CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Master thesis

A Meta-Analysis of the Effect of Minimum Wage Increases on Prices

Author: **Bc. Jana Vavřičková**

Supervisor: **PhDr. Tomáš Havránek, Ph.D.**

Prague, 2015

Declaration of Authorship

I hereby proclaim that I wrote my Master thesis on my own under the leadership of my supervisor and that the references include all resources and literature I have used.

I grant a permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 15, 2015

Jana Vavřičková

Acknowledgment

I would like to use this opportunity to express my gratitude to my supervisor, PhDr. Tomáš Havránek, Ph.D. for introducing me to the filed of meta-analysis and for his guidance through the extensive body of its methodology. I would like to thank him no less for the patience, time, valuable advice and remarks he has so kindly given me throughout my work on the thesis.

Furthermore, I would like to thank my colleagues for always trying to accommodate for my need of a flexible schedule as best they could in order for me to be able to focus on my university studies as well as on my job.

Last but not least I would like to thank my friends and my family for being so supportive and for always encouraging me in pursuing my goals.

Citation: VAVŘIČKOVÁ, Jana (2015) *"A Meta-Analysis of the Effect of Minimum Wage Increase on Prices"* Prague: Charles University, Faculty of Social Sciences, Institute of Economic Studies, 2015, Thesis supervisor: PhDr. Tomáš Havránek, Ph.D.

Abstract

As an economically as well as politically sensitive topic, labor market interventions stir up discussions among professionals as well as general public. Most economists take negative stance against minimum wage policies providing arguments backed by theoretical reasoning rather then sound empirical evidence. Knowledge of labor market outcomes and their transmission channel to other segments of the economy are till nowadays limited and inconsistent. Neither empirical research in the field contributes to a uniform consent on the impact of minimum wage hikes on the price level. Moreover, the reported estimates display large heterogeneity and after a brief inspection reveal that the field is infested with publication selectivity. A uniquely constructed dataset consisting of 469 estimates of the price effect of minimum wage changes and their associated characteristics is analyzed using a set of statistical tools generally known as meta-analysis. The method is a powerful tool nowadays widely used in empirical research to synthesize and systematically evaluate sometimes inconsistent research results. While the study finds no consistent evidence of an actual link between minimum wage hikes and inflationary pressures, the empirical results show strong presence of publication selectivity.

Keywords meta-analysis, publication selectivity, minimum wage, price level, inflation, prices

JEL classification C83, E31, J31

Abstrakt

Ekonomicky a politicky citlivé téma, jakým jsou zásahy na trhu práce, rozdmýchává diskuze jak mezi profesionály, tak mezi veřejností. Většina ekonomů zaujímá vůči politice minimálních mezd negativní stanovisko, přičemž jsou častěji používány argumenty založené na teoretické bázi spíše než na empirických důkazech. Poznatky o fungování trhu práce a přenosu vzniklých šoků do ostatních odvětví ekonomiky jsou dodnes limitovány. Ani empirické studie v oboru nepřispívají k jednotné shodě ohledně vlivu změn minimální mzdy na

cenovou hladinu. Dostupné odhady tohoto vlivu navíc vykazují významnou heterogenitu a při bližším prozkoumání poukazují na přítomnost silné publikační selektivity. Na unikátně sestrojenou datovou matice skládající se z celkem 469 odhadů této veličiny a jejich přidružených charakteristik používá tato studie soubor statistických metod známých pod pojmem meta-analýza. Tato metoda dnes nachází široké využití pro syntézu a systematické zhodnocení výsledků empirických článků. Zatímco existence vztahu mezi pohyby minimální mzdy a cenové hladiny není meta-analýzou potvrzena, nacházíme mezi odhady důkazy přítomnosti silné publikační selektivity.

Keywords meta-analýza, publikační selektivita, minimální mzda, cenová hladina, inflace, cena

JEL klasifikace C83, E31, J31

Master Thesis Proposal

Author	Bc. Jana Vavřičková
Supervisor	PhDr. Tomáš Havránek Ph.D.
Topic	A Meta-Analysis of the Effect of Minimum Wage Increase on Prices

Topic Characteristics: Various consequences of minimum wage changes have been the concern of empirical studies for decades, yet only a minor part of the literature considers the price effect of such changes. Standard economic theory predicts that firms in a competitive environment will reduce employment should the minimum wage increase. However, if such an increase happens to be industry wide, firms will pass the additional costs on to prices. The existing literature on this effect has been summarized and reviewed in a comprehensive survey by Lemos (2006b), yet the comparative character of the study leaves potential for a more precise and objective quantitative review. I intend to use meta-regression techniques to draw a clear picture of the effect. The main contribution of my diploma thesis should be the provision of an objective alternative to the narrative review of the available literature that has been published by Lemos (2006b).

Hypotheses:

1. Significant estimates of the price effect of minimum wage changes are published preferentially
2. The researchers are more likely to publish a positive estimate of the effect (presence of publication bias)
3. The large differences in the estimates can be explained by variation in methodology and datasets that have been used to obtain these results

Methodology: I intend to collect a data set of the empirical results (estimates and t-statistics) and important characteristics of academical and other studies that deal with the price effect of minimum wage increase. The meta-analysis as

a tool guarantees a greater objectivity of the reviewing process, as the studies are taken into consideration based on a fixed rule set by the reviewer beforehand. In order to clear the results of the meta-regression analysis of possible distorting factors I will also employ measures to mitigate the publication bias and deal with additional heterogeneity in the data.

Outline:

1. Introduction
2. Theoretical background and methods used in the literature for empirical testing of the effect
3. Meta-analysis introduction
   (a) History and its use
   (b) Methodology
4. Meta-regression
   (a) Data set collection
   (b) The actual analysis
   (c) Dealing with publication bias and heterogeneity
   (d) Results
5. Discussion and concluding remarks

Core Bibliography:

1. Lemos, S. (2008): *"A Survey of the Effects of the Minimum Wage on Prices."'* Journal of Economic Surveys 22(1): pp. 187-212.
2. Stanley, T. D. & Doucouliagos, H. (2012): *"Meta-regression analysis in Economics and Business"'*, 1st edn., Routledge
3. Stanley, T. D. (2001): *"Wheat from Chaf: Meta-analysis as Quantitative Literature Review."'* Journal of Economic Perspectives 15(3): pp. 131-150.
4. Stanley, T. D. & S. B. Jarrell (1989): *"Meta-Regression Analysis: A Quantitative Method of Literature Surveys."'* Journal of Economic Surveys (3)2: pp. 161-170
5. Stanley, T. D. (2005): *"Beyond Publication Bias."'* Journal of Economic Surveys 19(3): pp. 309-345

Contents

List of Figures

List of Tables

List of Acronyms

2SLS Two stage least squares

ECM Error correction model

FAIVEHR Funnel asymmetry instrumental variable heteroscedasticity robust estimator

FAIVE Funnel asymmetry instrumental variable estimator

FAT Funnel asymmetry test

FE Fixed effects

IV Instrumental variable

LD Labor demand

LS Labor supply

ME Mixed effects

MLE Maximum-likelihood estimator

MRA Meta regression analysis

OLS Ordinary least squares

PET Precision effect test

SMIC Salaire minimum de croissance (statutory minimum wage in France)

VAR Vector autoregression

WLS Weighted least squares

Introduction

Minimum wage, the controversial policy so often despised by economists for its allegedly distorting impact on labor market outcomes, was according to the 2014 Global Wage Report actively maintained in 118 states worldwide including the United States of America and most members of the European Union. Even when a statutory minimum wage policy is not enacted by law, substitutes in the form of sector-wide collectively agreed minimum wage rates are often in place.

