plus infinity.

The linear regression models all suggest that risk-taking raises the hourly wage of sex workers while being overweight reduces it. There is a premium on providing information about height or age, and there is a tendency of customers to favor medium-tall and younger, especially inexperienced (first-time) suppliers. This chapter's focus is on the magnitude of the impact of $Risk-taking_i$ and $Risk-taking_i \times \widetilde{Overweight}_i$ as captured by the parameter estimates $\hat{\beta}$ and $\hat{\gamma}$ in the OLS and 2SLS models.

First of all, it stands out that $\hat{\beta}$ in the OLS models is less than half of the size of the counterpart 2SLS models. An OLS parameter of 0.4958 corresponds to a semi-elasticity of about $100[\exp(0.4958) - 1] \simeq 64\%$ for an hourly wage premium of risk-taking. According to the same model, there is a semi-elasticity of being overweight of about $100[\exp(-0.2588) - 1] \simeq -23\%$ which is *raised* (rather than

Table 4.2: *Regression results*

	Probit	OLS	OLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Risk-taking	-	0.5063***	0.4958***	1.2277***	1.3478***	1.0673***	1.1666***
	-	(0.0214)	(0.0215)	(0.1181)	(0.1627)	(0.1657)	(0.2202)
Risk-taking×$\widetilde{\text{Overweight}}$	-	-	-0.2079***	-	0.4633	-	0.3359
	-	-	(0.0537)	-	(0.3348)	-	(0.3451)
Overweight	-0.1003**	-0.2716***	-0.2588***	-0.2609***	-0.2878***	-0.2579***	-0.2777***
	(0.0411)	(0.0130)	(0.0134)	(0.0136)	(0.0235)	(0.0137)	(0.0247)
Height is provided	15.1032***	6.4907***	6.4427***	3.2645	2.9386	3.5238	3.2677
	(4.7512)	(2.0643)	(2.0634)	(2.1965)	(2.2504)	(2.1736)	(2.2257)
Height in cm if height is provided	-0.1696***	-0.0842***	-0.0837***	-0.0465*	-0.0426	-0.0499*	-0.0468*
	(0.0583)	(0.0250)	(0.0250)	(0.0265)	(0.0272)	(0.0263)	(0.0269)
Height-squared in cm if height is provided	0.0005***	0.0003***	0.0003***	0.0002**	0.0002*	0.0002**	0.0002**
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Age is provided	1.1090***	0.5507***	0.5497***	0.4546***	0.4439***	0.4200***	0.4149***
	(0.3697)	(0.1216)	(0.1216)	(0.1267)	(0.1288)	(0.1284)	(0.1295)
Age in years if age is provided	-0.0478*	-0.0203***	-0.0202***	-0.0163**	-0.0159**	-0.0147*	-0.0146*
	(0.0237)	(0.0076)	(0.0076)	(0.0079)	(0.0080)	(0.0079)	(0.0080)
Age-squared in years if age is provided	0.0004	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000
	(0.0004)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
First-time supplier	-0.0094	0.3758***	0.3766***	0.3794***	0.3779***	0.3785***	0.3774***
	(0.0707)	(0.0245)	(0.0245)	(0.0253)	(0.0257)	(0.0250)	(0.0252)
Single	0.2391***	0.3215***	0.3185***	0.2956***	0.2987***	0.2894***	0.2921***
	(0.0613)	(0.0214)	(0.0214)	(0.0226)	(0.0229)	(0.0229)	(0.0233)
History of risk-taking in same 2-digit zipcode	5.9422***	-	-	-	-	-	-
	(0.9099)	-	-	-	-	-	-
History of risk-taking in same 2-digit zipcode squared	-1.4593	-	-	-	-	-	-
	(1.7785)	-	-	-	-	-	-
Control function	-	-	-	-	-	0.3125	0.2886
	-	-	-	-	-	(0.2503)	(0.2588)
Observations	16,583	16,583	16,583	16,583	16,583	16,583	16,583
Model degrees of freedom	88	87	88	87	88	88	89
F test on joint model significance (LR test for probit)	1,358.03	55.62	55.21	47.25	47.94	47.94	46.58
First stage							
Model degrees of freedom	-	-	-	87	88	88	89
Joint F test on $\hat{\Phi}$ and $\hat{\Phi} \times \widetilde{\text{Overweight}}$							
Equation for Risk-taking	-	-	-	598.78	315.46	296.87	163.10
Equation for Risk-taking×$\widetilde{\text{Overweight}}$	-	-	-	-	472.09	-	420.94

Notes: The table presents the main regression results. Column 1 presents the probit regression results used in the 2SLS procedures (column 4-7) and columns 2 and 3 present the baseline OLS results. Usual (observed information matrix) standard errors are given in brackets and ***/**/* denote significance at the 1%/5%/10% level.

reduced) to -37% when offering unprotected services.[9]

Second, clearly the 2SLS regressions suggest that the parameters of the OLS models are biased. The instruments work well. The F-statistics for the relevance of the single identifying instrument $\hat{\Phi}_i$ for $Risk{-}taking_i$ in the models excluding $Risk{-}taking_i \times \widetilde{Overweight}_i$ are about 599 and 297 in the first stage, and the ones for the two instruments $\hat{\Phi}_i$ and $\hat{\Phi}_i \times \widetilde{Overweight}_i$ in the other first-stage 2SLS models are between 163 and 472. Neither one of the estimated 2SLS models suggests that risk-taking has a different effect on hourly wages of overweight versus other suppliers (i.e., statistically, $\hat{\gamma}$ is not distinguishable from zero). This renders the models without the interaction term preferable for efficiency reasons. The model including the control function $\eta\hat{\phi}_i$ seems preferable from an econometric point of view, as it eliminates an endogeneity bias under less stringent assumptions than the 2SLS model without the control function. In the model with the control function, the premium on risk-taking is about $100[\exp(1.0673) - 1] \simeq 91\%$. Obviously this more than compensates being overweight for the average female sex worker in the data, which reduces the hourly wage according to the same specification by $100[\exp(-0.2579) - 1] \simeq -23\%$.

4.4.2 Sensitivity analysis

In this subsection, we assess the sensitivity of the 2SLS results along four lines. First, we assess the possible impact of using weight information that had been imputed by students. For this, we exclude all transactions, where the weight information was imputed or the self-reported body-mass index (BMI) of sex workers was in the interval $24 < BMI_i < 26$ (referred to as model 7A). Alternatively, we interact the two risk measures with a binary indicator variable which is unity if the weight information was self-reported and zero if it was imputed by students based on the provided photographs. Then, we include four risk measures – the original ones and the ones that were interacted with the aforementioned binary indicator variable (referred to as model 7B). Furthermore we run regressions, where $Instruments_i = (OthersHeight_i, OthersHeight_i^2)$ rather than $Instruments_i = (RegionHistory_i, RegionHistory_i^2)$ are used (referred to as model 7C). Finally, we present results of a 2SLS model with sex-worker-specific fixed effects (referred to as model 7D). The parameters corresponding to these regressions are reported in the aforementioned order as columns (7A)-(7D) in Table 4.3, and they should be compared to column (7) in Table 4.2.

The results from those sensitivity checks may be summarized as follows. First of all, the results for the extreme-BMI subsample in column (7A) of Table 4.3 suggest that the main effect and interaction effect are statistically insignificant. However,

[9] The correlation of the disturbances between latent process of the probability of risk-taking and the outcome (wage) equation is negative. Hence, the downward bias of the OLS parameter on risk taking suggests that risk-taking sex workers have on average lower gains from risk-taking than non-risk-taking ones.

Table 4.3: *Sensitvity analysis*

	Self reported without BMI 24-26 2SLS **(7A)**	**Self reported** 2SLS **(7B)**	**Full sample Height, Height^2 as IV** 2SLS **(7C)**	**Full sample Seller fixed effects** 2SLS **(7D)**
Risk-taking	0.1580	1.1925***	1.2371***	-1.6496
	(0.2419)	(0.2184)	(0.2801)	(1.0212)
Risk-taking×$\widetilde{\text{Overweight}}$	-0.6247	0.2397	0.4250	-0.0355
	(0.3893)	(0.3434)	(0.3218)	(1.0587)
Risk-taking × Self reported BMI		-0.5946**		
		(0.2028)		
Risk-taking×$\widetilde{\text{Overweight}}$ × Self reported BMI		-0.6115		
		(0.6053)		
Overweight	-0.2909***	-0.2611***	-0.2459***	0.0696
	(0.0376)	(0.0279)	(0.0241)	(0.0601)
Height is provided	-	2.6649	-0.4284	1.3415
		(2.2239)	(2.4237)	(7.1665)
Height in cm if height is provided	0.1860***	-0.0387	-0.0060	-0.0141
	(0.0637)	(0.0269)	(0.0291)	(0.0862)
Height-squared in cm if height is provided	-0.0005***	0.0001*	0.0001	0.0000
	(0.0002)	(0.0001)	(0.0001)	(0.0003)
Age is provided	0.7876***	0.4561***	0.0409	0.0365
	(0.2467)	(0.1300)	(0.1475)	(0.1445)
Age in years if age is provided	-0.0347***	-0.0168**	0.0020	0.0008
	(0.0153)	(0.0080)	(0.0086)	(0.0088)
Age-squared in years if age is provided	0.0002	0.0001	-0.0001	-0.0001
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
First-time supplier	0.3859***	0.3820***	0.3772***	0.0890
	(0.0442)	(0.0251)	(0.0256)	(0.0567)
Single	0.3001***	0.2899***	0.2119***	0.0044
	(0.0531)	(0.0232)	(0.0275)	(0.0338)
Control function	0.0772	0.4734*	2.3352***	0.5668***
	(0.3967)	(0.2699)	(0.4584)	(0.2162)
Observations	3405	16583	16583	16583
Model degrees of freedom	54	91	89	88
F test on joint model significance (LR test for probit)	41.70	46.3	45.54	11.40
First stage				
Model degrees of freedom	54	91	89	88
F test risk	45.34	163.10	79.85	11.22
F test interaction	190.81	420.94	389.02	44.26
F test risk self reported		19.58		
F test interaction self reported		9.12		

Notes: This table presents the results of the performed sensitivity analysis. Usual (observed information matrix) standard errors are given in brackets and ***/**/* denote significance at the 1%/5%/10% level.

these results are based on only 3,405 rather than 16,583 observations and only half of the regions (2-digit zip codes) included in the full sample are still represented in the corresponding subsample. Secondly, the results in column (7B) suggest that both in the full sample and in the subsample of contracts with self-reported weight only there is a statistically significant and positive main effect and a statistically insignificant interaction effect. The interaction term of risk and the self-reporting indicator variable enters negatively. Hence, in the self reported sample the risk premium is smaller than in the full sample, but it is still positive. Thirdly, with the alternative instrument set in column (7C), the results are similar to the ones in Table 4.2. Finally, with sex-worker-specific fixed effects, none of the risk-taking effects – neither the main effect nor the interaction term – is statistically significant. The latter flows from the low degree of variation of risk-taking within sex workers over time. However, as reported in the context of the discussion of Table 4.2, the Hausman test statistic does not support the fixed effects estimator relative to the random effects model, since estimation of the fixed effects involves an enormous loss of degrees of freedom, while the parameter vectors between the two model types are not statistically significantly different.

Notice that one could estimate the two-stage least-squares models alternatively by using a first-stage linear-probability model. Doing so results in a p-value for the Sargan over-identification test of 0.1358 (hence, instrument validity is not rejected in spite of a high instrument relevance as reflected in an F-statistic on the joint relevance of 46.43) and the following parameters (standard errors) of interest: 1.0573 (0.1669) for Risk-taking; 0.8702 (0.3096) for Risk-taking$\times \widetilde{\text{Overweight}}$, and -0.3173 (0.0227) for Overweight.[10] However, we should be very careful with the interpretation of these results, since the standard errors and test statistics are all biased due to the limited-dependent-variable character of Risk-taking and the involved corner solutions (see Wooldridge (2002, p. 637)).

4.5 Conclusions

This chapter analyses the price premium on unprotected sex offerings by sex workers in Germany. The data used for the study come from a large online database, where the authors downloaded the universe of transactions between January 16 of 2011 and September 9 of 2012. With an average hourly wage of about 140 Euros, the data at hand feature in the premium segment of the market. The issue of self-selection into risky behavior through unprotected sex received particular attention in the study. Using time-lagged characteristics of sex transactions by other sex workers in the geographical neighborhood, we found that disregarding endogenous

[10]The other model results are available from the authors upon request. Notice that further test statistics suggest that the instruments do not under-identify and are now weak. Yet, as with the other test statistics, these are biased with limited dependent endogenous right-hand-side variables.

risk-taking leads to down-ward biased risk-premia. The premium on unprotected services in the data is estimated at about 91% of the average hourly wage. There is no evidence in the data that over-weight sex workers receive a lower positive premium on unprotected services than other sex workers.

Chapter 5

Fun and games: How soccer games and conventions raise the demand for paid sex

5.1 Introduction

Earlier work on the economic effects of events at large comes to mixed conclusions regarding the desirability of hosting such events from a purely economic standpoint. This chapter sheds light on effects of such events on a margin in an economy's shadow: the demand for and supply of paid sexual services. While sexual services for pay are illegal – for customers and/or providers – in some countries (e.g., Finland, France, Ireland, Norway, Sweden, the United States), they are not in others (e.g., Austria, Belgium, Germany, the Netherlands, Switzerland, United Kingdom). Cho et al. (2012) provide an excellent overview on the legality and illegality of sexual services for pay around the globe.

This chapter relates to three distinct literatures: the one on economic consequences of sports events, the one on the economic consequences of trade fairs, and the one on the demand for and supply of sexual services for pay. The chapter uses a unique, novel data-set on the offers of, bidding about, and purchases of sexual services for pay through an online platform in Germany.[1] The data-set allows to geo- and time-reference the individual transactions and provides rich information. This permits matching the services offered and purchased with data on home and away games of local soccer teams in a region, with global soccer events (such as the European Championship), and of trade fairs. To evaluate the results of this chapter, it is useful to provide a short overview of the main findings of the three related literatures.

First of all, work on the economic effects of sports events comes to rather mixed

[1] The dataset has also been used in two other studies. See chapter 4 (Egger & Lindenblatt, 2015a) and 6 (Egger & Lindenblatt, 2015c).

conclusions. Only a small body of work focuses on regular events (such as national or regional league soccer, football, basketball, or baseball games), while most of the related work focuses on mega events (such as the soccer World Cup or European Championship). Only modest economic effects tend to be found for smaller, regular events, and positive effects occur mostly – if at all – for tourism services rather than retailing (see Allmers & Maennig, 2009).[2] Even the evidence on the economic effects of large, supranational soccer events is mixed (Rose & Spiegel, 2011, is a rare example of relatively large economic effects – on trade flows; Baade & Matheson, 2004; Hagn & Maennig, 2009; Feddersen & Maennig, 2012; Ferris et al., 2012; Lamla et al., 2012; for evidence on quite modest effects on other economic outcomes).

Second, the evidence on the economic effects of trade fairs seems to be rather positive. Related research points to significant effects on tourism services and even retailing (see Tourism Vancouver, 2007; Penzkofer, 2008; Bathelt & Spigel, 2012).

Third, the sales of sexual services has been shown to be driven by standard market forces such as the size of demand (population, average income, etc.) and the size of supply (lack of outside options for sexual workers, poverty, etc.). Della Giusta et al. (2008, 2009b) provide a model and analysis of the determinants of sexual work. Moffatt & Peters (2004) provide thorough evidence on the pricing of paid sexual services. Interestingly, earlier work found evidence of only a moderate impact of mega sports events on the demand for sex work (see The German Delegation to the Multidisciplinary Group on Organised Crime, 2007; Richter et al., 2012), while it pointed to large effects of political conventions (see Cunningham & Kendall, 2011a). Although there is a lot of evidence on the influence of tourism on the demand for sex work (Oppermann, 1998; Ryan & Hall, 2005), actual data on this specific issue seems to be hard to get.[3] Tannahill (1992), Bullough & Bullough (1987), and Sanger (2006) provide an overview of the determinants and consequences of sex work in history.

Contrary to other spheres of the economy, we find rather big effects of events on the sales of paid sex. In particular, those effects tend to be larger for bigger events (e.g., European Championship soccer games) than for smaller ones (e.g., standard soccer-league games or the average trade fair), and, for sports events, they are specific to outcomes (e.g., transaction numbers supplied vs. demanded and, alternatively, pricing). Using micro-level evidence based on transaction-level data from an internet platform on sex for sale in Germany, we document that, at the level of the zip code, events stimulate the extent of transactions offered and, in particular, demanded, and they even induce effects on the pricing of sexual services for pay.

The remainder of the chapter is organized as follows. Section 5.2 provides a short overview of some hypotheses emerging from earlier work on the matter. Section 5.3

[2]This is consistent with moderate effects of such events on tax revenues (see Baade & Matheson, 2004). Moreover, some research sheds even light on quite significant negative economic effects in case of losses at ball game events (see Priks, 2010; Berument & Ceylan, 2012).

[3]As sex work is legal in Germany and illegal in some neighboring countries, even Germany has become a "sex tourism" destination.

introduces the database the results in this chapter are based upon, Section 5.4 summarizes the estimation strategy and results, and the last section concludes.

5.2 Hypotheses

Economic theory would suggest that events, and, especially, bigger events, lead to a bigger number of possible customers of such services and, hence, they should raise the price of supplied services – in case of some constraints thereon – temporarily. Furthermore, soccer events influence people's sentiments (Edmans et al., 2007). According to Della Giusta et al., 2008, 2009b, this effect is mitigated to some extent since the consumption and supply of sexual services is deemed immoral, and the associated stigmatisation of the involved agents constrains the market outcome. However, an increase in local demand for such services still means a deterioration of the outside options to the suppliers such that one would expect an increase in the transactions of sexual services for pay at times of big events.

Especially, with sports events we might expect direct (local or home) and indirect (away) effects. Specifically, we expect a potentially differential effects of home and away sports events on home team fans as such and also on them versus on opponent team fans. For similar reasons local versus global sports events might differ in their effects on the demand for sexual services, since the attachment of local fans to different events, and the number of persons following those events might differ. Moreover, there might be a differential effect even among fans of the home team with home events (ones that watch the game on television versus ones watching it in the stadium or a public place). Finally, we would expect effects to be differential for won and lost games. Hence, the nature of an event itself may have an impact on the potential demand for sexual services for pay. For that reason, we distinguish between alternative types of sports (soccer) events and allow them to trigger different effects on the demand and supply of sexual services for pay.

Unlike sports events, conventions (e.g., trade fairs) should be expected to primarily trigger local (at least, in-the-neighborhood) effects on sexual work and its consumption (Cunningham & Kendall, 2011a).

5.3 Data

In this chapter, we employ information on all services offered through *www.gesext.de*, an internet auction site that operates exactly as ebay, except that sellers can only offer sexual services and related activities. Sellers describe their services in a short text (they give a self-description and a description of the services offered, including the duration) and can either offer their services through a buy-it-now offer at a fixed price, through a second-price auction with a pre-determined end-time, or both.[4] We utilize information on all transactions administered by the platform between January 16 in 2011 and September 9 in 2012. Let us introduce a regional (2-digit zip-code) index $z = 1, ..., Z$ and a time (i.e., the day an auction ended or a buy-it-now offer was bought) index $d = 1, ...D$.[5] Moreover, let us use index $s = 1, ..., S$ to identify each individual sexual service contract. Altogether, there are $Z = 95$ zip codes, $D = 572$, days and $S = 27,967$ services sales (i.e., completed transactions) covered by the data. Another 58,181 services were offered but no buyer was willing to make a contract. During the concerned time span, the total revenues from such sales of sexual services by female providers at *www.gesext.de* amounted to 5.76 million Euros. Data on spectators of soccer-league games were collected from *www.transfermarkt.de* and numbers of trade-fair visitors from *FKM*, a trade fair service provider.

Table 5.1: *Description of the generic outcomes and determinants of the supply and demand of sex work for pay*

Variable name	**Description**
$HourlyWage_{szd}$	Hourly wage at which service s is bought in Euros
$zipDailyOffers_{zd}$	Number of auction and buy-it-now offers of sexual services in z at d
$zipDailySold_{zd}$	Number of auction and buy-it-now purchases of sexual services in z at d
$zipDailyAuctions_{zd}$	Number of auction offers of sexual services in z at d
$zipDailyAuctionsSold_{zd}$	Number of auction purchases of sexual services in z at d
$SoccerMatch_{zd}$	The home team of z plays on day d (1) or not (0)
$HomeGame_{zd}$	The home team of z plays a home game on day d (1) or not (0)
$GameLost_{zd}$	The home team of z looses the game on day d (1) or not (0)
$LnNrSoccerVisitors_{zd}$	Log number of spectators at soccer stadium in zip code z on day d
$TradeFair_{zd}$	A trade fair takes place in z on day d (1) or not (0)
$LnNrTradeFairVisitors_{zd}$	Log number of visitors at any trade fair in zip code z on day d
$ECMatch_{d}$	A European Championship match is played on day d by any team (1) or not (0)
$ECMatchDEU_{d}$	A European Championship match is played on day d by Germany (1) or not (0)
$ECMatchDEULoose_{d}$	Germany lost a European Championship match on day d (1) or not (0)
$SoccerMatchNeighborZip_{zd}$	The home team of z's neighboring zip code plays on day d (1) or not (0)
$HomeGameNeighborZip_{zd}$	The home team of z's neighboring zip code plays a home game on day d (1) or not (0)
$TradeFairNeighborZip_{zd}$	A trade fair takes place in z' neighboring zip code on day d (1) or not (0)
$DWeight_{szd}$	The seller of s provides information on her weight (1) or not (0)
$Overweight_{szd}$	The seller of s is overweight (1) or not (0)
$DHeight_{szd}$	The seller of s provides information on her height (1) or not (0)
$Height_{szd}$	The seller's height
$Height^2_{szd}$	The seller's height squared
$DAge_{szd}$	The seller of s provides information on her age (1) or not (0)
Age_{szd}	The seller's age
Age^2_{szd}	The seller's age squared
$First_{szd}$	The seller of s claims to be a first-time supplier (1) or not (0)
$Single_{szd}$	The seller of s claims to be single (1) or not (0)
Weekday FE	A fixed effect for the weekday of d is included
Month FE	A fixed effect for the month of d is included
Zip code FE	A fixed effect for z is included

Notes: Indices s, z, and d refer to sold (or offered) services, 2-digit zip codes, and days, respectively.

Table 5.1 describes the variables used in this chapter. There are two types of variables, a generic outcome y_{szd} and generic determinants thereof, $x_{k,szd}$. As

[4] To obtain the data we used a self-written download program. Information on each transaction was then extracted by students we hired.

[5] Buy-it-now prices should differ from auction prices since a spatio-temporal equilibrium may be found with an auction but not a price. In the presence of a variable elasticity along the demand curve for sexual services, we would then also expect prices and demand to react differently to shocks on buy-it-now versus auctioned transactions.

generic outcomes, we consider the contracted hourly wage of a sexual service (*HourlyWage*$_{szd}$, the contracted price divided by the duration of the service offered) and the number of services contracted per zip code z and time d (*AuctionsBuyItNows*$_{zd}$ for auctions as well as buy-it-now contracts; *Auctions*$_{zd}$ for auctions only). Among the generic determinants, four types stand out, one related to soccer game events of teams that have their home stadium in zip code z (*SoccerMatch*$_{zd}$, *HomeGame*$_{zd}$, *GameLost*$_{zd}$ and *LnNrSoccerVisitors*$_{zd}$), one related to European Championship soccer game events (three variables: *ECMatch*$_{d}$, *ECMatchDEU*$_{d}$ and *ECMatchDEULoose*$_{d}$), one related to trade fairs held in zip code z at time s (*TradeFair*$_{zd}$ and *LnNrTradeFairVisitors*$_{zd}$), and one related to indirect effects of such events on neighboring zip codes (*SoccerMatchNeighborZip*$_{zd}$, *HomeGameNeighborZip*$_{zd}$, and *TradeFairNeighborZip*$_{zd}$). For identification of the parameters on the event variables of particular interest, we condition on a number of seller-specific characteristics (*DWeight*$_{szd}$, *Overweight*$_{szd}$, *DHeight*$_{szd}$, *Height*$_{szd}$, *Height*$^2_{szd}$, *DAge*$_{szd}$, *Age*$_{szd}$, *Age*$^2_{szd}$, *First*$_{szd}$ and *Single*$_{szd}$), time characteristics (weekday and month fixed effects), as well as regional characteristics (zip-code fixed effects).

Table 5.2: *Descriptive statistics*

		Sold services		Offered services	
Dep. variables	**Type**	**Mean**	**SD**	**Mean**	**SD**
(1)	(2)	(3)	(4)	(5)	(6)
$HourlyWage_{szd}$	continuous	102.2847	57.4246		
$zipDailyOffers_{zd}$	continuous			1.6567	2.2873
$zipDailySold_{zd}$	continuous	0.5403	0.9788		
$zipDailyAuctions_{zd}$	continuous			0.9151	1.4974
$zipDailyAuctionsSold_{zd}$	continuous	0.2853	0.6753		
Indep. variables	**Type**	**Mean**	**SD**	**Mean**	**SD**
$SoccerMatch_{zd}$	binary	0.0609	0.2391	0.0569	0.2316
$HomeGame_{zd}$	binary	0.0309	0.1730	0.0287	0.1670
$GameLost_{zd}$	binary	0.0064	0.0797	0.0051	0.0711
$TradeFair_{zd}$	binary	0.0860	0.2804	0.0781	0.2683
$ECMatch_{d}$	binary	0.0296	0.1695	0.0331	0.1788
$ECMatchDEU_{d}$	binary	0.0075	0.0860	0.0078	0.0879
$ECMatchDEULoose_{d}$	binary	0.0001	0.0090	0.0001	0.0085
$LnNrSoccerVisitors_{zd}$	continuous	10.2097	1.1944	10.1484	1.4046
$LnNrTradeFairVisitors_{zd}$	continuous	8.2836	1.7430	8.1918	1.7371
$SoccerMatchNeighborZip_{zd}$	binary	0.1066	0.3086	0.0990	0.2987
$HomeGameNeighborZip_{zd}$	binary	0.0642	0.2452	0.0602	0.2380
$TradeFairNeighborZip_{szd}$	binary	0.1618	0.3683	0.1522	0.3592
$DWeight_{szd}$	binary	0.3235	0.4678	0.3534	0.4780
$Overweight_{szd}$	binary	0.2617	0.4396	0.1680	0.3739
$DHeight_{szd}$	binary	0.5954	0.4908	0.5954	0.4908
$Height_{szd}$	continuous	166.9342	6.2472	166.7331	6.0491
$Height^2_{szd}$	continuous	27906.0600	2083.8930	27836.5000	2003.7300
$DAge_{szd}$	binary	0.6608	0.4735	0.6958	0.4601
Age_{szd}	continuous	31.7944	9.0073	29.7270	8.3915
Age^2_{szd}	continuous	1092.0150	626.8656	954.1062	567.2429
$First_{szd}$	discrete	0.0306	0.1724	0.0450	0.2073
$Single_{szd}$	binary	0.0531	0.2242	0.0847	0.2785

Notes: This table presents the descriptive statistics of the variables used in this chapter. Indices s, z, and d refer to sold (or offered) services, 2-digit zip codes, and days, respectively. The total number of services sold (offered) is 27,967 (86,148) and the underlying number of service suppliers is 3,206 (3,773). The total number of zip codes and days covered is 95, and the total number of days covered is 572. The first two columns correspond to sold services only, while Column 3 and 4 correspond to all services offered.

Table 5.2 summarizes the first and second moments of these dependent and independent variables in use. We mainly utilize data on actually sold services of which there are 27,967, covering 3,206 suppliers during the covered time span (columns 3 and 4). As can be seen from Column 2 of Table 5.2, six of the explanatory variables are continuous while the others are binary determinants of generic outcome. On average, contracted services are sold at about 102 Euros per hour. About 6% of the services are offered on the day a soccer team plays a game at home or elsewhere – whose home stadium is located in the area of the respective 2-digit zip code. Hence, about 3% of the services are offered at 2-digit zip-code areas where a team plays a home game.[6] Roughly 3% of the services were offered during game days of the European Championship 2012, which was hosted in Ukraine and Poland. When the German team played, less than 1% of the services were purchased. Nearly 9% of the services were sold while a trade fair took place in the same 2-digit zip code area.[7] About 3% and 5% of the sexual service suppliers claim to offer services for the first time and to be single, respectively. The descriptive statistics are similar between the samples of actually contracted (sold) and just offered services.

Sports events and conventions are known to induce geographically bound economic effects. Clearly, the magnitude and the geographical reach of the effects depend on the size of those events, i.e., the number of spectators and visitors. We would expect proportional effects on the demand and supply of sexual services for pay. In order to convey an idea about the geography and size of such events in Germany, we prepared Figures 5.1a and 5.1b which display 2-digit zip codes[8] in darker shade, if the maximum size of a soccer-league team's stadium or the number of trade fair visitors is larger, respectively.

The figures suggest that larger agglomerations (such as Munich, Berlin, Hamburg, Frankfurt, and some cities in Northrhein-Westphalia) host bigger soccer events, on average. However, bigger trade fairs with more than 200,000 visitors each are hosted also outside of the four largest agglomerations in Germany (e.g., in Hannover, Mannheim, Düsseldorf, Leipzig, or Nuremberg). The subsequent analysis will assess the effects of such events (within limited geographical and temporal boundaries) on the supply and demand of sex for sale in Germany.

[6]Each team in a pair has to play one home and one away game.

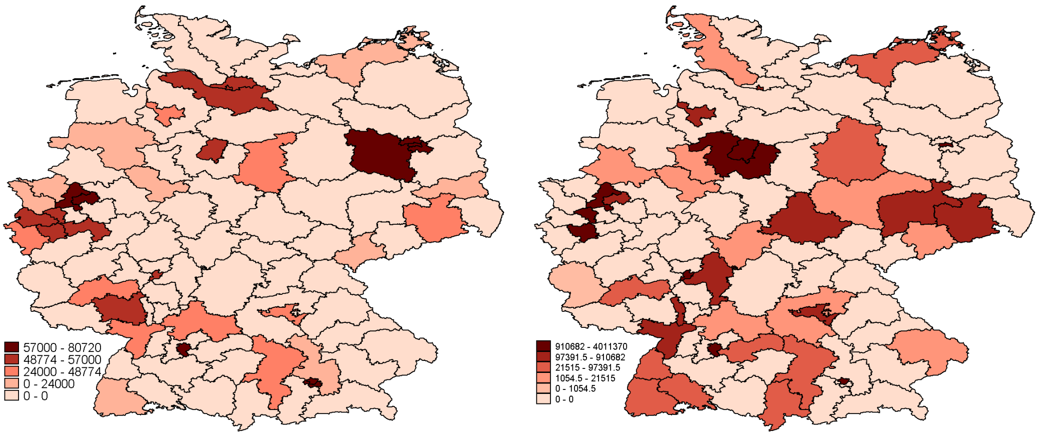
[7]Games in the German soccer league are generally played on Fridays, Saturdays, and Sundays. Since there are Fridays, Saturdays, and Sundays with and without league games, weekday fixed effects are identified separately from game effects.

[8]Germany uses a 5-digit system to identify postal areas. The data are aggregated to areas corresponding to the first two digits of this postal zip code.

Figure 5.1: *Soccer stadiums and trade fair visitors*

These two maps present the maximum capacity of soccer stadiums and annual number of trade fair visitors on the 2 digit zip code level (zip code z). Areas with a higher value of the variable are marked in darker color. *Source: Own representation.*

(a) Maximum capacity of soccer stadiums (b) Total number of trade fair visitors



5.4 Empirical analysis

5.4.1 Econometric approach

In the empirical model, we utilize information on generic outcomes y_{szd}. We model y_{szd} as an exponential function of $x_{szd}\beta$, where x_{szd} is a $1 \times K$ vector of determinants with kth element $x_{k,szd}$ (including fixed effects) and β is an unknown parameter vector. Notice that each contract s is uniquely related to a particular day d and zip code z. Days d and zip codes z are repeatedly observed in the data. The three main advantages of exponential-family models are that (i) the boundedness of the dependent variable is respected, (ii) that the variance-covariance matrix of the parameters may be corrected for heteroskedasticity of unknown form (Fisher, 1949; Nelder & Wedderburn, 1972; Wedderburn, 1974), and (iii) that the parameters β may be interpreted as elasticities and semi-elasticities (depending on whether x_{szd} is measured in logs or levels), unlike with most non-linear models. For that reason we employ Poisson quasi-maximum-likelihood estimators as advocated by Cameron & Trivedi (2005).

Defining the conditional mean of y_{szd} by $\overline{y}_{szd}$, the econometric models estimated and their first-order conditions (FOC) read

$$\text{Model: } y_{szd} = \exp(x_{szd}\beta) + u_{szd}, \tag{5.1}$$

$$\text{FOC: } \sum_{s=1}^{S} (y_{szd} - \overline{y}_{szd}) x_{szd} = 0, \tag{5.2}$$

where FOC is a $1 \times K$ vector. The parameters $\hat{\beta}$ can be found by following the procedure in Wedderburn (1974) which essentially searches for $\hat{\beta}$ in (5.1) by maintaining (5.2). Usual standard errors are calculated using the observed information matrix (OIM).