Card and Krueger (1995a) express surprise at how many economists can find common ground when it comes to opposing minimum wage policies, especially since economics are so often subject to dispute by so many. It should be recognized that most of the negative opinions so heavily expressed by economists are based on abstract theoretical reasoning more or less directly derived form the Neoclassical model rather than on sound empirical evidence. The question however is that even when empirical evidence exists, and and we know for a fact that it does, how do we separate the wheat from the chaff? Meta-analysis and in particularly meta-regression analysis is a type of literature review method tailored for use in analyzing empirical research and can therefore help us answer the question.

Over the years, meta-analysis has developed into a comprehensive statistical tool-set capable of supplementing or even substituting the widely used narrative reviews. Its power comprises mainly in the ability to rid the reviewing process of subjectivity by assessing individual studies based on a set of predefined criteria. Nevertheless, it can be used in far more versatile ways. Meta-regression analysis may help us explain why studies sometimes come to diametrically opposed conclusions or even uncover deliberate data manipulation in professional literature.

The following thesis is to the best of our knowledge the very first study to provide a quantitative review of professional literature focused on price response to changes in minimum wage policy. Besides the common synthesis of

reported findings, the thesis addresses any systematic selectivity by journals or authors themselves that could impair the overall perceived effect. Lastly, a comprehensive model is build that can explain some of the variation in the reported estimates by study-level characteristics.

The diploma thesis is organized as follows: Chapter 1 outlines a theory linking the policies introduced in the labor market to their outcomes observed in the aggregate goods market. Chapter 2 contains an overview of estimation techniques and models to be found in the minimum wage literature. Data definitions and the actual data collection process are described in Chapter 3 while the meta-analytical tool set applied in our analysis is briefly summarized in Chapter 4. Finally, Chapter 5 elaborates on the actual analysis and presentation of results.

Chapter 1

Price effect of minimum wages: theory

Increasing minimum wage levels requires that firms, particularly firms in the low wage sector, take action to cope with the increased labor costs. Common sense dictates, that this can be achieved in three ways - through a reduction in employment, profits, or by increasing the prices of final products.

Earlier essays, such as Stiegler's (1946), were quite positive about the detrimental effect of increased minimum wages on employment levels. However, the empirical evidence published over the years was scanty and often contradictory. In 1995, Card and Krueger's meta-analysis of minimum wage impact on unemployment levels found an at that time rather unexpected pattern in the literature, which did not provide any evidence for the actual existence of a real effect. Card and Krueger attributed the confusing results to publication bias and specification searching, or possibly to a structural change in the data. Doucaliagos and Stanley (2009) have conducted another meta-analysis in which they have partly confirmed the findings of Card and Krueger as well as discovered evidence of a strong publication bias in the minimum wage literature. Furthermore, they have challenged Card and Krueger's interpretation of the negative relationship between study's t-statistic and degrees of freedom, claiming that the negative relationship was not caused so much by publication selection as it was caused by the absence of any genuine effect whatsoever.

In contrast to the widely explored unemployment level elasticity of minimum wage increases, studies of the impact of minimum wage shift on companies' profits remain rather rare. In one such study, Card and Krueger (1995a) note

that minimum-wage paying companies tend to be smaller and concentrated in retail, especially in restaurant industry. Small companies in highly competitive sectors, however, often cannot absorb the extra labor costs into profits (Lemos, 2006b). Furthermore, Card and Krueger (1995a) point out, that the effect of the increase in minimum wages often tends to be of sectoral extent in which case the firms may answer the rising costs of labor by increasing the overall price of the industry output.

The price effect of an increase in the minimum wage is nonetheless hardly unambiguous. In the next few paragraphs, we aim to demonstrate that there is no uniform theoretical consent about the actual existence let alone about the direction of the price effect of a minimum wage increases. The theoretical outcome of introducing a minimum wage into a labor market varies among other things with the level up to which the analyzed labor market structure resembles perfect competition. It is therefore convenient for the purposes of describing this impact to pay special attention to extreme cases in the labor market. Building on the example Borjas (2005), we examine the way labor market reacts to minimum wage introduction under perfect competition and monopsony, i.e. a situation where the firms face an upward sloping supply curve of labor and therefore cannot hire unlimited amount of workers at a fixed wage without being forced to offer higher wages to attract more workers.

1.1 Labor market solution

Perfect competition

Attempts were made to broaden the spectrum of the analysis from the partial labor market solution to a much more general model which could embody the economy as a whole, naturally in a much simplified manner. For our needs of demonstrating the diversity of the possible outcomes in different market environments, we find the simple diagram model of Werner et al. (2013) suitable. Not unlike Borjas, Werner et al. illustrate the effect of introducing a minimum wage under perfect competition and monopsony into the labor market, but in addition to the original model, their paper interlinks the labor market with the goods market via a typically shaped production function with decreasing marginal returns of labor.

The original equilibrium employment choice under perfect competition L^*

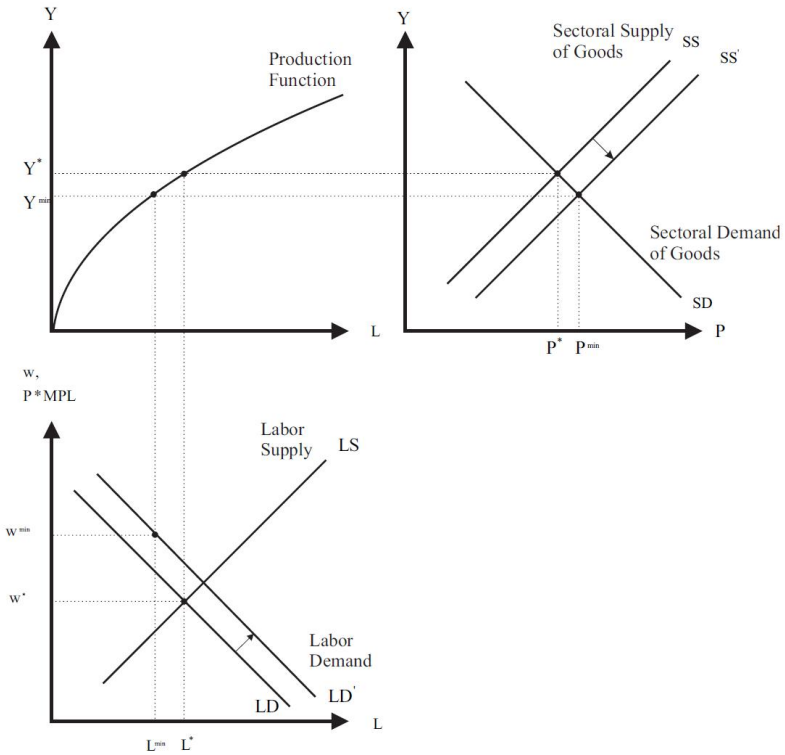

Figure 1.1: Price effects of a minimum wage in a competitive market (Werner, Sell, and Reinisch, 2013)

lies in the intersection of the LD and LS curves. Since Werner's model assumes the economy only in the short term, all other input factors beside the amount of labor remain fixed. The production function under consideration is therefore a function of a single variable - labor, and we can read off the amount of produced goods from the graph in the upper left corner. Finally, the initial price level is determined in the goods market in the upper right corner at the intersection of our well-behaved supply and demand curves.