5.4.2 Results

The regression results are summarized in Tables 5.3-5.5. Tables 5.3 and 5.4 contain five and Table 5.5 contains six numbered columns, which refer to different specifications in terms of explanatory variables used ($x_{k,szd}$). We first look at the influences of events on supply and demand on this platform,[9] being able to distinguish between the extensive and the intensive margin. The dependent variables are number of sellers, number of offered sexual services per seller, number of bidders, number of bids per bidder and number of contracted sexual services, each per zip code z and day d. While Table 5.3 presents result for the whole sample of auctions and buy-now offers, 5.4 presents results for auction offers only. Hence, Tables 5.3 and 5.4 use grouped data such that the index szd is generally replaced by the index zd in equations (5.1) and (5.2). We then consider the dependent variable hourly wages (Table 5.5).

Table 5.3: *Regression results A: Supply and demand, Auctions and Buy-now offers*

Dependent variable	Events per region and day ($Events_{zd}$) - Auctions and Buy-now				
	Nr. of sellers (1)	Nr. of offers per seller (2)	Nr. of bidders (3)	Nr. of bids per bidder (4)	Nr. of sold offers (5)
$SoccerMatch_{zd}$	0.4880**	0.3244	0.8316***	0.5481	0.6619***
	(0.2232)	(0.2892)	(0.1225)	(0.3503)	(0.2065)
$HomeGame_{zd}$	0.0541	0.0175	0.0357	0.0351	0.0409
	(0.0683)	(0.0812)	(0.0409)	(0.1041)	(0.0670)
$TradeFair_{zd}$	0.3486***	0.4678***	0.1166	0.4910***	0.3384***
	(0.1233)	(0.1495)	(0.0739)	(0.1917)	(0.1210)
$ECMatch_d$	0.5956***	0.6958***	0.7005***	0.7374***	0.6187***
	(0.0446)	(0.0512)	(0.0260)	(0.0663)	(0.0434)
$ECMatchDEU_d$	-0.1405*	-0.1257	-0.2284***	-0.2123*	-0.1352*
	(0.0815)	(0.0926)	(0.0487)	(0.1235)	(0.0787)
$LnNrSoccerVisitors_{zd}$	-0.0223	-0.0030	-0.0596***	-0.0234	-0.0394**
	(0.0214)	(0.0278)	(0.0116)	(0.0335)	(0.0198)
$LnNrTradeFairVisitors_{zd}$	-0.0037	-0.0179	0.0288***	-0.0211	-0.0033
	(0.0151)	(0.0185)	(0.0090)	(0.0237)	(0.0149)
$SoccerMatchneighborZip_{zd}$	0.6459***	0.7133***	0.6655***	0.7423***	0.6222***
	(0.0471)	(0.0525)	(0.0279)	(0.0684)	(0.0467)
$HomeGameneighborZip_{zd}$	0.0325	0.0084	0.0356	-0.0057	0.0455
	(0.0548)	(0.0613)	(0.0325)	(0.0798)	(0.0544)
$TradeFairneighborZip_{szd}$	0.7294***	0.8086***	0.7086***	0.8146***	0.7190***
	(0.0234)	(0.0266)	(0.0142)	(0.0352)	(0.0231)
Weekday FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Zipcode FE	Yes	Yes	Yes	Yes	Yes
Nr. of observations	45125	45125	45125	45125	45125
Nr. of events	19774	15236	55349	8930	20499

Notes: The table presents the regression results of the supply and demand regressions using the full sample of auction and buy-now offers. Indices z and d refer to 2-digit zip codes and days. Usual (observed information matrix) standard errors are given in brackets and ***/**/* denote significance at the 1%/5%/10% level.

[9] The data do not permit an analysis of displacement effects of offered off-line services by online ones. Moreover, they do also not permit an analysis of displacement effects of services offered at other online platforms by the considered one. However, it seems unlikely that there would be a bias exactly towards or away from the considered platform. Moreover, since we include zip code fixed effects, the results should be interpreted as to reflect direct effects within zip codes from events and other fundamentals within that or neighboring zip codes.

Table 5.4: *Regression results B: Supply and demand, Auctions only*

Dependent variable	Events per region and day ($Events_{zd}$) - Auctions only				
	Nr. of sellers (1)	Nr. of offers per seller (2)	Nr. of bidders (3)	Nr. of bids per bidder (4)	Nr. of sold offers (5)
$SoccerMatch_{zd}$	0.2034	0.1169	0.1310	0.6738	0.1801
	(0.3462)	(0.3850)	(0.3042)	(0.7638)	(0.3447)
$HomeGame_{zd}$	0.0648	0.0671	0.1157	0.1962	0.0433
	(0.1014)	(0.1074)	(0.0892)	(0.2702)	(0.1012)
$TradeFair_{zd}$	0.4658***	0.5956***	0.3705**	1.5383***	0.4715***
	(0.1785)	(0.1893)	(0.1571)	(0.4563)	(0.1767)
$ECMatch_{d}$	0.4985***	0.5412***	0.4270***	0.2597	0.4550***
	(0.0657)	(0.0696)	(0.0604)	(0.1768)	(0.0658)
$ECMatchDEU_{d}$	-0.0839	-0.0140	0.0020	0.1422	-0.0754
	(0.1185)	(0.1222)	(0.1065)	(0.2987)	(0.1187)
$LnNrSoccerVisitors_{zd}$	0.0009	0.0123	0.0031	-0.0449	0.0037
	(0.0334)	(0.0371)	(0.0293)	(0.0725)	(0.0332)
$LnNrTradeFairVisitors_{zd}$	-0.0193	-0.0281	-0.0049	-0.1597***	-0.0178
	(0.0218)	(0.0233)	(0.0191)	(0.0581)	(0.0216)
$SoccerMatchneighborZip_{zd}$	0.6410***	0.6697***	0.6293***	0.8198***	0.6141***
	(0.0681)	(0.0710)	(0.0604)	(0.1730)	(0.0686)
$HomeGameneighborZip_{zd}$	0.0386	0.0596	0.0687	-0.1989	0.0612
	(0.0791)	(0.0822)	(0.0701)	(0.2076)	(0.0796)
$TradeFairneighborZip_{szd}$	0.7396***	0.7563***	0.7561***	0.7520***	0.7411***
	(0.0330)	(0.0350)	(0.0292)	(0.0861)	(0.0328)
Weekday FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Zipcode FE	Yes	Yes	Yes	Yes	Yes
Nr. of observations	45125	45125	45125	45125	45125
Nr. of events	9614	8574	12088	1506	9689

Notes: The table presents the regression results of the supply and demand regressions using the sub-sample of auction offers only. Indices z and d refer to 2-digit zip codes and days. Usual (observed information matrix) standard errors are given in brackets and ***/**/* denote significance at the 1%/5%/10% level.

Supply and Demand estimations

Column (1) in Table 5.3 clearly indicates that if the home team of a zip-code area has a match day (there is no difference between a home or an away game) and if a convention is held there, the supply of sexual services both in the respective zip code and in neighboring zip codes rises. The number of sellers in the respective zip code per day rises on days with soccer-league games by $100(\exp(0.4880) - 1) \simeq 62.9\%$, and on convention days by $100(\exp(0.3486) - 1) \simeq 41.7\%$. In neighboring zip codes these numbers are 90.8% and 107.4%, respectively. Additionally, in all zip codes the number of sellers rises by 81.4% on European-Championship-match days (13.1% less when Germany played). While there are more sellers in the respective zip code areas on days with soccer-league games, sellers do not offer more services (see Column 2). In contrast to that, not only the number of sellers, but also the number of offered services per seller rises in a zip code if conventions are taking place there (59.6%), and it also rises in neighboring zip codes on days with soccer-league games (104.1%) and trade-convention days (124.5%), as well as in all zip codes on European-Championship-match days (100.5%).

On the demand side, when a local soccer team has a match day, the number of men placing at least one bid is 129.7% higher in the respective zip code and 94.5% higher in neighboring areas (see Column 3). The average trade fair has an effect of 25% on demand at the zip code where the convention takes place ($TradeFair_{zd}$ is not significant, but $LnNrTradeFair_{zd}$ is) and 103.1% in neighboring zip codes. European Championship matches raise the number of bidders by 101.5% in all zip codes (German European Championship matches lessen this effect by 20.4%). When looking at the intensive margin (see Column 4), soccer-match days do not raise the number of bids per bidder, but these rise due to trade fairs taking place in the zip code itself (63.4%) and in neighboring zip codes (125.8%), as well as in neighboring zip codes on soccer days (110.1%) and in all zip codes on European-Championship-match days (109.1%).

Finally, the number of contracted services in this market is raised by all five event variables (see Column 5). Interestingly, the direct effect of a soccer match is more than twice as high as that of a trade fair (93.8% vs. 40.3%), but the indirect one is smaller (86.3% vs. 105.2%).

According to Table 5.4, when considering auctions rather than auctions together with buy-it-now transactions, there is no direct effect of soccer matches on the five supply and demand variables. While direct effects of trade fairs on all five variables are slightly larger than in the whole sample, European-Championship-match effects are slightly smaller. Indirect effects of soccer games and trade fairs on neighboring zip codes are of the same size as in the whole sample.

Summing up and accounting for direct as well as indirect effects, the aggregate effect of soccer-league games on the number of sellers and the number of paid sex services offered per seller through *www.gesext.de* is about 113.7% and 35.2%. The combined effect on the number of bidders, bids per bidder, and contracted services is 66.5%, 36.8%, and 69.2%, respectively. The effect of soccer games plus conventions in Germany on the numbers of services supplied and demanded amounts to about 277.0% (number of sellers), 175.2% (number offers per seller), 155.4% (number of bidders), 185.9% (number of bids per bidder), and 192.4% (number of contracted services), respectively. Sex workers seem to anticipate the higher demand and offer more services on days of soccer events and trade fairs.

Price estimations Table 5.5 looks at the second dimension of effects of events on sexual services for pay: log hourly wages at which these services are contracted. Columns (1) and (4) employ the zip code z-specific soccer-team and trade-fair variables. Columns (2) and (5) add the European Championship soccer variables and Columns (3) and (6) add information on visitor numbers of soccer games and trade fairs as well as indicators regarding events in neighboring zip codes. Columns (1) - (3) are based on data where sexual service contracts are an outcome of a true auction or a buy-it-now sale, while Columns (4) - (6) include auctions only. Columns (1)-(6) in Table 5.5 suggest that both types of soccer-game events (local teams' matches and European Championship matches) as well as trade fairs matter for the contracted price of sexual services for pay. Events seem to have a relatively bigger effect on the price determined by auctions than by buy-it-now. To see this, compare the coefficients in Columns (1) and (2) with those in Columns (4) and (5), respectively. Moreover, as expected, adding European Championship events to local soccer team events does not change the coefficients on the latter. The reason is simply that the European Championship happened out of the German league games season. Let us first focus on a discussion of the results in Columns (2) and (4).

Table 5.5: *Regression results C: Price estimations*

Dependent variable			$HourlyWage_{szd}$			
	(1)	(2)	(3)	(4)	(5)	(6)
$SoccerMatch_{zd}$	0.0133***	0.0129***	-0.2138***	0.0669***	0.0665***	-0.1736***
	(0.0038)	(0.0038)	(0.0208)	(0.0053)	(0.0053)	(0.0263)
$HomeGame_{zd}$	-0.0168***	-0.0167***	-0.0143***	-0.0487***	-0.0485***	-0.0466***
	(0.0049)	(0.0049)	(0.0050)	(0.0069)	(0.0069)	(0.0069)
$GameLost_{zd}$	-0.0307***	-0.0308***	-0.0281***	-0.1047***	-0.1048***	-0.1013***
	(0.0086)	(0.0086)	(0.0086)	(0.0120)	(0.0120)	(0.0120)
$TradeFair_{zd}$	0.0208***	0.0210***	0.1086***	0.0174***	0.0181***	0.1510***
	(0.0024)	(0.0024)	(0.0104)	(0.0036)	(0.0036)	(0.0154)
$ECMatch_{d}$		-0.0643***	-0.0658***		-0.0907***	-0.0921***
		(0.0043)	(0.0043)		(0.0062)	(0.0062)
$ECMatchDEU_{d}$		0.0443***	0.0432***		0.1346***	0.1339***
		0.0078	(0.0078)		(0.0108)	(0.0108)
$ECMatchDEULoose_{d}$		-0.1108	-0.1045		-0.2241**	-0.2163**
		(0.0686)	(0.0686)		(0.1008)	(0.1008)
$LnNrSoccerVisitors_{zd}$			0.0229***			0.0242***
			(0.0020)			(0.0026)
$LnNrTradeFairVisitors_{zd}$			-0.0110***			-0.0167***
			(0.0012)			(0.0019)
$SoccerMatchNeighborZip_{zd}$			0.0014			0.0094**
			(0.0033)			(0.0048)
$HomeGameNeighborZip_{zd}$			-0.0386***			-0.0526***
			(0.0039)			(0.0057)
$TradeFairNeighborZip_{zd}$			0.0148***			0.0161***
			(0.0018)			(0.0027)
$DWeight_{szd}$	0.0091***	0.0095***	0.0092***	0.0075***	0.0083***	0.0077***
	(0.0016)	(0.0016)	(0.0016)	(0.0024)	(0.0024)	(0.0024)
$Overweight_{szd}$	-0.2158***	-0.2161***	-0.2162***	-0.2262***	-0.2264***	-0.2268***
	(0.0018)	(0.0018)	(0.0018)	(0.0026)	(0.0026)	(0.0026)
$DHeight_{szd}$	-3.2891***	-3.1832***	-3.2336***	-5.0934***	-4.9864***	-5.0325***
	(0.3069)	(0.3071)	(0.3073)	(0.3973)	(0.3977)	(0.3978)
$Height_{szd}$	0.0336***	0.0323***	0.0329***	0.0566***	0.0553***	0.0559***
	(0.0037)	(0.0037)	(0.0037)	(0.0048)	(0.0048)	(0.0048)
$Height^2_{szd}$	-0.0001***	-0.0001***	-0.0001***	-0.0002***	-0.0002***	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$DAge_{szd}$	0.5171***	0.5152***	0.5157***	0.7065***	0.7060***	0.7097***
	(0.0107)	(0.0107)	(0.0107)	(0.0163)	(0.0163)	(0.0163)
Age_{szd}	-0.0201***	-0.0200***	-0.0200***	-0.0361***	-0.0361***	-0.0363***
	(0.0007)	(0.0007)	(0.0007)	(0.0010)	(0.0010)	(0.0010)
Age^2_{szd}	0.0001***	0.0001***	0.0001***	0.0004***	0.0004***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$First_{szd}$	0.1211***	0.1207***	0.1204***	0.1716***	0.1709***	0.1712***
	(0.0027)	(0.0027)	(0.0027)	(0.0036)	(0.0036)	(0.0036)
$Single_{szd}$	-0.1303***	-0.1306***	-0.1302***	-0.0794***	-0.0800***	-0.0792***
	(0.0024)	(0.0024)	(0.0024)	(0.0035)	(0.0035)	(0.0035)
Weekday FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes	Yes
Nr. of observations	27967	27967	27967	14768	14768	14768
Observations included	All	All	All	Auctions only	Auctions only	Auctions only

Notes: This table presents the results of the price estimations. The first three columns use the whole data set, while columns (4) to (6) use the sub-sample of auction offers only. Indices s, z, and d refer to sold (or offered) services, 2-digit zip codes, and days, respectively. Usual (observed information matrix) standard errors are reported in brackets below the regression coefficients and ***/**/* denote significance at the 1%/5%/10% level.

Column (2) of Table 5.5 suggests that won away games raise the hourly wage by $100(\exp(0.0129) - 1) \simeq 1.3\%$, while won home games in the German soccer league lower the average contracted price per sexual service offer for pay by about $100(\exp(0.0129 - 0.0167) - 1) \simeq -0.4\%$. Moreover, lost away games in the German soccer league reduce the average contracted price per sexual service offer for pay by $100(\exp(0.0129 - 0.0307) - 1) \simeq -1.7\%$ and lost home games even reduce it by about $100(\exp(0.0129 - 0.0167 - 0.0307) - 1) \simeq -3.4\%$. The average price effects are a lot stronger for European Championship games, where the average effect on games not involving Germany was negative ($\simeq -6.2\%$), and the average effect of German games was less negative for won games ($100(\exp(-0.0643 + 0.0443) - 1) \simeq -2.0\%$) and more negative for lost games ($100(\exp(-0.0643 + 0.0443 - 0.1108) - 1) \simeq -12.3\%$). The effect of the average trade fair on prices is about 2.1%.

In general, the effects on auction prices (where the price is not determined ex ante by the service supplier) are much bigger. To see this, compare Column (5) with Column (2) in Table 5.5. For instance, on days where the local soccer team won an away league game, customers were willing to pay a price premium of about 6.9% per auction relative to other days. On the other hand, on days where the local soccer team lost a home game, paid sexual service suppliers had to take a pay cut of more than 8%. Also the predatory effect of European Championship games is much bigger on auction prices than on buy-it-now prices, according to the comparison of Column (5) with Column (2). With regard to auction prices, the local price premium on local paid sex services associated with a won away game in the German soccer league was more than three times bigger than that of a local trade fair (6.9% vs.1.9%)! Clearly 50% of the German league games in the data were played at home. On average (i.e., across home and away games), about 38% of these games were lost by the local team. Hence, the weighted average direct effect of German league events on the local hourly price (within the same zip code) sex workers could charge was about −0.7%. That effect was about 0.2% for auction prices.

Columns (3) and (6) include indicators for events (soccer games and trade shows) taking place at a neighboring zip code (indirect effects) as well as visitor numbers of soccer games and trade fairs. Marginal effects of an away game are not significant, while they are negative for a home game (−3.8%). In contrast, marginal effects of visitor numbers at trade fairs in a neighboring zip code are positive and significant (1.5%). For events taking place within a zip code, marginal effects were larger for smaller soccer events, but smaller for larger trade fairs. As we have seen, soccer games and conventions display direct and indirect, local and neighborhood effects which may reinforce and mitigate each other. One may calculate aggregate effects of soccer games (and soccer games plus trade fairs) on hourly wages in Germany – i.e., accounting for direct as well as indirect effects. Aggregate effects of soccer events and soccer events plus conventions, respectively, on hourly wages of auctions and buy-it-nows are relatively small (−0.14% and 2.5% respectively). These effects are slightly larger if only auctions are considered (3.1% and 5.3%). In general, evidence of larger effects on auctions than on buy-it-now prices suggests that the demand effect is bigger than the supply effect when considering prices.

5.5 Conclusion

This chapter analyses the influence of soccer and trade fair events on the local supply and demand of sexual services for pay in Germany. To provide evidence that events are relatively important, we use a transaction-level data-set on the offers and purchases of sexual services, extracted from a German online platform. Soccer game (German league or European Championship) and trade fair events both raise the number of sellers, bidders and contracted services on this platform in the high double-digit percentage range. Effects on prices are smaller, but significant as well. For soccer events, these effects depend on whether events happen locally (home games versus away games), and how the local team fairs (looses or wins). Overall, the estimated effects are rather large for the numbers of active sex workers (113% for soccer games and 277% for local soccer events and conventions together) as well as bidders (66% and 155%). These effects are geographical in scope and display some contagion across 2-digit zip code boundaries. Effects on prices are more moderate with e.g. $3-5\%$ for auctions.

While we only look at the tip of the iceberg (5.6 million Euros earned by $3,206$ suppliers), the total market for sexual services offered in Germany is huge. In 2003 it has been estimated to include $400,000$ sex workers earning 14.5 billion euros (or 0.4% of the German GDP) and is likely to be even larger today (Reichel & Topper, 2003). From a broader perspective, insights into this matter are relevant since (especially, unprotected) sexual work for pay is known to contribute to the spreading of certain diseases (see Rao et al., 2003; Gertler et al., 2005) and the supply of sexual services is linked with criminal activities related to the control of sexual workers and human trafficking (see Cho et al., 2012). Also in the data at hand, about 8% of the services offered are unprotected services. Hence, we might conclude that a shadow consequence of big events is that they contribute to the generation of adverse health effects and the existence of criminal activity via their stimulating effect on the demand and supply of sex work. With a view on policy implications, we might conclude that, due to a higher supply and demand in relatively narrow spatial aggregates and specific time intervals around conventions and sports events, the monitoring of sexual workers could be relatively effectively administered and improved.

Chapter 6

The long shadow of the iron curtain for female sex workers in German cities: Border effects and regional differences

6.1 Introduction

In contrast to many other countries in Europe or in the United States, sex work is legal in Germany, and, according to various sources it constitutes activity of nontrivial economic magnitude (see Reichel & Topper, 2003; Kavemann & Steffan, 2013). Interest in sex work is broad in the social sciences, and it has three major roots: first, fostering the understanding of the driving forces of the supply and demand of sex work, since it is associated with social stigma, a lack of outside opportunities, a complete lack of or reduced access to social security and pension systems, and with economic restraints (see Cameron & Collins, 2003; Moffatt & Peters, 2004; Sanders, 2004; Arunachalam & Shah, 2008; Collins & Judge, 2008; Hubbard & Whowell, 2008; Döring, 2014); second, understanding how the aforementioned stigmatization of sex workers and their activity, in many countries, at the boundaries of legality, lead to a coexistence of sex work with criminal activity at large, ranging from drug use and trading over human trafficking to tax fraud (see Albrecht et al., 1999; Kubrin & Weitzer, 2003; Cho et al., 2012; Döring, 2014); third, the understanding of the circumstances under which sex workers expose themselves and their customers to risky services through – in particular, unprotected – sex work and an associated transmission of diseases (see Rao et al., 2003; Gertler et al., 2005; Cunningham & Kendall, 2010b; de la Torre et al., 2010; Robinson et al., 2011; Arunachalam & Shah, 2013).

Apart from individual characteristics of people selecting into sex work, the social sciences point to regionally or spatially bound social and economic conditions which stimulate or deter the supply and demand of sexual services for pay. For instance, research in sociology points to a geographical dimension of morality beyond legality (see Hubbard, 2001; Hubbard & Sanders, 2003; Hubbard & Whowell, 2008; Cameron, 2004): local authorities or subgroups of citizens with relatively homogeneous views determine the local level of acceptance of sex work. This view is consistent with Green's theory of *sexual fields* (Green, 2008, 2011), the aggregation of individual sexual desires in some space such as geography. To the extent that social fields induce a gravitational pull, they induce a diversification of regions and places, lead to a selection of individuals in and to those places, and they shape the habit and behavior of individuals living therein. Moreover, regions and places are not only characterized by differences in preferences, norms, and social conditions, they also differ in terms of the enforceability of norms on non-compliers as emphasized by the theory of *social disorganization* (Kubrin & Weitzer, 2003). Hence, there is a geographical dimension to the emergence and acceptance of sex work, which has to do with differences in preferences across places but also with the enforceability of norms.

One regional aspect of attitudes towards and prevalence of sex work in Germany clearly has to do with the location of the former Iron Curtain along the boundaries of what was called the German Democratic Republic (GDR) and the Federal Republic of Germany (FRG). The former border between the two historical Germanies marks a multidimensional space of discontinuities in preferences and norms which still persist more than 20 years after the reunification (for instance, see Alesina & Fuchs-Schündeln, 2007, for general differences in preferences between Germans in the former East and West). Bauernschuster & Rainer (2010) show that Germans in the so-called *New Länder* (the former East) hold more egalitarian sex role attitudes, and there is no sign of this difference to vanish yet. Heineck & Süssmuth (2010), Becker & Hainz (2011), Vosskort & Necker (2014) as well as Dohmen et al. (2011) show that differences in preferences and attitudes may generally be persistent for generations. In the context of sex work the demarcation of regions and places along the boundaries of the former Iron Curtain is interesting, since the two historical Germanies differed substantially by legal treatment of sex work as well as the tolerance and social status of sex workers, as will become clear in the next section.

The present chapter uses online auction data of 3,204 sex workers in Germany across more than 27,946 transactions over the period of January 16 2011 to September 9, 2012. The results suggest that, indeed, there is an important regional component in the price sex workers can achieve and in the probability of their engaging in risky behavior through offering unprotected (unsafe) sexual services. We find evidence of a specific component accruing to larger cities in general and to Berlin in particular. Moreover, we find that, conditional on other determinants, prices are higher where sex work was illegal during the separation of the two Germanies, namely in the New Länder (the East). This gap in prices is larger in larger agglomerations along the former Iron Curtain. Moreover, we find a strong network or peer-group component in both the pricing as well as the unprotected sex services. The latter supports theories claiming a formative effect of regions due to the prevalence of certain norms and the ability to enforce them, e.g. through existing social fields.

The remainder of the chapter provides an outline of sex work in post-World-War-II Germany in Section 6.2. It introduces and describes data on sex work activity collected from a large online platform which features ebay-type (second-price auction) mechanisms for sex work in Section 6.3. The chapter presents a set of methods in pursuit of specific research questions and a summary of the associated findings in Section 6.4, and it offers a short set of conclusions in the last section.

6.2 Sex work and sex workers in Germany

6.2.1 Sex work in the GDR versus the FRG

In the early days of the Soviet empire, prostitution was deemed inconsistent with the role of women in society, and prostitutes were "encouraged" to engage in "regular" work. Prostitution was free and tolerated in most Western European countries (and still is to date in many), including Germany, at the time. However, the formation of the Soviet bloc and the occupation of six of Germany's federal states with the result of the formation of the Federal Republic of Germany (FRG) in the West and the German Democratic Republic (GDR) in the East led to unconstrained activity of prostitutes in the former and to constrained activity of prostitutes in the latter. At first, prostitutes in the GDR were only allowed to pursue their activity part-time, and ultimately prostitution was declared illegal with a threat of prosecution from 1968 onward.[1]

In spite of its illegality, prostitution existed and was tolerated in urban areas such as Leipzig or Rostock, primarily in hotels that accepted foreign currency (and where tourists and members of the GDR's political and military establishment were hosted). With their access to foreign and domestic sensitive information, prostitutes became strategic elements in the information gathering by the political establishment (see Falck, 1998). Hence, in spite of their official illegality, prostitutes – with their access to foreign currency and their tolerance and even use by GDR's political and executive establishment – were part of the elite and could maintain a relatively high standard of living, not only relative to average citizens in the GDR but also compared to prostitutes in the FRG relative to average citizens there: prostitution was largely seen as a *service for survival* in the FRG but as a *service for luxury* in

[1] Until 1967, the FRG and GDR employed the same framework to penal law codex.

the GDR (see Falck (1998, p. 107)).[2] With the removal of the Berlin Wall in 1989 and the re-unification of the two Germanies in 1990, a process of drastic societal and economic change started which progressed impressively since but is not finished yet.

6.2.2 Other regional aspects of sex work in Germany

Although Germany has a nationally identical legal framework for sex work, regions differ due to the varying implementation of this law, community laws, and a heterogeneity of the regional population.

In 2002, Germany introduced a new legal framework to strengthen the rights of sex workers on the national level. In general, sex workers can now legally sell sexual services and are able to settle legal disputes on service fees, rental, etc., in front of a court. At the same time, procuration or human trafficking are still illegal. In theory, with one of the most liberal laws in Europe, sex work in Germany is now accepted as a "normal" job with all the rights (health insurance, social benefits) and duties (taxes) workers in other sectors have. Nevertheless, due to the federal structure of Germany, the regions have to enforce the new law and communities are able to set their own 'ban zones' (while Munich has had such a ban zone for quite some time and sex work takes place in the outskirts of the city, the red-light district at Frankfurt on Main is very close to the main station) and taxes (the city of Cologne introduced an "enjoyment" tax with a yearly revenue of about 700.000 Euros, Stuttgart's revenues are estimated to be about 1 million Euros per annum, VG Cologne, 2007; Preiss, 2012). This has only an indirect effect on online sex workers, but a very direct one on the offline part of the market.

In general, regions in Germany differ in many aspects that have been shown to be correlated with the supply or demand of sexual services such as population structure, religion, income, job opportunities, inmoving men, ethnic networks (which are necessary to provide infrastructure for legal foreign sex workers or illegal procuration and human trafficking). E.g., Cameron & Collins (2003) argue that male participation in the market for sexual services is amongst others correlated with religion (due to the relative lack of "free" sex opportunities made available by religious women) and men that have recently moved to a new area (thereby missing well developed personal networks). Supply and demand both depend on local preferences and influence each other. Differentiated sexual desires on the demand side aggregate to a market whose demand will be fulfilled by sex workers and, vice versa, existing differenciated service offerings attract new customers. New sellers entering the market have to recognize the local market structure and rules and adapt their service offerings to local habits (Döring, 2014).

[2] Lucas (2005) records that, while the primary motive for prostitution is economic, other motives are flexibility and freedom, enjoyment of variety, available options, and job satisfaction.

To sum up, in an analysis of the determinants of the supply and demand of sexual services, paying attention to regional facets and characteristics appears important.

6.3 Data

This chapter utilizes the universe of data of the internet platform *www.gesext.de*, a digital market place (or matching site) where mainly women offer sexual services for pay[3]. For provision of their services, the platform owners charge a fee of 15 percent. The activity is legal and, in the case of platform owners, does not even count as procuration. Identities of suppliers and customers are known to the platform administrators such that contract enforcement is not an issue, similar to *ebay*, and identities can be revealed to legal (judicial and executive) authorities in case of problems (e.g., pregnancy and associated claimed alimony payments, infections with diseases associated with sexual service provision, etc.). Between January 16 in 2011 and September 9 in 2012, the platform administered 27,946 transactions offered by females – of which 11,114 were offered only as a second-price auctions in a narrow sense, 10,602 as buy-it-now at a fixed price offer, and 6,230 both as an auction in a narrow sense *and* as buy-it-now, akin to *ebay*. The nominal revenue total (net services costs gross of taxes plus administration fees) adds up to 5.59 million Euros within the considered time span.[4]

Apart from specifying a starting price for auctions, a fixed price for buy-it-now sales together with the duration and nature of the service, and some of their physical characteristics, women determine the location of the service by way of a two-digit German zip code which is known to the customer.[5] Moreover, it is known when the service offer was posted, and whether and when it was contracted. Anonymized digital identities of both suppliers and buyers and, hence, the repeated activity (sequentially or within-time) of one and the same individual under the same nickname is known. Altogether, there are 95 unique two-digit zip codes (Berlin and Munich are composed of four and two such zip codes, respectively), 76 of which are located in Germany's 10 *Old Länder* which formed the FGR prior to 1990 (West Germany) and 19 are located in the country's 6 *New Länder* which formed the GDR (East Germany) prior to 1990. Of the 27,946 transactions analysed here, 24,542 (or 88%) and 3,404 (or 12%) belonged in the Old and New Länder, respectively.

Table 6.1 summarizes information about the first and second moments of the main variables and some control variables, for the pooled sample as well as for the

[3]E.g., Hubbard & Whowell, 2008 emphasize that considering male sex workers is important for gaining broader insights into the industry. Unfortunately, the data available to us do not permit such an analysis, since male offers on the platform are scarce and (the sometimes negative) prices reveal peculiarities of that market segment in the data at hand.

[4]For the U.S. market, Cunningham & Kendall (2011b) report a significant shift from offline to online prostitution over recent years, especially, for the upper price end of the market. Hence, the lower end of the market (e.g., street prostitution) is not represented in this data set.

[5]Clearly, the location of the service offer may differ from the (unknown) residence of the provider.