Let us now consider a binding minimum wage w_{min}. For simplicity we assume no migration between the covered and uncovered sectors, i.e. sectors affected and unaffected by the minimum wage policy. The basic scenario starts with the firms reacting by lowering their employment levels. This naturally leads to a lower supply of goods to which the goods market responds by increasing the final price of the product.

Werner's model explores the sequence of market reactions one step further by assuming the increased price expectations and the following increased value of the marginal product of labor reflect back to the decision making on the labor market. By assuming the existence of such backward projection, we demonstrate

that even on purely theoretical basis, the conclusions drawn from economic models critically hinge on the underlying assumptions. As the labor market expects increased goods prices, the labor demand curve shifts to its left position LD'. Contrary to the expected outcome when no backwards projection is assumed, the increased demand for labor LD' reduces the resulting price effect.

Nondiscriminating monopsony

The non-competitive labor market structure analyzed by Werner et al. (2013) is that of a nondiscriminating monopsonist. As it is stated in Borjas (2005), a nondiscriminating monopsonist is not able to differentiate between his employees according to their reservation wage and therefore must raise all wages in case he wishes to attract more workers. As a result, the cost of labor is represented by an upward sloping schedule, which is steeper than the supply curve and lays above it (see the lower left graph in figure 1.2). A profit-maximizing monopsonist then finds his optimal employment by setting his value of marginal product (i.e. labor demand) equal to his marginal cost of labor. Note that in order to attract his optimal amount of labor L_{mon}, the monopsonist only needs to pay the wage W^{mon}. Naturally, as in the perfect competition scenario, the amount of labor fully determines the production level Y^{mon}, supply of goods and consequently also the price level P^{mon}.

In the next step, Werner et al. (2013) present the outcome of imposing a binding minimum wage that is set higher than the optimum minimum wage of the monopsonist W^{mon}, but lower than the equilibrium wage under perfect competition W^*. In case the employer adheres to the new minimum wage law, he acts as a price taker and is able to hire up to L^{min} workers at the price of W^{min}. However, in order to attract a higher amount of labor, the monopsonist needs to start raising the wages to all of his employees once again. The resulting schedule of his marginal costs of labor is marked in grey in figure 1.2. Due to the introduction of the minimum wage, the monopsonist raises employment to L^{min} which subsequently reflects in the higher level of production and supply SS'. The goods market reacts to the increased supply by decreasing the final price level.

As in the prefect competition scenario, Werner, Sell, and Reinisch go slightly further by working the resulting lower value of the marginal product of labor into the decision making of agents in the labor market in the form of a downward shifted labor demand schedule LD'.

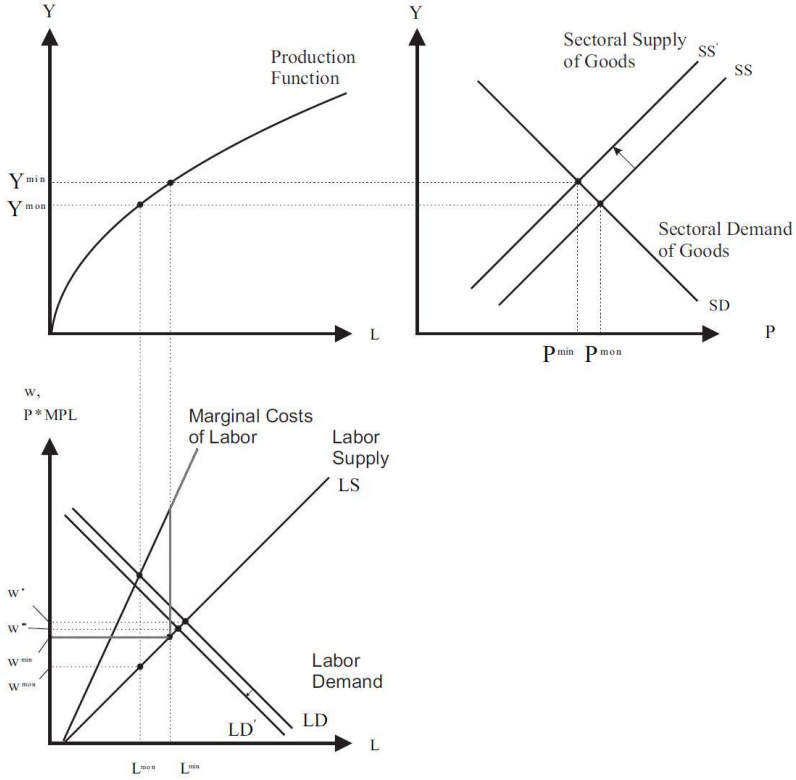

Figure 1.2: Price effects of a minimum wage under monopsony (Werner, Sell, and Reinisch, 2013)

With the help of Werner's analysis, we have shown how much the predicted outcomes of introducing a minimum wage can vary under different levels of competition in the market. However, in reality many other factors such as the time horizon, the elasticities of the schedules in the goods market, even factors related to geographical affiliation contribute to the final outcome of the equation. At that point, where the literature gets too contradictory and complex to be grasped without a systematic concept, a meta-analysis can help us identify some of the factors driving the differences in the literature.

1.2 Price pass-through channels

In light of the preceding analysis of labor market outcomes, the aim of the following paragraph is to pinpoint and recap the overview given by Lemos (2006b) and demonstrate on a thought experiment the many ways changes in the price level can be observed as a consequence of introducing a new policy into the labor market.

Aggregate perceived price level is affected by minimum wage increases via several channels. Principally, aggregate supply and labor demand create an upward pressure on prices whereas the reaction of labor supply works in the opposite direction. On the other hand, aggregate demand can in fact work either way.

Raising costs of labor contribute to an increase in output prices through labor demand - if firms adjust to the new conditions on the labor market by decreasing their optimum employment level and thus creating less output, the overall decrease in the aggregate supply will inevitably push the prices upwards. Furthermore, the perceived higher aggregate income of households reinforces the upward pressure on prices through an upward shift in the aggregate demand. However, the direction of aggregate demand shift is not as unambiguous as it seems on the first sight. Unless the producers of cheap-labor-intensive goods are able to absorb the increased labor costs into their profits, prices of such goods will grow putting a downward pressure on the aggregate demand for goods. Finally, it can be expected that the increased minimum wages improve labor productivity which in the end creates downward pressure on prices via a decrease in labor demand (Lemos, 2006b).

All these factors work together in a transmission mechanism which could be in a simplified way described as follows. The first to be affected by the new minimum wage are the workers to whom the increase in wages directly applies, with possible spill-over effects to higher wage categories. Firms are subsequently forced to increase prices to compensate for the higher labor costs. They also adjust the levels of input and output so that the cost minimization condition is maintained under the new conditions. New equilibrium income level is determined from the new level of employment and wages paid. The level of demand then depends on the newly determined income level. Finally, production level adjusts in order to be in an equilibrium with the aggregate demand. The cycle then closes when the new level of inflation and unemployment stabilizes over time, until a new change of labor market policy becomes effective (Lemos, 2006b).