Table 6.1: *Descriptive statistics*

		Total		West		East	
Variable	Nature	Mean	SD	Mean	SD	Mean	SD
Price per hour	continuous	102.3603	57.3697	102.2971	57.8617	102.8159	53.6958
Log price per hour	continuous	4.4793	0.5896	4.4780	0.5886	4.4889	0.5972
Indicator for East Germany	binary	0.1218	0.3271	0	0	1	0
Log Distance to inner german border × East Germany	continuous	0.1670	0.4769	0	0	1.3709	0.4657
Log Distance to inner german border × East Germany squared	continuous	0.2553	0.7839	0	0	2.0962	1.0892
Log Distance to inner german border	continuous	-2.9675	1.7442	-3.5692	0.6794	1.3709	0.4657
Log Distance to inner german border squared	continuous	11.8481	5.1088	13.2007	3.8125	2.0962	1.0892
Log population in same 2-digit zipcode	continuous	13.6915	0.6603	13.7590	0.3083	13.2052	1.6203
Log population in same 2-digit zipcode squared	continuous	28.2898	0.8457	28.3922	0.5850	27.5520	1.6686
Log population density in same 2-digit zipcode	continuous	6.2086	1.1066	6.3336	1.0804	5.3074	0.8436
Log population density in same 2-digit zipcode squared	continuous	39.7712	14.4334	41.2818	14.2077	28.8799	10.9496
Log GDP per capita in same 2-digit zipcode	continuous	10.3671	0.3019	10.4096	0.2956	10.0606	0.1057
Log GDP per capita in same 2-digit zipcode squared	continuous	107.5676	6.3709	108.4471	6.2636	101.2261	2.1378
Indicator for the ten largest 2-digit zipcodes	binary	0.1619	0.3684	0.1751	0.3801	0.0667	0.2495
Log number of cars in same 2-digit zipcode	continuous	13.1754	0.1283	13.1868	0.0849	13.0935	0.2751
Log number of restaurants with Michelin stars in same 2-digit zipcode	continuous	-13.0411	0.7730	-12.9937	0.7611	-13.5865	0.6961
Duration of the service offered	continuous	3.1108	8.1845	3.0948	8.2884	3.2261	7.3930
Duration of the service offered squared	continuous	76.6616	2373.6210	78.2725	2517.2490	65.0480	754.7917
Weight is provided	binary	0.3534	0.4780	0.3589	0.4797	0.3137	0.4641
Overweight	binary	0.1681	0.3740	0.1713	0.3768	0.1414	0.3486
Height is provided	binary	0.5955	0.4908	0.5968	0.4906	0.5867	0.4925
Height in cm if height is provided	discrete	166.7367	6.0477	166.6415	5.9640	167.4352	6.5889
Age is provided	binary	0.6958	0.4601	0.6938	0.4609	0.7103	0.4537
Age in years if age is provided	discrete	29.7308	8.3929	29.7076	8.3808	29.8941	8.4771
First-time supplier	binary	0.0451	0.2074	0.0475	0.2127	0.0273	0.1630
Single	binary	0.0847	0.2785	0.0854	0.2795	0.0796	0.2707
Unprotected sex offering	binary	0.0741	0.2620	0.0650	0.2465	0.1401	0.3472
History of hourwage in same 2-digit zipcode	continuous	4.4815	0.1992	4.4851	0.1947	4.4554	0.2276
History of unprotected sex offerings in same 2-digit zipcode	binary	0.0544	0.0738	0.0501	0.0719	0.0852	0.0795
Number of contracts		27,946		24,542		3,403	
Number of sellers		3,204		2,742		511	

Notes: The table presents the descriptive statistics of all variables used in this chapter for the sample as a whole, as well as for Western and Eastern Germany seperately.

East and the West, separately.[6] Looking at West (East) Germany, the average hourly wage is 102.30 (102.82) Euro, 6.50% (14.01%) of the offers include unprotected sex, the average population size in the same zip code area is $945,009$ $(543,165)$, the average population density per square kilometer in the same zip code area is 1022 (419) and average GDP per capita in the same zip code is $34,853$ $(23,537)$ Euro. 17.51% (6.61%) of the contracts are made in one of the ten largest regions, the average number of cars in the same 2-digit zip code area is 535.15 (498.21) and the average number of Michelin star restaurants in the same 2-digit zip code area is 3.55 (1.61). The average duration of an offered service is 3.09 (3.23) hours. In 35.9% (31.37%) of the auctions the seller posts her weight. Of these 17.13% (14.14%) report to be overweight. In 59.7% (58.7%) of the auctions the seller posts her height, which is on average 166.64 cm (167.43 cm). 69.4% (71.0%) provide their age and the average age is then 29.7 (29.9) years. 4.75% (2.73%) of the sellers claim to be offering for the first time, 8.5% (8.0%) to be single. Hence, except for the propensity of unprotected sexual service offerings, the nature of auctions offered (including the sellers' attributes) seem to be rather similar in formerly West and East Germany.

The data at hand contain information about the time of the posting of the service offer (an auction or a buy-it-now) as well as the concluded contract. Each offer or contract represents a cross-sectional unit so that we may suppress the time index (we still can employ time-specific determinants as contract-specific attributes) without loss of generality. We use an index $i = 1, \ldots, N = 27,946$ to denote such cross-sectional units (transactions).

6.4 Empirical analysis

In this section, we address various aspects of sex worker behavior along the lines of the pricing of and the acceptance of risk in terms of unprotected sexual services offerings. In particular, we are interested in four dimensions of the question: (i) how large the part of the variance is in the measures of behavior considered that can be explained by time, region (2-digit zip codes) and individual (sex worker) attributes, as well as the residual; (ii) to which extent the regional component can be explained by observable regional variables; (iii) to which extent there is a discontinuity at the former East-West-German border and by how much it varies with observables; and (iv) to which extent there are regional network (peer group) effects that affect the habit formation or the scope of initially supplied services. For these dimensions, we allude to the methods employed for analysis first and then present the associated results.

[6] For studies on the determinants of hourly wages of sex workers see, e.g., Cunningham & Kendall (2011b) or Moffatt & Peters (2004).

6.4.1 Variability of sex work characteristics in Germany: Geography, Time, and Individuals

Empirical approach

The variability of multi-dimensional or multi-indexed data can suitably be analysed in a basic way by means of an analysis of variance (ANOVA), a method which is designed to linearly decompose the overall variance in data into its components (for a detailed discussion see Scheffé, 1999).

For the question at stake, we may think of an outcome characterizing demand and supply of a sexual service by y_i, where $i = 1, \ldots, N$ refers to an individual transaction which is unique per time, region, and individual sex worker. However, transactions i are nested within time $t = 1, \ldots, T$, regions $r = 1, \ldots, R$, and sex workers $s = 1, \ldots, S$. Hence, we may write outcome y_i also as y_{srti} and decompose its variance by the model

$$y_{srti} = \alpha_t + \beta_r + \gamma_s + \varepsilon_i \tag{6.1}$$

where α_t is a time (month) fixed effect of which there are $T = 21$ in the data, β_r is a region (2-digit zip code) fixed effect of which there are $R = 95$ in the data, γ_s is an individual sex-worker effect of which there are $S = 3,204$ in the data, and ε_i is a remainder (residual) term which captures all effects that cannot be attributed to time, regions, or sex workers but are specific to transactions or contracts of which there are $N = 27,946$ in the data.

The model in 6.1 may now be used to decompose the variance (i.e., the demeaned average sum of squares) of y_{srti} into the components that can be attributed to time, regions, sex workers, and the residual. Such an analysis is informative about the relative importance of these components in the data.

Analysis of variance of log hourly wages charged for sexual work and the probability of offering unprotected sexual services

In this subsection, we undertake an analysis of variance on two alternative outcome variables, y_{srti}: log hourly wages for sex work as contracted in the data, and a binary indicator variable which is unity if a sex worker offers unprotected sex services and zero else. The results of the decomposition of the variance in these two variables for all $27,946$ transactions in the data based on an ANOVA as in equation 6.1 are summarized in Table 6.2.

Table 6.2 is organized in two bigger horizontal blocks for the two outcomes which contain four columns each (these are for the sum of squared deviations from the mean, the number of degrees of freedom, the F-statistic of the joint significance of the respective fixed effects in the dimensions t, r, and s,, and the p-values of the

Table 6.2: *ANOVA results*

	Log hourly wage				**Unprotected sex**			
	Partial SS	df	F	p-value	Partial SS	df	F	p-value
Model	6,307.01	3,315	1.90	0.00	1,527.34	3,315	29.02	0.00
Month	7.33	20	0.37	0.00	0.94	20	2.95	0.00
Region	43.99	94	0.47	0.00	4.06	94	2.72	0.00
Sex worker	5,554.19	3,201	1.74	0.00	1,335.56	3,201	26.28	0.00
Residual	3,408.60	24,630	0.14		391.04	24,630		
Total	9,715.62	27,945	0.35		1,918.38	27,945		

Notes: The table presents the results of the ANOVA variance decomposition of the two variables *log hourly wages* and *unprotected sex*. The sample contains 27,946 contracts, sold within 21 months, in 95 regions and by 3,204 sex workers.

latter) and in three bigger horizontal blocks for the total sum of squared deviations from the mean, the residual, and the model accruing to $\alpha_t + \beta_r + \gamma_s$ in equation 6.1.

The results may be summarized as follows. First, the residual makes up about one-third of the variation in the data on log hourly wages and about one-fifth of the variation in the probability of unprotected sex offerings. In either dimension, the variance within the three dimensions of the model (time, region, and sex worker) is dominated by sex worker-specific effects within time and regions. However, the regional dimension is much more important than time for either outcome. It is also interesting to see that the relative importance of the regional dimension for the offering of unprotected sexual services is more important in relative terms than for log hourly wages (to see this, compare $4.06/1,918.38 = 0,2\%$ with $43.99/9,715.62 = 0,4\%$).

In what follows, we will shed further light on the role of regional characteristics with an emphasis on urban agglomerations on the factors behind regional factors as a determinant of hourly wages and unprotected service offerings by sex workers.

6.4.2 Explaining the geographical component of sex work characteristics: Evidence of social disorganization of places

Empirical approach

In this subsection, we are interested in observable measures behind the region fixed effects, β_r. Such an assessment may be based either on a regression of parameter estimates of these fixed effects, say, $\hat{\beta}_r$, or of averaged outcome per region, say, $\overline{y}_r$. An advantage of such an analysis is that we may gain insights into specific characteristics of places (zip codes) as drivers of certain aspects of sexual work outcome. Denoting observable zip-code characteristics by x_r (including a constant), what we have in mind is estimating a model of the type

$$\overline{y}_r = \delta x_r + \varepsilon_r \tag{6.2}$$

where δ is a conformable parameter vector and ε_r are all residual (not observed or observable), region- or place-specific factors. Hence we first calculate the averaged outcome $\overline{y}_r$ of log hourly wages and the probability of offering unprotected services. Then a between estimation is used to regress these two new variables against several explanatory variables.

The role of observable place-specific factors for sexual worker outcome

Table 6.3 summarizes the results for two regressions, one for the place-specific component of log hourly wages and one for that of unprotected sex services.

Table 6.3: *Regression results: Geographical component of sex work characteristics*

Dependant variable	Log hourly wages	Offering service unprotected
Indicator for the ten largest 2-digit zipcodes	-0.1634*	0.0131
	(0.0896)	(0.0291)
Log number of cars in same 2-digit zipcode	0.1263	0.0328
	(0.0874)	(0.0284)
Log number of Michelin star restaurants in same 2-digit zipcode	0.0306	-0.0030
	(0.0359)	(0.0117)
Log population in same 2-digit zipcode	-0.1372*	-0.0038
	(0.0735)	(0.0239)
Log population in same 2-digit zipcode squared	0.0940	0.0007
	(0.0569)	(0.0185)
Log population density in same 2-digit zipcode	0.1947	-0.1044
	(0.3006)	(0.0978)
Log population density in same 2-digit zipcode squared	-0.0139	0.0079
	(0.0247)	(0.0080)
Log GDP per capita in same 2-digit zipcode	-3.9631	0.7102
	(5.2985)	(1.7234)
Log GDP per capita in same 2-digit zipcode squared	0.1984	-0.0325
	(0.2547)	(0.0828)
Number of observations	95	95
R^2	0.2074	0.0377

Notes: The table presents regression results of the geographical components of the two variables *log hourly wages* and *offering services unprotected*. Usual (observed information matrix) standard errors are given in brackets below the coefficients and ***/**/* denote significance at the 1%/5%/10% level.

In every one of the regressions we use the following observables: log population and log population squared; log population density and log population density squared; log per-capita income (GDP) and log per-capita income squared; the log number of passenger cars; the log (one-plus) the number of Michelin-star restaurants; a binary indicator variable which is unity for the 10 largest 2-digit zip code regions of Germany, and a constant. We ran further specifications which also included information on the number of hospital beds, five-star hotels, the number of hotel visits and nights booked, the number of immigrants, of males commuting in, the Theil index about income inequality in the region, the share of people with more than 125,000 Euros of annual income, the average price per square meter of land sold, the number of internet domains, the share of inhabitants with completed secondary school education, the share of inhabitants in day care and kindergarten, the share of street surface or of recreational surface in a place's total surface, and an indicator for formerly East-German regions. At the level of 95 regions only, these characteristics together did not contribute a significant share of explanatory power to the factors included in the specifications in Table 3.

In any case, the explanatory power of the specification is much higher for the continuous log hourly wage per service (with an R^2 of 0.207) than for the binary unprotected services (with an R^2 of 0.038). With the latter outcome, none of the included regressors is significant even at a significance level of 15%. With log hourly wages, being in one of the ten largest cities contributes the lion's share in explanatory power, and population size (through linear and squared terms) also contribute a decent share (the squared term having a level of statistical significance of 10.2%).

Hence, big cities stand out – with lower prices of contracted services per hour – and other amenities are apparently too collinear with the large-city indicator and do not have a discernible impact.

6.4.3 Discontinuities in sex work characteristics at the former East-West-German border

Empirical approach

To estimate the impact (long shadow) of the former Iron Curtain on behavior of sexual service providers, we may utilize a regression discontinuity design (RDD). Such a design is helpful for avoiding an endogeneity bias in pricing, since the location of the former border between East and West Germany can not be influenced by sex workers nowadays (see Imbens & Lemieux, 2008, and Lee & Lemieux, 2010, for a general discussion of RDDs and Egger & Lassmann, 2014, for a discussion of spatial RDDs). The general idea of a spatial RDD is to compare two spatial observations' outcome which are virtually identical and just located on different sides of the same border (e.g., being part of former East Germany or not). Clearly, the Old and the New Länder differ in many regards such that an unconditional mean comparison of variables in the two aggregates may be misleading due to omitted confounding factors, and even conditioning on observable characteristics may not allow for a complete eradication of that problem. Since the border between the Old and the New Länder entails a sharp boundary, we can use an RDD and take the limit of the distance to the internal border in order to avoid differences between the West and the

East beyond the historical political, societal, and economic separation (e.g., related to the relative importance of urban relative to rural areas, the performance of local labor markets, the availability of infrastructure, etc.) to the largest possible extent. This RDD identifies $LATE_{east}$, the local average treatment effect of auctions in the East (relative to the West) at the internal East-West border, the historical location of the Iron Curtain, as outlined in Apppendix A.3.1.

Estimating the long shadow of the Iron Curtain in general and in Berlin

The RDD results are summarized in Table 6.4. The table is organized horizontally in two big blocs: one for log hourly wages and one for binary unprotected services offerings. Within each bloc, one column on the left summarizes the findings for a parsimonious model, assuming homogeneity of $LATE_{east}$, and one column summarizes the findings when permitting $LATE_{east}$ to be different for Berlin and to generally vary with log population. Each RDD regression includes several control variables which pertain to the nature of the service (see Table 6.1) as well as a quadratic control function based on log distance to the former Iron Curtain which is estimated separately for West and East. In fact, we have estimated polynomial control functions for first-order, second-order, etc., up to fifth-order polynomials. However, the second-order (quadratic) polynomial function is selected by the Akaike Information Criterion as the preferred specification. All reported standard errors are clustered at three levels: the 2-digit zip codes, East versus West, and the month of the offer.

Looking first at the zip-code and seller control variables results are not surprising. Contracted hourly wages rise with a higher population density and also with a higher GDP per capita and higher historical hourly prices in the respective zip code. Sellers seem to achieve lower hourly wages the more hours the "date" will last and if they claim to be single. Higher prices can be reached if they are not overweight, provide their age, are young and claim to offer a service for the first time. The parameter on the indicator variable of interest – $east_i$ – is 0.2724, respectively, significant at the 1% level. Hence, conditional on the factors mentioned above, the result suggests that prostitutes in Germany's East, the *New Länder* are still able to charge significantly higher (27%) *relative* service premia than their counterparts in the Western, *Old Länder*. Figure 6.1a nicely illustrates this strong discontinuity of the log hourly wages at the former Iron Curtain. Observations are categorized by the log-distance towards the border (with negative values in the West) and each distance category is represented by a dot. The size of the dot corresponds to the amount of observations in the respective category and log hourly wages are depicted on the ordinate. The polynomial function jumps at the former border and the same is true if the control variables are included by looking at the residuals instead of hourly wages (Figure 6.1b).

The second column of Table 6.4 suggests that the $LATE_{east}$ is indeed smaller

Table 6.4: *Regression results: East vs. West*

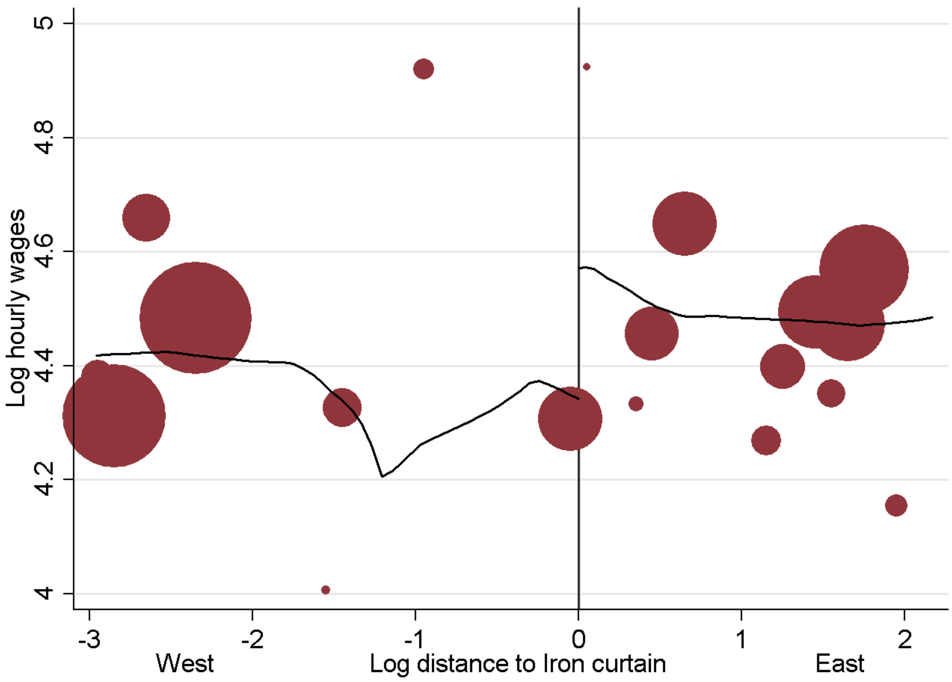
Dependent variables	**Log hourly wages**		**Unprotected service offering**	
	(1)	**(2)**	**(3)**	**(4)**
Indicator for East Germany	0.2724***	0.3068***	0.1379***	0.0209
	(0.0327)	(0.0290)	(0.0169)	(0.0239)
Log Distance to inner german border × East G.	-0.3403***	-0.4534***	-0.0866***	0.1144***
	(0.0588)	(0.0323)	(0.0280)	(0.0358)
Log Distance to inner german border^2× East G.	0.1227***	0.1820***	0.0008	-0.0508**
	(0.0217)	(0.0141)	(0.0128)	(0.0215)
Log Distance to inner german border	0.0120	0.0179	0.0236***	0.0054
	(0.0132)	(0.0174)	(0.0036)	(0.0097)
Log Distance to inner german border^2	0.0036	0.0061*	0.0082***	0.0041*
	(0.0026)	(0.0033)	(0.0009)	(0.0025)
Log population size		-0.0597***		0.0215***
		(0.0101)		(0.0044)
Log population size × East Germany		0.0706***		-0.0528***
		(0.0129)		(0.0081)
Indicator for Berlin		0.1719***		0.2038***
		(0.0395)		(0.0253)
Indicator for Berlin × East Germany		-0.2163***		-0.2255***
		(0.0585)		(0.0390)
Log population density in same 2-digit zipcode	0.1220*	0.1036	0.0518*	0.0573**
	(0.0696)	(0.0806)	(0.0276)	(0.0257)
Log population density in same 2-digit zipcode^2	-0.0097*	-0.0079	-0.0051**	-0.0056***
	(0.0055)	(0.0065)	(0.0021)	(0.0015)
Log GDP per capita in same 2-digit zipcode	0.1272***	0.4107***	0.0533***	0.0078
	(0.0248)	(0.0103)	(0.0084)	(0.0187)
Duration of the service offered	-0.0551***	-0.0552***	0.0029**	0.0033***
	(0.0037)	(0.0055)	(0.0015)	(0.0006)
Duration of the service offered^2	0.0001***	0.0001***	0.0000	0.0000**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Weight is provided	0.0093	0.0102	-0.0137	-0.0263*
	(0.0078)	(0.0103)	(0.0145)	(0.0145)
Overweight	-0.2363***	-0.2372***	-0.0070	-0.0348***
	(0.0153)	(0.0158)	(0.0069)	(0.0043)
Height is provided	-0.9921***	-1.0378***	-0.0632	0.0232
	(0.2472)	(0.2408)	(0.3541)	(0.2695)
Height in cm if height is provided	0.0059***	0.0061***	0.0000***	0.0000***
	(0.0016)	(0.0000)	(0.0000)	(0.0000)
Age is provided	0.4054***	0.4107***	0.0020	0.0023
	(0.0104)	(0.0103)	(0.0250)	(0.0357)
Age in years if age is provided	-0.0123***	-0.0124***	0.0007***	0.0009***
	(0.0008)	(0.0008)	(0.0002)	(0.0004)
First-time supplier	0.1659***	0.1647***	-0.0086	-0.0084
	(0.0231)	(0.0240)	(0.0091)	(0.0176)
Single	-0.1012***	-0.1047***	0.0099	0.0185
	(0.0325)	(0.0301)	(0.0127)	(0.0150)
History of hourwage in same 2-digit zipcode	0.3942***	0.4045***	0.0240	0.1011***
	(0.0482)	(0.0417)	(0.0189)	(0.0255)
AIC	1.014	1.009	0.169	0.615
BIC	-23891.1	-23883.4	-124704.2	-25212.5
Number of observations	27811	27811	27811	27811

Notes: The table presents the results of the RDD at the former inner German border. The first two columns look at log hourly wages, columns (3) and (4) look at the probability of offering unprotected services. Standard errors are clustered at three levels: 2-digit zip codes, East versus West, as well as month of the offer and are given in brackets below the coefficients. ***/**/* denote significance at the 1%/5%/10% level.

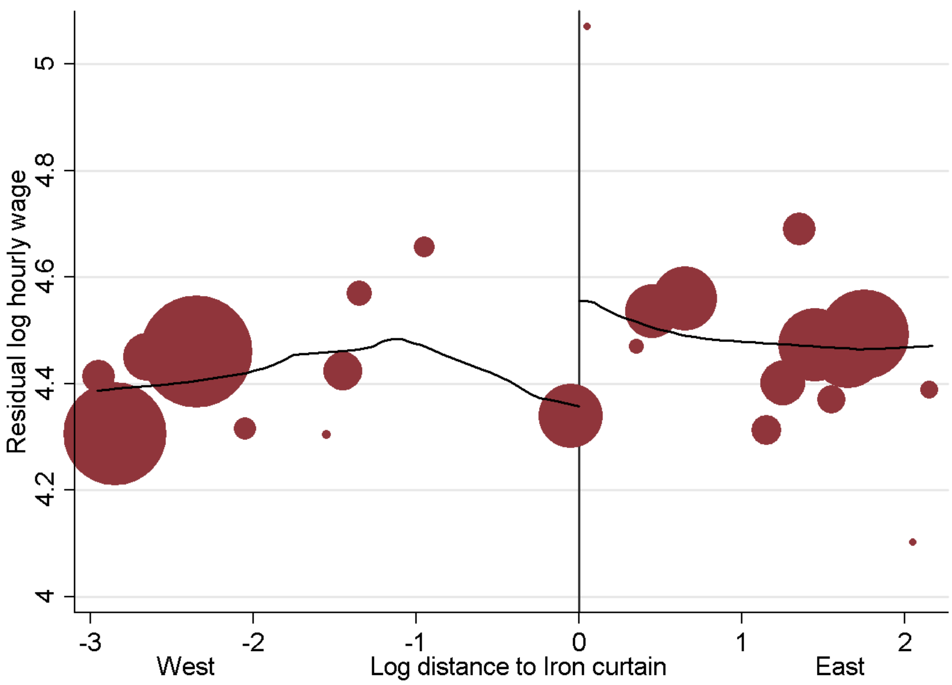
Figure 6.1: *Discontinuity at the former Iron Curtain*

The graph illustrates the discontinuity of (residual) log hourly wages at the former German border. The ordinate reflects the average (residual) log hourly wages in the respective distance-to-border category. The size of the dots corresponds to the number of observations in the respective distance category. *Source: Own representation.*

(a) Log hourly wages



(b) Residual log hourly wages



in absolute value for Berlin – we interpret this as evidence of, e.g., greater competition in the capital as well as of a greater exposure of Berlin to western standards historically – and that it is otherwise bigger with population. For an interpretation of the $LATE_{east}$ we have to recognize that the value of log population in east Berlin takes on a value of 12.026. Hence, the overall $LATE_{east}$ for east Berlin is $0.307 + 0.071 \cdot 12.026 - 0.216 \approx 0.945$. In other words, the log hourly wage markup for sexual services in Berlin is about $100exp(0.945) - 100 \approx 157\%$. Had this been just another city of the same size, it would be estimated at $100exp(0.307 + 0.071 \cdot 12.026) - 100 \approx 219\%$.

The results for the (linear) probability of unprotected sexual service offerings suggest that the effect on prices is not accompanied with one of differences in risky supply behavior: neither the assumed-homogeneous nor the heterogeneous $LATE_{east}$ are significant at customary levels for unprotected sexual services supply (Table 6.4, columns 3 and 4).

6.4.4 Habit formation in versus location of specific sexual work types in space

Empirical approach

In this sub-section, our interest is focused on regionally bound so-called contagion, network, or peer-group effects which lead to an impact of other sex workers' hourly wage outcomes or unprotected sex services offerings on a given sex worker's outcome in the same region. In particular, we are interested in how such contagious effects differ between ones on first-time users of the platform versus continuers or stayers on the platform. The latter suggests an adjustment of the habit of existing suppliers while the former is indicative of a combined effect on location and the initial scope of services supplied.

For the empirical analysis, such an approach asks for regression models of the form

$$y_{srti} = \overline{y}_{sr(t-1)i}\lambda_{first} + x_{srti}\delta_{first} + \varepsilon_i \qquad \text{if } first_{srti} = 1,$$

$$y_{srti} = \overline{y}_{sr(t-1)i}\lambda_{continued} + x_{srti}\delta_{continued} + \varepsilon_i, \quad \text{if } first_{srti} = 0. \tag{6.3}$$

where $\overline{y}_{sr(t-1)i}$ is a weighted average of outcomes of all sex workers recorded in the same region as sex worker s and transaction i but at time t-1. In the discussion of the results, we will primarily focus on estimates of λ_{first} versus $\lambda_{continued}$.

Estimates of peer-group effects on habit formation and first services postings

The results of the analysis outlined in equation 6.3 are summarized in Table 6.5 for the two outcomes of interest: log hourly wage and the probability of unptotected sex offering. The table is organized horizontally in four columns, two for each outcome. The first column for any outcome focuses on outcome of first-time suppliers on the platform while the second column focuses on continuing suppliers' outcome. Each regression includes time (month) and region (2-digit zip-code) fixed effects.

Table 6.5: *Regression results: Habit formation*

Dependent variables	**Log hourly wages**		**Service offered unprotected**	
	First time (1A)	Continuing (2A)	First time (1B)	Continuing (2B)
History of hourwage in same 2-digit zipcode	0.2710**	0.3551***		
	(0.1214)	(0.0472)		
History of offering unprotected sex in same 2-digit z.			1.3500***	0.7410***
			(0.4164)	(0.1313)
Duration of the service offered	-0.0911***	-0.0561***	0.0024	0.0027***
	(0.0068)	(0.0049)	(0.0038)	(0.0007)
Duration of the service offered squared	0.0005***	0.0001***	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Weight is provided	0.0348	0.0156	0.0497	-0.0313
	(0.0279)	(0.0355)	(0.0353)	(0.0197)
Overweight	-0.3349***	-0.2448***	0.0030	-0.0247**
	(0.0603)	(0.0399)	(0.0254)	(0.0112)
Height is provided	-1.2316*	-0.9167**	0.7924**	0.0403
	(0.6620)	(0.4392)	(0.3824)	(0.2367)
Height in cm if height is provided	0.0074*	0.0054**	-0.0051**	-0.0004
	(0.0040)	(0.0026)	(0.0023)	(0.0014)
Age is provided	0.4818***	0.4295***	0.1265**	-0.0064
	(0.1023)	(0.0621)	(0.0641)	(0.0448)
Age in years if age is provided	-0.0192***	-0.0132***	-0.0024	0.0012
	(0.0032)	(0.0018)	(0.0021)	(0.0014)
Single	0.0374	-0.1065*	0.0638	0.0177
	(0.0532)	(0.0564)	(0.0476)	(0.0290)
Region Fixed effects	Yes	Yes	Yes	Yes
Time Fixed effects	Yes	Yes	Yes	Yes
Number of observations	1250	26561	1250	26561
R^2	0.6001	0.4060	0.2680	0.1288

Notes: The table presents the results of the habit formation regression and compares first time users to non-first time users. The first two columns look at log hourly wages and columns (3) and (4) look at the probability of offering unprotected services. Usual (observed information matrix) standard errors are given in brackets below the coefficients and ***/**/* denote significance at the 1%/5%/10% level.

The results in the table suggest the following conclusions. First, irrespective of the outcome (log hourly wage versus unprotected service supply) and the type of provider considered (first-time versus continuing supplier) considered, contagion matters and sex workers are affected by other sex workers' prior outcomes in the same place (2-digit zip code area). However, while spillovers or peer-group effects are stronger for the habit formation of continuers rather than the pricing of first suppliers regarding the log hourly wage, the opposite is true for the offering of unprotected services. Sex worker-specific control variables affect, in particular, the hourly wage in an expected way: overweight and older sex workers have to accept a lower price, and sex workers offer longer durations of services at lower prices.

6.5 Conclusions

It is well documented that people's preferences and attitudes display a spatial pattern and are influenced by the local political regime and institutions. Political and institutional perturbations lead to only a sluggish adjustment of those preferences and attitudes. After the reunification, sluggish adjustment of preferences and attitudes is documented among the people in the former East and West.

At the time when East and West Germany were politically separated after World War II, the classification of, the access to, and the attitudes towards normal versus luxury goods diverged since the War in the two Germanies. Falck (1998) documented for the former East Germany that prostitution – while only tolerated and actually illegal for most of the time – was an activity which generated relatively much higher income and provided access to much more luxury in the East than in the West of Germany. This chapter documents that sexual services for pay nowadays still generate higher relative income to providers in what was East Germany than in what was West Germany. Moreover, we demonstrate that pattern is even more pronounced for larger cities such as Berlin. Finally, it is partly driven by network, peer-group, and other spillover effects which leads to a clustering in both the nature (offering unprotected sex) as well as the pricing of services by sex workers.

One conclusion which might be distilled for policy makers from this evidence is that the application of risky practices of sex work, due to their clustered emergence, might be addressed and prevented at relatively low cost when considering their spatial pattern.

Bibliography

Albrecht, G., Groenemeyer, A., & Stallberg, F. (1999). *Handbuch soziale Probleme*. Springer VS.

Alesina, A., & Fuchs-Schündeln, N. (2007). Good-bye Lenin (or not?): The effect of communism on people's references. *American Economic Review*, *97*(4), 1507–1528.

Alfred, J. (2011). *World History Encyclopedia [21 volumes]*. vol. 2. ABC-CLIO, Santa Barbara, CA.

Allmers, S., & Maennig, W. (2009). Economic impacts of the FIFA soccer World Cups in France 1998, Germany 2006, and outlook for South Africa 2010. *Eastern Economic Journal*, *35*(4), 500–519.

Arunachalam, R., & Shah, M. (2008). Prostitutes and brides? *American Economic Review*, *98*(2).

Arunachalam, R., & Shah, M. (2012). The prostitute's allure: The return to beauty in commercial sex work. *The BE Journal of Economic Analysis & Policy*, *12*(1).

Arunachalam, R., & Shah, M. (2013). Compensated for life: sex work and disease risk. *The Journal of Human Resources*, *48*(2), 345–369.

Baade, R., & Matheson, V. (2004). The quest for the Cup: Assessing the economic impact of the World Cup. *Regional Studies*, *38*(4), 343–354.

Badinger, H., & Egger, P. (2014). Fixed effects and random effects estimation of higher-order spatial autoregressive models with spatial autoregressive and heteroskedastic disturbances. *CESifo Working Paper*, *4847*.

Barnett, L. (2013). Germany a 'sex tourism hotspot', *The Telegraph, 15.11.2013. Last access on 20.02.2015.*
URL `http://www.telegraph.co.uk/news/worldnews/europe/germany/10452246/Germany-a-sex-tourism-hotspot.html`

Barro, R., & Sala-i-Martin, X. (1991). Convergence across states and regions. *Brookings Papers on Economic Activity*, *1991*(1), 107–182.

Barro, R. J., & Sala-i-Martin, X. (1995). *Economic growth*. McGraw-Hill advanced series in economics. McGraw-Hill, New York, NY.

Bathelt, H., & Spigel, B. (2012). The spatial economy of North American trade fairs. *Canadian Geographer / Le Géographe canadien*, *56*(1), 18–38.