It should be noted that not all studies in the field of minimum wage research have the ambition to estimate the size of the effect after the minimum wage change propagated through all stages of the transmission mechanism. In fact, there is a rather clear distinction between studies focusing on the big picture of the economy and those concentrating more on the effects on a micro-level, studying price level responses of particular sectors or even single items.

Chapter 2

Price effect of minimum wages: empirical methods

When designing the empirical model, some authors put more emphasis on developing a sound theoretical base, whereas others abandon standard economic theory and instead turn to pure statistical modeling and their own judgement. The resulting specification of the empirical model not only affects its interpretation, but often directly requires that a particular type of dataset is used. Both of the factors have the potential to dramatically alter the final outcome of the analysis.

The downside to the particular branch of minimum wage literature, previously noted also by Lemos (2006b), is that especially earlier studies often provide only scanty information about the models applied to estimate the reported relationships (Frye and Gordon, 1981). Interestingly, to the best of our knowledge the very first empirical publication in the field written by Fels and Hoa (1981) employs the causality test procedures proposed by Granger (1969). From today's perspective, knowing the pitfalls brought on by frequent non-stationarity in economic time series, the pre-whitening measures taken in the paper, namely the *ad-hoc* inclusion of a constant term into the fitted regression equations, might be perceived as insufficient. Data non-stationarity is addressed in a much more correct manner in the later studies (see e.g. L'horty and Rault (2004) and Andreica, Aparaschivei, Cristescu, and Cataniciu (2010)) using differentiation and vector-error correction models.

The following chapter contains an overview of the methods and models appearing in the empirical literature in the field of minimum-wage price effect research.

2.1 General equilibrium model

As the name of the model suggests, the concept of a general equilibrium was designed as a simplified yet complete model of the economy with several interacting markets. The theory seeks to find a set of prices which would bring the markets to equilibrium. Use of highly aggregated data is typical for studies employing the approach.

Empirical specification derived from a General equilibrium model was employed in the two studies by Lemos (2003, 2004b) on Brazilian price and wage level data. Both studies use a simplified general equilibrium model composed of a set of structural equations for labor demand, labor supply, aggregate demand and aggregate supply. Assuming the production function $F(L_s, L_u, K)$ to be a function of skilled and unskilled labor and capital, we can specify the demand for labor as a function of the prices of the corresponding output and input factors. Lemos (2006b) further differentiated between the demand for skilled and unskilled labor $L_s^d(W, W^M, r, P)$ and $L_u^d(W, W^M, r, P)$. The standard labor-leisure theory leads to a labor supply function depending on paid wages and output price. Again, Lemos (2006b) uses a different specification for skilled and unskilled labor $L_s^s(P, W)$ and $L_s^s(P, W^M)$ respectively, the second depending on the level of minimum wage rather than on the equilibrium wage.

Lemos (2006b) borrows the aggregate supply and demand formulations from Romer (2005) in the form of two functions of output prices and various supply or demand shocks $Y^s(P, Z)$ and $Z^d(P, X)$, where an increase in the minimum wage level can be reflected in one of the shock variables represented by Z and X.

For the purposes of estimation, all of the equations in the system are approximated by their logarithmic form as presented below and estimated simultaneously.

$$\ln L_{st}^d = \alpha_1 + \beta_1 \ln W_t^M + \gamma_1 \ln W_t + \delta_1 r_t + \rho_1 \ln P_t + \nu_{1t} \tag{2.1.1}$$

$$\ln L_{ut}^d = \alpha_2 + \beta_2 \ln W_t^M + \gamma_2 \ln W_t + \delta_2 r_t + \rho_2 \ln P_t + \nu_{2t} \tag{2.1.2}$$

$$\ln L_{st}^s = \alpha_3 + \gamma_3 \ln W_t + \rho_3 \ln P_t + \nu_{3t} \tag{2.1.3}$$

$$\ln L_{st}^s = \alpha_4 + \gamma_4 \ln W_t + \rho_4 \ln P_t + \nu_{4t} \tag{2.1.4}$$

$$\ln Y_t^s = \alpha_5 + \rho_5 \ln P_t + \lambda_5 \ln Z_t + \nu_{5t} \tag{2.1.5}$$

$$\ln Y_t^d = \alpha_6 + \rho_6 \ln P_t + \lambda_6 \ln Z_t + \nu_{6t} \tag{2.1.6}$$

Note that all of the equations above can be inverted so that the logarithm of price level P_t is on the right hand side and thus allow for the estimation of the effect of a change in minimum wage level on prices holding different variables constant at a time. In fact, Lemos (2003) used this particular approach as a robustness check for her reduced form general equilibrium equation model.

A reduced form equation is obtained by substituting the endogenous variables out of the system. We start by using the equilibrium identities $L_u^d = L_u^s$, $L_s^d = L_s^s$, and the equation $L_s + L_u = 1$ to eliminate the wage level W. Substituting the equilibrium amounts of skilled and unskilled labor into the production function $F(L_s, L_u, K)$ yields a modified version of the aggregate supply which leads to the final aggregate equilibrium price equation $P(W^M, r, K, X)$ after using the goods market identity $Y^s = Y^d = Y$. The following price equation is our reduced form and is again estimated in logarithmic approximation as

$$\ln P_t = \alpha_7 + \beta_7 \ln W_t^M + \delta_7 r_t + \kappa_7 \ln K_t + \chi_7 + X_t + \nu_{7t} \tag{2.1.7}$$

The single equation 2.1.7 tells us what happens to the price level after a minimum wage change while encorporating the effect of all parts of the transmission channel through workers, firms and consumers.

2.2 Models of imperfect competition

Another approach utilized in a set of studies by Sara Lemos (2003, 2004a, 2004b, 2005, 2006a) makes use of the profit maximizing condition under imperfect competition where price is assumed to be determined as a mark-up over costs 2.2.1.

$$P = \left(\frac{e}{1+e}\right) C \tag{2.2.1}$$

The price level is denoted by P, C are the costs of production and e is the elasticity of demand. As in the previous case, a logarithmic transformation is used and the equation is estimated in differences. Furthermore, Lemos enhances the model by adding qualitative variables f_i and f_t to control for the regional and time fixed effects. The remaining question is how do we model the right hand side of the equation 2.2.2.

$$\Delta \ln P_{it} = \alpha + \varsigma \Delta C_{it} + f_i + f_t + v_{it} \tag{2.2.2}$$

Lemos assumes that the main components of production costs are the prices of production input factors, i.e. wages and interest rate. A differentiation is made between the average and the minimum wage variables W_{it} and MW_t thus enabling us to control for the effect of the minimum wages alone. Finally, Lemos includes a measure of power consumption E_{it}, productivity A_{it} and introduces dynamics into the model by adding the lags of the dependent as well as minimum wage variables. The complete model, as it is used in the studies included in our dataset, is described by equation 2.2.3.

$$\begin{aligned}\Delta \ln P_{it} = \alpha + \Sigma_{l=-k}^{L} \beta_l \Delta \ln MW_{t-l} + \gamma \Delta \ln W_{it} + \delta \Delta r_{it} + \epsilon \Delta \ln E_{it} + \mu \Delta \ln A_{it} + \\ + \Sigma_{m=1}^{M} \rho_m \Delta \ln P_{it-m} + f_i + f_t + v_{it}\end{aligned} \tag{2.2.3}$$

2.3 Difference-in-difference estimation

Depending on the empirical methodology, the narrative review by Lemos (2006b) proposed a classification of the publications into several groups, one of which referred to as the difference-in-difference studies. No matter how accurate the name might be for the overall character of the listed studies, the estimates of the price response reported within rarely come from regressions designed as difference-in-difference in the prevalent sense of the word. With regard to that, the methodological information in our sample does not always coincide with the original Lemos differentiation.

Disregarding the Lemos classification, we have identified two studies that employ the classical difference-in-difference strategy as we know it from other econometrical literature.

Let us have a two-period panel dataset with observations drawn from two groups, one where the analyzed policy change becomes active in the second period and one where the policy does not apply at all. We will continue by calling the groups the treatment and the control group respectively. The difference-in-difference regression will take the following form

$$price = \beta_0 + \delta_0 P2 + \beta_1 MW_{treatment} + \delta_1 P2MW_{treatment} + other\ factors \quad (2.3.1)$$

where *price* describes the price level, $P2$ is a binary variable assuming the value 0 for the observations collected before the increase in minimum wages and 1 after it. $MW_{treatment}$ is a measure of the minimum wage change in the affected area.

The coefficient of interest which describes the isolated effect of minimum wage change on the price level is the δ_1. In case we drop the vector *other factors* from the regression, the exact same estimate of the minimum wage impact can be obtained by simply averaging the independent variable separately for each period, control, and treatment group and substituting into the formula in 2.3.2

$$\hat{\delta_1} = (\overline{price_{2,T}} - \overline{price_{2,C}}) - (\overline{price_{1,T}} - \overline{price_{1,C}}) \quad (2.3.2)$$

The indices 1,2 and T,C denote the periods, treatment and control group respectively.

2.4 Vector autoregression

Finally, a fraction of the authors decide to take the path of pure econometrical modeling and, as stated by L'horty and Rault (2004), let the data speak for themselves by estimating a vector autoregression or alternatively vector error-correction models (VAR-ECM).

In a setup, where the goal is to uncover any mutual relationships in a group of stationary time series variables without making any a priori assumptions about the existence and direction thereof, the more complex structure of vector autoregressive models presents a in many ways convenient transition state between the univariate time series models on one hand and the systems of simultaneous equations on the other (Cipra, 2008).

Vector autoregressive model $VAR(p)$ of degree p in its reduced form as it is applied by Andreica, Aparaschivei, Cristescu, and Cataniciu (2010), generally

follows the specification of

$$y_t = \varphi_0 + \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + \varepsilon_i, \quad t = 1, ..., n \tag{2.4.1}$$

where n is the number of observations, y_t is a $(m \times 1)$ vector of stationary $I(0)$ time series and the φ_0 and ε_i are the m-dimensional intercept and a residual white nose process respectively. The reduced form of the model is characteristic for including the contemporaneous form of variable y_i to the left hand side of the i-th equation only. For a $VAR(p)$ model to be stationary, it is necessary and must be verified that all roots of the polynomial equation

$$\Phi(z) = I - \Phi_1 z - ... - \Phi_p z^p \tag{2.4.2}$$

fall outside of the unity circle in a complex plane.

However, since many economic time-series tend to behave as integrated (unit-root) rather than stationary processes, the validity of the assumptions underlying classic linear as well as vector autoregressive models becomes questionable at best. As the knowledge of time-series modeling progressed, other approaches have been devised for the cases where non-stationarity of the data cannot be ruled out[1].

One solution, found for example in the paper by Andreica, Aparaschivei, Cristescu, and Cataniciu (2010), is transforming the non-stationary time series, often by differencing, into a stationary one. As a byproduct, such transformation unfortunately leads to a loss of valuable information and changes the interpretation of the estimated coefficients Cipra (2008). This often ultimately also means that by transforming the time series, the regression loses its ability to provide a satisfactory answer for the original research question - the effect in the long run.

2.5 Vector error correction

Engle and Granger (1987) point out, and it is typical especially in economics and finance, that even when certain time-series are individually non-stationary, or integrated in other words, it is often possible to find their linear combination which itself meets the conditions of a stationary process. The authors elaborate on an autoregressive as well as an error-correction representation of such

[1] A combination of the tests proposed by (Dickey and Fuller, 1979) and (Shin and Schmidt, 1992) is commonly advised.

system of co-integrated variables which, contrary to the differencing approach, can be interpreted as a long run equilibrium relationship. Its application was demonstrated, among others, by L'horty and Rault (2004) on French minimum wage and price level data.

To find the co-integration relationships in a set of variables, Engle and Granger (1987) propose a simple *EG test* based on the idea that should the non-stationary variables $y_t \sim I(1)$, $x_{1t} \sim I(1)$, $x_{2t} \sim I(1)$,..., $x_{kt} \sim I(1)$ be co-integrated such that

$$y_t = \beta_1 + \beta_2 \cdot x_{2t} + \beta_3 \cdot x_{3t} + ... + \beta_k \cdot x_{tk} + \varepsilon_t, \tag{2.5.1}$$

the OLS residuals $\hat{\varepsilon}_t$ must pass as an $I(0)$ process. An alteration of the original Dickey-Fuller test with distinct critical values[2] is used to evaluate the null hypothesis of a unit root in the residuals. In practice, however, when there are more than one of such relationships, it becomes increasingly complicated to make sure the correct co-integration relationship is identified. For that reason, the iterative tests proposed by Johansen (1991) are commonly preferred (Cipra, 2008).

One further step in the form of a vector error correction model is needed in order to analyze the response to short term fluctuations. The interpretation of co-integration after all is closer to a long run equilibrium relationship. The error-correction representation derived for the m-dimentional VAR(p) model from equation 2.4.1 takes the form of

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \epsilon_i, \tag{2.5.2}$$

where $y_{1t} \sim I(1), \ldots, y_{mt} \sim I(1)$ and

$$\Pi = \Phi_1 + \cdots + \Phi_p - I = -\Phi(1), \tag{2.5.3}$$

$$\Gamma_1 = -\Phi_1 - \ldots - \Phi_p, \ldots\ldots, \Gamma_{p-2} = -\Phi_{p-1} - \Phi_p, \Gamma_{p-1} = -\Phi_p. \tag{2.5.4}$$

It is essentially derived form the vector autoregressive model for first differences of the y_t time series by including a corrective term for each co-integrative relationship in the system. The corrective term then adjusts the model in case the short term fluctuations in Δy_t cause a divergence from the long term equilibrium, indicated by the co-integrative relationship between the variables.

[2]tabulated by Monte Carlo simulation in Engle and Granger (1987)

2.6 Identification issues

The rather low frequency of price changes as well as the lack of regional variation are common issues for the whole minimum wage literature independently of the chosen empirical approach. Since such variation is crucial for the identification of the estimated models and the lack of it often completely prevents the original model from being estimated, we will briefly describe the various approaches scientists choose in order to deal with the issue.