Bauernschuster, S., & Rainer, H. (2010). Political regimes and the family: How sex-role attitudes continue to differ in reunified Germany. *Journal of Population Economics*, *25*(1), 5–27.

Becker, G. S., Murphy, K. M., & Grossman, M. (2004). The economic theory of illegal goods: The case of drugs. NBER Working Papers 10976, National Bureau of Economic Research, Inc.

Becker, S. O., & Hainz, C. (2011). The empire is dead, long live the empire! Long-run persistence of trust and corruption in the bureaucracy. *IZA*, (5584).

Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review*, (pp. 67–72).

Berument, M. H., & Ceylan, N. B. (2012). Effects of soccer on stock markets: The return-volatility relationship. *The Social Science Journal*, *49*(3), 368–374.

Bliss, C. (1988). A theory of retail pricing. *Journal of Industrial Economics*, *36*(4), 375 – 391.

Brenke, K., & Grabka, M. (2011). Schwache Lohnentwicklung im letzten Jahrzehnt. *DIW Wochenbericht*, (2), 3–16.

Bullough, V. L., & Bullough, B. (1987). *Women and Prostitution: A social history. New Concepts in Human Sexuality*. Prometheus Books, Amherst, NY.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press, Cambridge, UK.

Cameron, S. (2002). *The Economics of sin: Rational choice or no choice at all?*. Edward Elgar Publishing, Cheltenham, UK.

Cameron, S. (2004). Space, risk and opportunity: The evolution of paid sex markets. *Urban Studies*, *41*(9), 1643–1657.

Cameron, S., & Collins, A. (2003). Estimates of a model of male participation in the market for female heterosexual prostitution services. *European Journal of Law and Economics*, *16*(3), 271–288.

Carree, M., & Klomp, L. (1997). Testing the convergence hypothesis: A comment. *Review of Economics and Statistics*, *79*(4), 683–686.

Cassel, G. (1918). Abnormal deviations in international exchanges. *The Economic Journal*, (pp. 413–415).

Cawley, J. (2014). The impact of obesity on wages. *The Journal of Human Resources*, *39*(2), 451–474.

Cecchini, P., Catinat, M., & Jacquemin, A. (1988). *The European Challenge, 1992: the Benefits of a Single Market*. Gower Publishing Co., Brookfield, VT.

Chamberlin, E. H. (1933). *Theory of Monopolistic Competition*. Harvard University Press, Cambridge, MA.

Chang, H.-H., & Weng, Y. (2012). What is more important for prostitute price? Physical appearance or risky sex behavior? *Economics Letters*, *117*(2), 480–483.

Cho, S.-Y., Dreher, A., & Neumayer, E. (2012). Does legalized prostitution increase human trafficking? *World Development*, *41*, 67–82.

Collins, A., & Judge, G. (2008). Client participation in paid sex markets under alternative regulatory regimes. *International Review of Law and Economics*, *28*(4), 294–301.

Cumby, R. E. (1996). Forecasting exchange rates and relative prices with the hamburger standard: Is what you want what you get with McParity? *NBER Working Paper* 5675, National Bureau of Economic Research.

Cunningham, S. (2011). Monopolistic competition with product differentiation in prostitution markets: explaining the returns to BMI, 2.12.2011, Last access on 20.02.2015. *Working paper*.
URL `http://econfaculty.gmu.edu/pboettke/workshop/Fall2011/Obesity_2.0.pdf`

Cunningham, S., & Kendall, T. D. (2010a). Moonlighting: Skill premia in commercialized sex markets, April, 2010, Last access on 20.02.2015 *Working paper*.
URL `http://ssrn.com/abstract=1583510`

Cunningham, S., & Kendall, T. D. (2010b). Risk behaviours among internet-facilitated sex workers: Evidence from two new datasets. *Sexually Transmitted Infections*, *86 Suppl 3*(Suppl 3), iii100–105.

Cunningham, S., & Kendall, T. D. (2011a). Men in transit and prostitution: Using political conventions as a natural experiment. *The BE Journal of Economic Analysis & Policy*, *11*(1), Article 30.

Cunningham, S., & Kendall, T. D. (2011b). Prostitution 2.0: The changing face of sex work. *Journal of Urban Economics*, *69*(3), 273–287.

de la Torre, A., Havenner, A., Adams, K., & Ng, J. (2010). Premium sex: Factors influencing the negotiated price of unprotected sex by female sex workers in Mexico. *Journal of Applied Economics*, *13*(1), 67–90.

Della Giusta, M., Di Tommaso, M. L., Shima, I., & Strom, S. (2009a). What money buys: Clients of street sex workers in the US. *Applied Economics*, *41*(18), 2261–2277.

Della Giusta, M., Di Tommaso, M. L., & Strom, S. (2009b). Who is watching? The market for prostitution services. *Journal of Population Economics*, *22*(2), 501–516.

Della Giusta, M., Tommaso, M. L., & Strom, S. (2008). *Sex markets: A Denied Industry*. Routledge, London.

Destatis (2014). Genesis online.
URL `https://www-genesis.destatis.de/genesis/online`

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2011). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, *79*(2), 645–677.

Döring, N. (2014). Prostitution in Deutschland: Eckdaten und Veränderungen durch das Internet. *Zeitschrift für Sexualforschung*, *27*(02), 99–137.

Dreger, C., Kholodilin, K., Lommatzsch, K., Slačálek, J., & Wozniak, P. (2008). Price convergence in an enlarged internal market. *Eastern European Economics*, *46*(5), 57–68.

Edlund, L., Engelberg, J., & Parsons, C. A. (2009). The wages of sin, June 4, 2009, *Working Paper*.

Edlund, L., & Korn, E. (2002). A theory of prostitution. *Journal of Political Economy*, *110*(1), 181–214.

Edmans, A., Garcia, D., & Norli, O. (2007). Sports sentiment and stock returns. *The Journal of Finance*, *LXII*(4), 1967–1998.

Egger, P., Gruber, S., & Pfaffermayr, M. (2009). The law of one price: A new econometric approach and evidence from disaggregated data. *Applied Economics*, *41*(26), 3345–3357.

Egger, P., & Lindenblatt, A. (2015a). Endogenous risk taking and physical appearance of sex workers. *The European Journal of Health Economics*.

Egger, P., & Lindenblatt, A. (2015b). Fun and games: How soccer games and conventions raise the demand for paid sex. *Mimeo*.

Egger, P., & Lindenblatt, A. (2015c). The long shadow of the iron curtain for female sex workers in german cities: Border effects and regional differences. *Mimeo*.

Egger, P., & Pfaffermayr, M. (2009). On testing conditional sigma-convergence. *Oxford Bulletin of Economics and Statistics*, *71*(4), 453–473.

Egger, P. H., & Lassmann, A. (2014). Cultural integration and export variety overlap. *CesIfo Working Paper*, *4800*.

Eigendorf, J., Meyer, S., Nagel, L.-M., & Neller, M. (2013). Augsburg mit höchster Dichte von Prostituierten, *Die Welt, 13.11.2013. Last access on 20.02.2015.* URL `http://www.welt.de/politik/deutschland/article121480296/Augsburg-mit-hoechster-Dichte-von-Prostituierten.html`

European Commission (1990). One market, one money: An evaluation of the potential benefits and costs of forming an economic and monetary union. *European Economy*, (44).

Eurostat (2007). *Eurostat-OECD Methodological Manual on Purchasing Power Parities*. Statistics (Organisation for Economic Co-operation and Development). OECD Publishing, Paris.

Eurostat (2010). Price levels for food, beverages and tobacco across the European market differ significantly. *Statistics in Focus*, (30).

Falck, U. (1998). *VEB Bordell: Geschichte der Prostitution in der DDR*. Ch. Links Verlag, Berlin, DE.

Feddersen, A., & Maennig, W. (2012). Sectoral labour market effects of the 2006 FIFA World Cup. *Labour Economics*, *19*(6), 860–869.

Feige, E. L. (1990). Defining and estimating underground and informal economies: The new institutional economics approach. *World Development*, *18*(7), 989 – 1002.

Ferris, S. P., Koo, S., Park, K., & Yi, D. T. (2012). Gold, silver, bronze or tin? The short and long term effects of mega sporting events. *KAIST Business School Working Paper*.

Fischer, C. (2012). Price convergence in the EMU? Evidence from micro data. *European Economic Review*, *56*(4), 757 – 776.

Fisher, R. (1949). A biological assay of tuberculins. *Biometrics*, *5*(4), 300–316.

Fletcher, J. M. (2009). Beauty vs. brains: Early labor market outcomes of high school graduates. *Economics Letters*, *105*(3), 321–325.

Friedman, M. (1988). *Market mechanisms and central economic planning. In Ideas, Their Origins, and Their Consequences*. American Enterprise Institute for Public Policy Research, Washington, DC.

Funke, K., & Koske, I. (2008). Does the law of one price hold within the EU? A panel analysis. *International Advances in Economic Research*, *14*(1), 11–24.

Gertler, P., & Shah, M. (2011). Sex work and infection: What's law enforcement got to do with it? *Journal of Law and Economics*, *54*(4), 811–840.

Gertler, P., Shah, M., & Bertozzi, S. (2005). Risky business: The market for unprotected commercial sex. *Journal of Political Economy*, *113*(3), 518–550.

Glazebrook, A., & Henry, M. (2011). *Greek Prostitutes in the Ancient Mediterranean, 800 BCE–200 CE*. Wisconsin Studies in Classics. University of Wisconsin Press, Madison, WI.

Goldberg, P. K., & Verboven, F. (2004). Cross-country price dispersion in the euro Era: A case study of the European car market. *Economic Policy*, (40), 483 – 521.

Green, A. I. (2008). The social organization of desire: The sexual fields approach. *Sociological Theory*, *26*(1), 25–50.

Green, A. I. (2011). Playing the (sexual) field: The interactional basis of systems of sexual stratification. *Social Psychology Quarterly*, *74*(3), 244–266.

Hagn, F., & Maennig, W. (2009). Large sport events and unemployment: The case of the 2006 soccer World Cup in Germany. *Applied Economics*, *41*(25), 3295–3302.

Hamermesh, D., & Biddle, J. E. (1994). Beauty and the labor market. *The American Economic Review*, *84*(5), 1174–1194.

Haskel, J., & Wolf, H. (2001). The law of one price – a case study. *Scandinavian Journal of Economics*, *103*(4), 545 – 558.

Heckman, J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, *46*(4), 931–959.

Heckman, J. J., & Vytlacil, E. J. (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(8), 4730–4734.

Heineck, G., & Süssmuth, B. (2010). A different look at Lenins legacy: Trust, risk, fairness and cooperativeness in the two Germanies. *CESifo Working Paper no. 3199*.

Hubbard, P. (2001). Sex zones: Intimacy, citizenship and public space. *Sexualities*, *4*(1), 51–71.

Hubbard, P., & Sanders, T. (2003). Making space for sex work: Female street prostitution and the production of urban space. *International Journal of Urban and Regional Research*, *27*(3), 75–89.

Hubbard, P., & Whowell, M. (2008). Revisiting the red light district: Still neglected, immoral and marginal? *Geoforum*, *39*(5), 1743–1755.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635.

Jevons, W. (1871). *The Theory of Political Economy*. Macmillan and Co., London, UK.

Johnston, D. W. (2010). Physical appearance and wages: Do blondes have more fun? *Economics Letters*, *108*(1), 10–12.

Kavemann, B., & Steffan, E. (2013). Zehn Jahre Prostitutionsgesetz und die Kontroverse um die Auswirkungen. *Aus Politik und Zeitgeschichte*, *9*, 9–15.

Kubrin, C. E., & Weitzer, R. (2003). New directions in social disorganization theory. *Journal of Research in Crime and Delinquency*, *40*(4), 374–402.

Kuersteiner, G. M., & Prucha, I. R. (2013). Limit theory for panel data models with cross sectional dependence and sequential exogeneity. *Journal of Econometrics*, *174*(2), 107–126.

Lamla, M. J., Straub, M., & Girsberger, E. M. (2012). On the economic impact of international sport event - Microevidence from survey data at the EURO 2008. *KOF Working Papers*, (317).

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in Economics. *Journal of Economic Literature*, *48*(June), 281–355.

Lindenblatt, A., & Feuerstein, S. (2015). Price convergence after the eastern enlargement of the EU: Evidence from retail food prices. *European Review Of Agricultural Economics*.

Lucas, A. (2005). The work of sex work: Elite prostitutes' vocational orientations and experiences. *Deviant Behavior*, *26*(6), 513–546.

MacDonald, R., & Ricci, L. A. (2005). The real exchange rate and the Balassa–Samuelson effect: The role of the distribution sector. *Pacific Economic Review*, *10*(1), 29–48.

Maddala, G. S. (1983). *Limited-dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge, UK.

Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review*, *15196*(1), 222–235.

Moffatt, P. G., & Peters, S. a. (2004). Pricing personal services: An empirical study of earnings in the UK prostitution industry. *Scottish Journal of Political Economy*, *51*(5), 675–690.

Nagel, L.-M. (2013). Prostitution - hier noch mehr Zahlen. *Die Welt*, 03.11.2013. Last access on 20.02.2015.
URL `http://investigativ.welt.de/2013/11/03/black-box-prostitution/`

Nelder, A. J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society*, *135*(3), 370–384.

Niehans, J. (1994). *A History of Economic Theory: Classic Contributions, 1720-1980*. Johns Hopkins University Press (Baltimore, MD.

Oppermann, M. (1998). Sex tourism and prostitution: Aspects of leisure, recreation and work. *Cognizant Communication Corporation, Putnam Valley, NY*.

Parsley, D., & Wei, S.-J. (2008). In search of a euro effect: Big lessons from a Big Mac meal? *Journal of International Money and Finance*, *27*(2), 260 – 276.

Penzkofer, H. (2008). Messen und Veranstaltungen der Messe München lösen bundesweit jährlich einen Umsatz von 2,17 Mrd. Euro aus. *ifo Schnelldienst*, *61*(10), 38–43.

Plankey, M. W., Stevens, J., Flegal, K. M., & Rust, P. F. (1997). Prediction equations do not eliminate systematic error in self-reported body mass index. *Obesity Research*, *5*(4), 308–14.

Preiss, G. (2012). Sexsteuer: Käuflicher Sex füllt die klammen Kassen, *Stuttgarter Nachrichten*, 25.11.2012. Last access on 20.02.2015.
URL `http://www.stuttgarter-nachrichten.de/inhalt.sexsteuer-kaeuflicher-sex-fuellt-die-klammen-kassen.fd0ab5e1-606f-4e35-99fb-3d65f5ad8e31.html`

Priks, M. (2010). Does frustration lead to violence ? Evidence from the Swedish hooligan scene. *Kyklos*, *63*(3), 450–460.

Quah, D. T. (1996). Empirics for economic growth and convergence. *European Economic Review*, *40*(6), 1353–1375.

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and Longitudinal Modelling using Stata*. STATA press, College Station, TX.

Rao, V., Gupta, I., Lokshin, M., & Jana, S. (2003). Sex workers and the cost of safe sex: the compensating differential for condom use among Calcutta prostitutes. *Journal of Development Economics*, *71*(2), 585–603.

Reichel, R., & Topper, K. (2003). Prostitution: verkannte Wirtschaftsfaktor. Aufklärung und Kritik. *Zeitschrift für freies Denken und humanistische Philosophie*, *10*(3), 2–30.

Reynolds, H. (1986). *The Economics of Prostitution*. Charles C Thomas Pub Ltd, Springfield, IL.

Richter, M., Luchters, S., Ndlovu, D., Temmerman, M., & Chersich, M. F. (2012). Female sex work and international sport events - no major changes in demand or supply of paid sex during the 2010 Soccer World Cup: A cross-sectional study. *BMC Public Health*, *12*, 763.

Robinson, J., Yeh, E., American, S., Journal, E., Economics, A., January, N., Caldwell, P., & Caldwell, J. C. (2011). Transactional sex as a response to risk in Western Kenya. *American Economic Journal: Applied Economics*, *3*(1), 35–64.

Rose, A. K., & Spiegel, M. M. (2011). The olympic effect. *Economic Journal*, *121*(553), 652–677.

Ruggiero, P., Ruggiero, V., South, N., & Taylor, I. (2002). *The New European Criminology: Crime and Social Order in Europe*. Taylor & Francis, London, UK.

Ryan, C., & Hall, M. C. (2005). *Sex Tourism: Marginal People and Liminalities*. Routledge, London, UK.

Sanders, T. (2004). The risks of street prostitution: Punters, police and protesters. *Urban Studies*, *41*(9), 1703–1717.

Sanger, W. W. (2006). *The History of Prostitution: Its Extent, Causes, and Effects Throughout the World. Being an Official Report to the Board of Alms-house Governors of the City of New York*. Ulan Press.