Several minimum wage studies rely on the within sample regional variation in minimum wage changes, however the usage of such sort of variation is possible almost exclusively in the context of federal states, where the federal minimum wage represents only the lower bound of state-level minimum wages (MacDonald and Aaronson, 2006; Aaronson, 2001; Basker and Khan, 2013; Khan, 2013). Other researchers rather estimate the effect using standard time-series estimation techniques on non-federal/non-US data without the cross-sectional dimension (Andreica, Aparaschivei, Cristescu, and Cataniciu, 2010; Fels and Hoa, 1981; L'horty and Rault, 2004).

In general, scientists have to find a variable that would approximate the movements in minimum wage levels as well as ensure sufficient variation in the cross-sectional dimension of the dataset. Fougère, Gautier, and Le Bihan, who analyze the effects of the changes in SMIC, the French minimum wage, note that a simple regression model between prices and the minimum wage levels suffers from severe collinearity since the absolute majority of wage changes occurs in July and therefore makes it impossible to disentangle the effect of the minimum wage control variable from the July seasonal effect. As a solution, they suggest replacing the original minimum wage variable by cumulative increase in the minimum wage since the last price change.

The single most frequently utilized variable included in place of the original minimum wage movements is the so called *spike* or *fraction at* defined as the proportion of population in a given region that are currently paid the minimum wage. Needless to say, the interpretation of the estimates obtained from such regressions would be somewhat altered, however the estimates are hardly ever published in the raw form. Instead, researchers use an additional simple regression to estimate the minimum wage elasticity of *fraction at* and report the results obtained from the original model multiplied by this constant. Such approach is followed not only by Lemos (2003, 2004a,b, 2005, 2006a) in all of her studies but

also by Dube, Naidu, and Reich (2006).

Fraction affected, defined as the proportion of workers below the newly implemented minimum wage level, and various other specifications of minimum wage gap are sometimes used as an alternative to *fraction at*, even though these measures can suffer from several deficiencies. Brown (1999) argues that these so called "degree of impact" measures do not perform well in periods where the minimum wage stays constant and so its expected impact is declining. Furthermore, Machin, Manning, and Rahman (2003) stress out other possible identification issues arising from the existence of a statistical relationship between the initial level of minimum wage and its increase, in other words they emphasize that the estimated coefficient is a true measure of the minimum wage impact only if the initial wages follow a random walk.

Chapter 3

Data collection

3.1 Choice of literature sample

The initial idea to write a meta-analysis on the price effects of minimum-wage changes was inspired by the work of Sara Lemos (2006b) who in her *"Survey of the effects of the minimum wage on prices"* provided a first-of-a-kind overview and assessment of the existing literature on price effects of minimum wage changes and the empirical findings reported within. Lemos draws attention to the shortage of academic literature reviews (quantitative or qualitative) concerning this subject that persists despite its indisputable contemporary relevance. She claims that besides the 25 publications, the list of which was conveniently included in her narrative review, there were no other works available at the time the study was published. By completing this study we mean to follow up on her narrative review and extend it by more up-to-date publications.

We begin our meta-analysis by searching for studies suitable for our meta dataset. The first search for empirical publications was focused on finding studies listed in the literature overview by Lemos. Unfortunately, only 16 out of the 25 listed papers were publicly available on-line or in the university library. After consideration of the titles and key words of the publications in the original sample, we have decided to look for additional publications via the *Google Scholar* and *EBSCO host* databases using the following search parameters:

- *The publication was found in a search with the words min*, wage, price*/inflation* as parameters of the search.*
- *The publication includes an abstract or an introduction hinting at its empirical*

nature.

- *The title or the abstract of the publication involves the words min* and wage.*

If a study at this point explicitly mentioned the estimation of the minimum wage effect on prices, it was directly included into the broader sample. Nonetheless, after a preliminary search for publications that could extend the list provided by Lemos, we came to realize that almost ten years after the publication of her narrative overview, the field has seen only a modest rise in the number of available studies. In fact, a significant part of the research was carried out by Lemos herself. Moreover, the estimation of the price effect often shows to be of secondary importance for many studies, which is another reason why the words *price** and *inflation** are sometimes absent from the titles and abstracts even in studies which in the end include a relevant estimate.

After a consideration, we decided to somewhat loosen the parameters of the search in order to be able to find a sufficient number of empirical works. Therefore, if the study fulfilled the first three conditions except for explicitly mentioning the estimation of a minimum-wage price effect in the abstract, the fourth criteria was also applied to determine whether the study should be included into the broader sample or not.

- *After a quick visual scan of the text, a regression output table with a label suggesting it included estimates of the desired effect was found.*

Additionally, we have extended our search to studies cross-referenced in our preliminary sample. Altogether, the search has yielded 44 publications, the oldest one dating back to 1976 while the most recent one was not published until 2013. Unfortunately, as we started to examine the individual studies in greater detail we had no choice but to discard almost half of the studies due to various deficiencies.

Some of the studies that were downloaded on the basis of being listed by Lemos in her review lacked the characteristics of an empirical study and often provided only a reproduced estimate of the effect-size rather than estimating the effect all anew[1]. Other studies have proven to be based on computer simulations rather than real-life data, see for example Wolff and Nadiri (1980).

[1]This was the case for the papers written by Wilson, Beach, Rector, and Hederman (1999) as well as MaCurdy and O'Brien-Strain (1997); MaCurdy and O'Brien-Strain (2000); MaCurdy and McIntyre (2001); Falconer (1979); Lee, Schluter, and O'Roark (2000)

Two studies from the Lemos' sample were left out on account of employing an input-output (IO) analysis (Lee, Schluter, and O'Roark, 2000; Lee and O'Roark, 1999). Instead of being the product of an estimation, the spill-over coefficients in IO analysis are computed from the available information on the size of minimum wage changes, the share of workers earning the minimum wage throughout different industries, and the share of labor costs relative to total production costs. As such, the values have no standard errors and are therefore unsuitable for being used in meta-analytical reviews.

The rest of the issues that led to excluding several papers were somewhat more subtle. Cuong (2011), Wadsworth (2010) as well as Draca, Machin, and Van Reenen (2005) used a dummy variable regressor indicating that an increase in the minimum wage level occurred in a particular period. Letting the minimum wage change be represented by a solely qualitative variable naturally leads to an altered interpretation of the regression coefficient incompatible with the rest of the studies in our sample. A reversed problem appeared in Werner, Sell, and Reinisch (2013) who chose to estimate a probit model with a dummy dependant variable indicating whether an increase in price level has occurred in reaction to changes in the minimum wage level or not. For the sake of avoiding distortions in the collected data, we decided to exclude studies that applied either of the approaches.

3.2 Treatment of exceptions

Once the final sample of relevant literature is assembled, we arrive at the core and most time-consuming part of the analysis. The relevance of any potential findings critically hinges on the quality standards met when building the underlying dataset. In compliance with the approach laid out in Havranek and Irsova (2011), we collect all of the estimates reported in our final literature sample. Narrowing the collection process only to those estimates deemed to be the best by us or the author of the publication might have counterproductive impact, introducing additional bias into the meta-dataset. Similarly, representing the studies by the average estimate reported within the studies has a significant downside when the amount of information needlessly lost by the averaging is considered.

In general, we focus on studies that provide information on both, the size of the empirical effect and its precision. For papers, where only the information

on the level of statistical significance was found (Fougère, Gautier, and Le Bihan, 2010), we have followed the approach for imputing p-values outlined by Greenberg, Michalopoulos, and Robins (2003). In principle, when an estimate was reported as significant on 5% level but not on 1% level, we have assumed its p-value was equal to 0.03, which is the exact midpoint of the interval between the p-values representing the borderline significance levels. The same logic was applied to estimates significant on 10% level resulting in an imputed p-value of 0.075. Finally, for insignificant estimates, we have assumed the p-values lied within the interval from 0.1 to 0.5, again represented by its 0.3 midpoint.

A number of studies work with a lag or a lead structure of the minimum wage variable to capture temporal patterns in the data that are more complex than a mere contemporaneous impact. As a result, the total effect splits between several variables. We have collected the estimates of the total effect given by the sum of all right-hand side minimum-wage-impact coefficients as well as their approximated standard errors. Since using the sum of reported standard errors as a measure of variation in the combined effect is not statistically correct, an approximation is realized by using the *Delta-method* based on the first-order Taylor sires expansion of the summed coefficients (Powell, 2007). The downside of the *Delta-method* approach in this case is, that it implicitly assumes we are working with independent random variables. By applying the method to compute the variance of a sum of two or more possibly correlated variables, we are knowingly introducing an error into our collected data.

3.3 Coding of variables

The essential question we need to ask ourselves at this point is what information about the primary studies are we about to code into our dataset. Besides the effect size and a measure of its precision, both imperative for the purposes of meta-analysis, Stanley and Jarrell (1989) and Havranek and Irsova (2011) give us an idea of what meta-independent variables might help us uncover patterns in scientific empirical literature. In the remainder of this section, we will explain the meaning and discuss the motives that led us to collect the particular variables as they are used in the remainder of the thesis.

3.3.1 Effect size and its precision

As was already stated before, each estimate of the effect size reported in the literature sample needs to be weighted by its precision. Luckily, most of the studies report the standard deviations and/or the t-statistics alongside the estimates.

Collecting the effect size data is further complicated by the need for comparability of the estimates within and between the individual studies. Stanley and Jarrell (1989) point out it is essential that the effect size is a standard measure of the empirical effect which is assumed constant across the literature. In reality, this means that special attention must be paid to the precise functional form of the regression variables so that coefficients with different interpretation, such as semi-elasticities, elasticities, and level-level estimates are not treated as equal.

Unfortunately, further restriction of our dataset to studies where the functional form of the empirical equation is identically defined would in the best case mean discarding almost 38% of our observations. For that reason we chose to abandon the idea of using the otherwise preferable effect-size that would measure the genuine economic effect and turned to analyzing the statistical relationship instead. The collected data was used to calculate the partial correlation coefficient r and its standard error Se as

$$r = \frac{t}{\sqrt{t^2 + df}} \qquad Se = \sqrt{\frac{1 - r^2}{df}} \tag{3.3.1}$$

For the remainder of the thesis, partial correlations are used (Stanley and Doucouliagos, 2012, p.23).

3.3.2 Meta-independent variables

Quality characteristics

Naturally, not all studies can meet the same quality standards. For example, not all of them were submitted and accepted for publication in academic journals. Even so, not all academic journals are peer reviewed and have comparable entry requirements in terms of research quality.

Our dataset includes altogether four variables to proxy for the varying quality of individual studies. These are namely the number of RePEc and Google Scholar citations recorded as of June 2014, RePEc impact factor of the outlet, and

a dummy variable assuming the value 1 in case the study has been in fact published and 0 otherwise. Both RePEc and Google Scholar citation counts were discounted by study age to ensure the parameters were not overestimated for earlier studies.

Data characteristics

Adhering to the advice of Stanley and Jarrell (1989), who emphasize that information on the primary datasets should be accounted for in all meta-analyses, we created five dummy variables together controlling for the type of dataset in three distinct ways.

First, we use two dummy variables to record whether the primary dataset consists of time series or panel data. Second, we create a dummy variable assuming the value 1 for datasets that were manually collected by the author himself specifically for the purposes of his research. Finally, we include two dummy variables carrying information on the level of aggregation of prices in the primary dataset. The reason we create such a variable at all is that the size of the empirical effect recognized in aggregated data may prove to be lower compared to when item-level data is used. Despite being sector-wise significant, the impact on the aggregate price level may dissolve if only few sectors are notably affected by the new minimum wage regulation. Moreover, studies employing item-level data commonly focus on the prices of goods the production of which is intensive in cheap labor. It is expected that such studies are more likely to report a larger estimate of the effect-size as a result.

Estimation characteristics

Being used to obtain more than 66% of all reported estimates makes the ordinary least squares method by far the most frequent estimation procedure to be encountered in our branch of minimum wage literature. Next in line is the weighted least squares regression which produced approximately 20% of the effect-size coefficients. Other less common yet still represented methods include instrumental variable, median and robust regressions, and maximum likelihood estimators.

For each group of estimators, a separate dummy variable was created to control for the effect of method heterogeneity.

Specifics of the dependent variable

Another potential source of heterogeneity lies in the varying composition of the price-level-related dependant variable used in the primary studies.

When we talk about price-level, one would usually assume we are referring to the price of a standardized bundle of consumption goods. In reality, three quarters of estimates reported in our studies apply to food industry only, especially to restaurants and fast-food chains known for paying out the minimum wage to a relatively large percentage of the work force in comparison to other sectors. This makes the industry particularly appealing to researchers who hope to find a significant and pronounced effect as a result of the amplified effect of labor costs' shifts. We added one general dummy to control for all food-related estimates along with two additional dummies allowing for a detail decomposition between restaurant and fast-food prices.

Several studies[2] focus on the impact of minimum wage rate shifts on different social classes. By testing the underlying hypothesis that a larger portion of the consumption bundle of those less wealthy consists of cheap-labor-intensive goods such as food, we aim is to show that not only does the rise in minimum wage level have a counter-productive effect on the living standards of lower social classes through increased inflation, but also that the inflation is experienced more intensively by those whom the change in policy was meant to help in the first place. We stored the information on social affiliation of estimates into two dummy variables representing its lower and upper quantiles.

Underlying model

Our literature sample is characterized by being assembled from studies of varying quality, not all of which abide by the leading practice of including detailed information on the utilized econometrical methods and underlying models. After all, the correct specification and not the actual choice of the model should be the driving factor determining the regression outcomes. Evidence that the estimated models have passed specification tests are unfortunately rare, which is why the choice of empirical model may skew the obtained results one way or the other.

The information on the underlying empirical model is collected only for studies that provide a clear idea of how the underlying model was constructed and

[2] MacDonald and Aaronson (2006, 2000); Aaronson (2001); Aaronson, French, and MacDonald (2008); Lemos (2003, 2004a)

estimated. That way our dataset makes a distinction between each of the four types of empirical models described in section 2.

Real factors: time and geographical affiliation

Taking into account the trade-off relationship between parsimony and retaining as much information as possible, we sort the primary studies according to the geographical origin of the data used within into four larger groups in an attempt to avoid the need for creating a separate dummy for each state. This way we manage to distinguish between estimates extracted from North American, South American, European, and Australian data[3]. Separate dummies, which partially overlap with the previously mentioned controls, are used of US data and OECD membership.