Scheffé, H. (1999). *The Analysis of Variance*. Wiley-Interscience, Hoboken, NJ.

Schefold, B. (1994). *Wirtschaftsstile: Studien zum Verhältnis von Ökonomie und Kultur*. Fischer Wissenschaft, Frankfurt, DE.

Schefold, B., & Carstensen, K. (2002). Die klassische politische Ökonomie, in *Issing, Geschichte der Nationaloekonomie*. Verlag Franz Vahlen, Munich, DE.

Schinzinger, F. (2002). Vorlaeufer der Nationaloekonomie, in *Issing, Geschichte der Nationaloekonomie*. Verlag Franz Vahlen, Munich, DE.

Schmitter, R. (2013). Prostitution - Das "älteste Gewerbe der Welt"? *Aus Politik und Zeitgeschichte*, *9*, 22–28.

Tannahill, R. (1992). *Sex in History*. Scarborough House, Lanham, MD.

The German Delegation to the Multidisciplinary Group on Organised Crime (2007). Experience Report on Human Trafficking for the Purpose of Sexual Exploitation and Forced Prostitution in Connection with the 2006 Football Worldcup in Germany. *Council of the European Union*, *5006/1/07*(REV 1).

Tourism Vancouver (2007). Economic impact of conventions in Metro Vancouver. Last access on 20.02.2015.
URL `http://www.tourismvancouver.com/includes/content/images/media/docs/mc_economic_impact.pdf`

Vella, F., & Verbeek, M. (1999). Estimating and interpreting models with endogenous treatment effects. *Journal of Business & Economic Statistics*, *17*(4), 473–478.

VG Cologne (2007). AZ 23 K 4180/04 - Kölner "Sexsteuer" ist im wesentlichen rechtmäßig, 11.07.2007. Last access on 20.02.2015.
URL `http://www.vg-koeln.nrw.de/presse/pressemitteilungen/archiv/2007/14_070711/index.php`

Vosskort, A., & Necker, S. (2014). Politics and parents – Intergenerational transmission of values after a regime shift. *European Journal of Political Economy*, *36*(December), 177–194.

Weber, T. (2012). Price theory in Eonomics, in *Ozer & Phillips, The Oxford Handbook of Pricing Management*. Oxford University Press, Oxford.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, *61*(3), 439–447.

Wilson, N. (2012). Economic booms and risky sexual behavior: Evidence from Zambian copper mining cities. *Journal of Health Economics*, *31*(6), 797–812.

Wolszczak-Derlacz, J., & De Blander, R. (2009). Price convergence in the European Union and in the new member states. *Bank i Kredyt*, *40*(2), 37–59.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

Chapter A

Appendices

A.1 Appendix for Chapter 2

A.1.1 Convergence decomposition - Deduction

Price indices are in natural logs. There are $m+n$ countries in two subgroups of size m and n, and prices of a specific product group are normalized such that $\bar{p}_{m+n} = 0$. For better readability we first omit the product group index k.

$$\underbrace{p_1, ..., p_m}_{m}, \underbrace{p_{m+1}, ..., p_{m+n}}_{n}$$

$$\bar{p}_{m+n} = \frac{1}{m+n}\sum_{i=1}^{m+n} p_i = 0 \quad ; \quad \bar{p}_m = \frac{1}{m}\sum_{i=1}^{m} p_i \quad ; \quad \bar{p}_n = \frac{1}{n}\sum_{i=m+1}^{m+n} p_i$$

We use the standard textbook equation of the variance:

$$\begin{aligned}
\sigma^2 &= E(X^2) - (E(X))^2 \\
\sigma^2_{m+n} &= \frac{1}{m+n}\sum_{i=1}^{m+n} p_i^2 - \underbrace{(\bar{p}_{m+n})^2}_{=0} \\
&= \frac{1}{m+n}\sum_{i=1}^{m+n} p_i^2 \\
\sigma^2_m &= \frac{1}{m}\sum_{i=1}^{m} p_i^2 - (\bar{p}_m)^2 \\
\sigma^2_n &= \frac{1}{n}\sum_{i=m+1}^{m+n} p_i^2 - (\bar{p}_n)^2 \\
m \cdot \sigma^2_m + n \cdot \sigma^2_n &= \sum_{i=1}^{m+n} p_i^2 - m \cdot (\bar{p}_m)^2 - n \cdot (\bar{p}_n)^2 \\
&= (m+n) \cdot \sigma^2_{m+n} - m \cdot (\bar{p}_m)^2 - n \cdot (\bar{p}_n)^2
\end{aligned}$$

Hence:

$$\boldsymbol{\sigma^2_{m+n} = \frac{m}{m+n} \cdot \sigma^2_m + \frac{n}{m+n} \cdot \sigma^2_n + \frac{m}{m+n} \cdot (\bar{p}_m)^2 + \frac{n}{m+n} \cdot (\bar{p}_n)^2}$$

The variance of the whole sample consists of the subgroups' weighted variances and their weighted quadratic means.

Since $m \cdot \bar{p}_m + n \cdot \bar{p}_n = 0$, and therefore both means are either moving towards or away from zero, we can further simplify using $\bar{p}_m = -\frac{n}{m} \cdot \bar{p}_n$:

$$\begin{aligned}
\frac{m}{m+n} \cdot (\bar{p}_m)^2 + \frac{n}{m+n} \cdot (\bar{p}_n)^2 &= \frac{m}{m+n} \cdot (-\frac{n}{m} \cdot \bar{p}_n)^2 + \frac{n}{m+n} \cdot (\bar{p}_n)^2 \\
&= (\frac{m}{m+n} \cdot \frac{n^2}{m^2} + \frac{n}{m+n})(\bar{p}_n)^2 \\
&= \frac{mn^2 + m^2 n}{(m+n)m^2}(\bar{p}_n)^2 \\
&= \frac{(m+n)mn}{(m+n)m^2}(\bar{p}_n)^2 \\
&= \frac{n}{m}(\bar{p}_n)^2
\end{aligned}$$

Thus:

$$\sigma^2_{m+n} = \frac{m}{m+n}\sigma^2_m + \frac{n}{m+n}\sigma^2_n + \frac{n}{m}(\bar{p}_n)^2,$$

or equivalently

$$\sigma^2_{m+n} = \frac{m}{m+n}\sigma^2_m + \frac{n}{m+n}\sigma^2_n + \frac{m}{n}(\bar{p}_m)^2.$$

Expressed in changes and adding the product group index k, we get:

$$\Delta\sigma^2_{k,m+n} = \frac{n}{m+n}\Delta\sigma^2_{k,m} + \frac{n}{m+n}\Delta\sigma^2_{k,n} + \frac{n}{m}\Delta(\bar{p}_{k,n})^2$$

Finally, we aggregate over all product groups k:

$$\Delta\sigma^2_{m+n} = \frac{1}{K}\sum_{k=1}^{K}\sigma^2_{k,m+n} = \frac{m}{m+n}\Delta\sigma^2_m + \frac{n}{m+n}\Delta\sigma^2_n + \frac{n}{m}\frac{1}{K}\sum_{k=1}^{K}\Delta(\bar{p}_{k,n})^2$$

In the EU23 case with $K = 34$ product groups:

$$\Delta\sigma^2_{k,EU23} = \frac{15}{23}\Delta\sigma^2_{k,EU15} + \frac{8}{23}\Delta\sigma^2_{k,EU8} + \frac{8}{15}\Delta(\bar{p}_{k,EU8})^2$$

$$\Delta\sigma^2_{EU23} = \frac{15}{23}\Delta\sigma^2_{EU15} + \frac{8}{23}\Delta\sigma^2_{EU8} + \frac{8}{15}\frac{1}{34}\sum_{k=1}^{34}\Delta(\bar{p}_{k,EU8})^2$$

A.1.2 Confidence Intervalls

For better readability, here we use $\mathfrak{p}_t = \frac{1}{34}\sum_{k=1}^{34} \Delta\bar{p}^2_{k,EU8,t}$

Table A.1: *Paired t-test and Wilcoxon matched-pairs signed-ranks test for* $\mathfrak{p}_t$

Period	N	$\mathfrak{p}_{t-1}$	$\mathfrak{p}_t$	$\Delta\mathfrak{p}_t$	$\frac{8}{15}\Delta\mathfrak{p}_t$	t	[99% C.I.] of $\Delta\mathfrak{p}_t$		z
2001-2003	34	0.1001	0.1217	0.0216	0.0115	-3.4388	- 0.0388	- 0.0044	- 3.1200
2003-2006	34	0.1217	0.0792	- 0.0425	- 0.0227	7.656	- 0.0273	- 0.0577	4.8640
2006-2009	34	0.0792	0.0481	- 0.0311	- 0.0166	5.5055	- 0.0157	- 0.0466	4.9150
2003-2009	34	0.1217	0.0481	- 0.0736	- 0.0394	8.6095	- 0.0503	- 0.0970	5.0860

The table presents the results of the paired t-test and Wilcoxon matched-pairs signed-ranks test for $\mathfrak{p}_t$. Both tests show significance at the 99% level. t is the t statistic of the paired t-test z is the z statistic of the Wilcoxon matched-pairs signed-ranks test. The confidence interval of $\mathfrak{p}_t$ has been calculated based on the paired t-test.

A.2 Appendix for Chapter 3

A.2.1 A comparison of professional and non professional sex workers

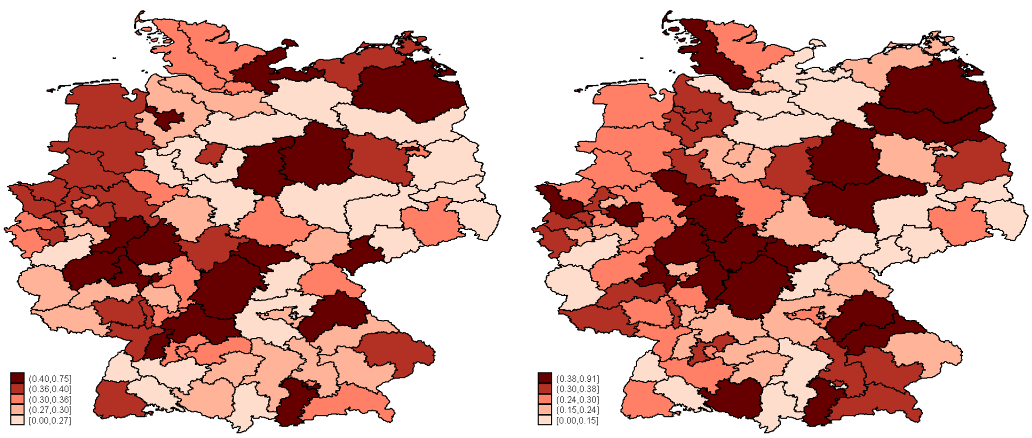
On *www.gesext.de* not only professional sex workers offer services, but also non-professional ones (claiming to offer services only occasionally and for whom sex work is not their main source of income). This is a group of sex workers, which is uncovered in nearly all other data sets on sex work, but which makes up a large part of our data: 27.2% of the 102.096 offers have been made by a woman that claims to be a non-professional seller (1355 out of 4070 sellers).

In general, at the time of making the contract the potential customer does not know if the claim is true and neither do we, when one can only look at self descriptions of the sellers. In this data set the sellers' ratings are close to 100%, which might indicate that after having met the sellers, customers seem to be convinced that the information they were given – including the claim to be a non-professional seller – was valid. However, this cannot be tested. In general, if the information was worthless it would not be worth mentioning. As there is no other data set able to compare these sub-groups, we use the dummy variable *seller claims to be no professional sex worker* to create two subsamples and compare them. So what we actually are able to look at are two things: 1) If sellers that claim to be non-professionals have different attributes and offer different services, and 2) if customers value services offered by these non-professional sellers at different prices.

Figure A.1: *Maps: Non professional sex workers*
Sex work in Germany at the 2 digit zip code level. Darker colors represent higher values of the variable in the given region. *Source: Own representation.*

(a) Share of non-professional sellers per zip code area and figure

(b) Share of offers made by non-professional sellers per zipcode area.



Figures A.1a and A.1b show how high the percentage of non-professional sellers and the percentage of offers made by them is in each 2 digit zip code. In the areas marked dark red this percentage is higher than 40% and 31% respectively, with the maximum of 75% non-professional sellers and 91% non-professionally offered services in Suhl (D-98). Nearly all of the dark red areas are rural areas and in most cities the share of non-professional sellers and offers is close to the average of 27.2% and 33.4%. However, there are also many rural areas with few or even no non-professional sellers. There seems to be no clear pattern.

When looking at attributes of the sellers and services offered, some differences appear.[1] First of all, non-professional sellers put more effort into their self descriptions. The average text length is 85 words longer for non-professional sellers (252 words in total) and there are less misspellings per word. This could be partly due to a higher educational status of the non-professional sellers: 14% of them claim to be students, while no seller in the other subsample makes this claim. The longer free texts result in information on personal attributes such as height, weight, age, etc. being specified more often. Interestingly, on average these personal attributes are very similar to the ones of non-professional sellers. The latter just weigh a bit more (70 kg, BMI 24.7 vs. for non-professional vs. 63 kg, BMI 22.9 for professionals), claim less often to be single (10% vs. 30%) and more often to have a partner (8% vs. 4%). On the pictures non-professional sellers are more often naked (10.2% vs. 7.2%), but show their face less often (18.0%, vs. 27.5%). This is unsurprising, as non-professional sellers have more to loose if they are recognized by other people, but 18% is still a high share.

When we look at sold prices, per sale non-professionals earn more money than professionals, but this can be explained by non-professional sellers offering slightly different services. The difference in hourly wages for the same service is not statistically significant (e.g., 101.20 euro vs. 100.80 euro for vaginal sex). Non-professional sellers offer longer sessions (3.1 vs. 2.7 hours) and less often only massages, manual or oral stimulation (6.8% vs. 15.3%), but rather vaginal sex. Professional sellers make up part of this by offering more unprotected (5.7% vs. 6.7%) and anal sex (14.9% vs. 16.9%). In total, the average professional seller earns 30 euro less per sale than a non-professional one, but the first group makes up for this by offering services more often (8.6 vs. 7.2 sold services per seller). Finally, when looking at the location the meeting will take place, non-professional sellers offer less often to meet at their own homes and more often at the seller's home or a hotel. Again, this is unsurprising as non-professional sellers might want to protect their "moonlighting" from neighbors and friends and might not have enough customers to afford renting a second location, as professional sellers might do.

When looking at the customers side, only four percent of the men that have bought at least one service from a non-professional seller bought all other services from a non-professional seller as well. Hence, most of them do not have a preference for one of the two groups. To sum up, these two groups differ, but not too much.

[1] Statistical significances in this paragraph have been tested by using unpaired t-tests. Unless stated otherwise, results are significant at the 1% level.

A.3 Appendix for Chapter 6

A.3.1 RDD: General Methodology

First, denote a generic outcome for unit i by y_i and a continuous so-called forcing variable $d_i \in (-\infty, \infty)$ – distance to the internal East-West border – which displays a continuous impact on outcome and determines where the internal border is, at which outcome may be discontinuous. For convenience, use $d_i \in (-\infty, 0)$ in the West and $d_i \in [0, \infty)$ in the East and define the binary variable $east_i = 1$ if $d_i \geq 0$ and $east_i = 0$ if $d_i < 0$. Second, define two flexible and smooth (e.g., polynomial) functions about d_i at $d_i < 0$ and $d_i \geq 0$ as

$f_0(d_i)$ at $east_i = 0, f_1(d_i)$ at $east_i = 1,$ and $f_1^*(d_i) \equiv f_1(d_i) - f_0(d_i)$

Then, the conditional expectations of outcome y_i may be specified as

$$E[y_i|y_i d_i, east_i = 0] = \alpha + f_0(d_i) \qquad (A.1)$$

$$E[y_i|y_i d_i, east_i = 1] = \alpha + f_0(d_i) + LATE_{east} + f_1^*(d_i) \qquad (A.2)$$

Where $LATE_{east}$ is the local average treatment effect of interest. Technically,

$$LATE_{east} \equiv \lim_{\Delta \to 0} \left(E[y_i \mid 0 < d_i < \Delta] - E[y_i \mid -\Delta < d_i < 0] \right). \qquad (A.3)$$

For a further discussion of an RDD, see Imbens & Lemieux (2008) as well as Lee & Lemieux (2010). In the approach pursued here, we present two types of results: one is based on the main effect, assuming homogeneity of $LATE_{east}$ across all places along the border, and an alternative one permits $LATE_{east}$ to vary with log population and, to permit for a further discontinuity, with Berlin.