Just as important is the effect of temporal heterogeneity especially considering an unrestricted meta-dataset where primary studies elaborate on data that cover a period long enough to allow the market to undergo a structural change. Recording the dataset span measured in months and the average year standardized by subtracting its mean 1995.04 from each observation enables us to cover for the caveat by controlling for time trends in the data.

Controls in primary regressions

The last set of information collected for our MRA addresses the need for describing the primary regression model in a greater detail particularly in terms of what explanatory variables besides the minimum wage shift control it consists of, and of the lag and lead structure thereof.

For that purpose, each observation lists the number of lagged and leading months of the minimum wage change control included in the primary regression along with two dummy variables to identify estimates that either are or incorporate lagged and leading effects. Furthermore, two additional dummy variables distinguish between estimates of the contemporaneous effect and those that capture the response of the price level in the long term.

Hand in hand with collecting estimates from regressions with a lag and lead structure comes the necessity of using the Delta method for aggregating the partial coefficients into a single point estimate to get an overall effect that can be utilized in an MRA. Since the usage of the method possibly has a distortive

[3]Note that the last two make up for less than 10% of the total sample

effect on the standard errors and, like any other systematic influence, should be controlled for, we create an auxiliary dummy variable attaining the value 1 every time Delta method is used.

As for the remainder of the controls, the decision of whether or not to note down their presence in the primary regressions was based on the practical frequency of their appearance in the literature sample.[4] In over 70% of all regressions researchers control for the impact of seasonality and location; the presence of either of the controls is recorded in two dummy variables *seasonalityctrl* and *cityareactrl*. Aaronson (2001), Lemos (2004a) alongside with other studies in total constituting about one third of our meta dataset acknowledge the potential impact of *input-price* shocks on the observed inflation level. By including an input-shock control, the studies separate the minimum-wage-shock-induced changes in the prices from the overall effect of raising production costs. Slightly less common in absolute numbers, *mean-revertion* is controlled for in about 15% of studies, including Aaronson, French, and MacDonald (2008), Lemos (2004a), and MacDonald and Aaronson (2006). It refers to a situation where the observed variable has a tendency to revert back to its average value in the long term. If that is true in addition to the empirical model being designed to reflect the long term impact of minimum wage changes as opposed to the short term hikes until the mean reversion kicks in, the mean reversion could put a downward bias on the the reported estimates. Finally, the dummy variable *unemployment change control* tags the 10% of estimates that were obtained from regressions where changes in unemployment level appear among the independent variables and their effect on the price inflation is thus held constant in the regression.

A description of the complete dataset is included as part of Appendix B. Table 3.1 offers a brief overview of the additional controls for better clarity.

[4]Creating separate dummy variables for controls with seldom occurrence could introduce further data issues into our sample, specifically concerning the identification and collinearity in our panel data models.

Table 3.1: Additional controls in primary regressions

Variable	(Definition)	Dummy	(Mean)	(Std. error)
short	=1, if estimate relates to contemporaneous effect	yes	0.356	(0.479)
long	=1, if estimate relates to longer-term effect	yes	0.644	(0.479)
lag	=1, if estimate is or includes lagged effect	yes	0.488	(0.505)
lead	=1, if estimate is or includes leading effect	yes	0.418	(0.494)
#lag	no. of lagged months covered by the estimate	no	1.5946	(2.226)
#lead	no. of leading months covered by the estimate	no	1.222	(1.542)
combination	=1, if delta method was used to produce the estimate	yes	0.235	(0.424)
cityareactrl	=1, if city area control present in primary regression	yes	0.727	(0.446)
unempchangectrl	=1, if unemployment change control present in primary regression	yes	0.096	(0.295)
seasonalityctrl	=1, if seasonality control present in primary regression	yes	0.797	(0.402)
meanrevertingctrl	=1, if mean reversion control present in primary regression	yes	0.156	(0.363)
inputpriceshockctrl	=1, if input-price-shock control present in primary regression	yes	0.292	(0.445)

Source: own analysis

3.4 Initial analysis of the data

To mention at least the basic specifications of our meta-dataset, the final matrix numbers 469 observations reported altogether in 23 empirical studies published between 1981 and 2013. The primary data used by the studies cover a period of almost 60 years starting from 1953 and spanning over to 2012. The median dataset dates between 1989 and 2000. The highest number of observations collected from a single study was obtained from the article by Aaronson (2001) whose paper contained an impressive total of 164 estimates of the price effect of minimum wage changes. The second largest number of estimates published in one study was 118. On the other hand, four studies contained only one suitable estimate each. The imbalance in the contribution of individual studies towards the size of the population is perhaps best illustrated by the relatively large difference between the mean and median number of estimates reported per study which adds up to 8 and 20.391 estimates respectively.

In addition to that, there is a large variability in the overall size of the underlying datasets of primary studies. The largest panel data study by Aaronson, French, and MacDonald (2008) works with 71,077 individual observations, while the time-series dataset of the smallest study by Andreica, Aparaschivei, Cristescu, and Cataniciu (2010) contains only 43 observations. It should be noted that besides giving an idea of the precision of the estimated results, the size of the dataset is to a certain extent predetermined by the type of the primary data (longitudinal vs. time-series) and, therefore, also by the methodology used for

the estimation.

We have calculated the arithmetic and weighted means of the partial correlation coefficients using both the inverted standard error and square roof of the number of observations in the primary study as weights. Both weighted means equal approximately 0.027 with a standard deviation of 0.044 and 0.0054 respectively. The notable difference between the weighted and arithmetic means, the first amounting to nearly the double of the later, can be understood as a preliminary test of certain imbalance in the data.

Table 3.2: Summary statistics

Variable	**Min.**	**Max.**	**Mean**	**Median**
Partial correlation	-0.113	0.273	0.042	0.0295
♯ estimates per study	1	164	20.391	8
♯ observations per study	43	71077	7842.414	3082
Dataset span (years)	0.583	26	12.251	12
Statistic	**Value**		**Std. dev.**	
Weighted mean ($\frac{1}{Se}$)	0.0273		0.0442	
Weighted mean ($\sqrt{n}$)	0.0272		0.0054	
Arithmetic mean	0.0422		0.0054	

Source: own analysis

Once transformed into partial correlations, none of the observations appears to lay markedly off the main data cluster. In order to support the claim with objective criteria, we applied the built-in Stata procedure for identification of outliers in multivariate data proposed by Hadi (1992) to the variables representing the partial correlation coefficients and its precision. Using the default 0.05 significance level for outlier cut-off, only a single observation was labeled by the procedure as an outlier. With respect to the lack of outlying observations and also to the argument made by Doucouliagos and Stanley (2009) who advocates maintaining the integrity of the collected data, claiming that imprecise outliers have little weight and their effect on the analysis' outcome is thus limited, we decided to use the full data sample in the subsequent analysis. Table 3.2 contains an overview of the descriptive statistics of the collected dataset.

Chapter 4

Meta-analysis: methodology

With the ever growing volume of empirical research the ability to integrate and synthesize its findings constantly gains on importance. The role has been traditionally assumed by literature surveys, nonetheless these are often criticized for deliberately excluding findings contradictory with the author's beliefs and for being unable to unambiguously define a framework for evaluating and interpreting the published results.

In 1976, Glass used the term meta-analysis to refer to the analysis of analyses, a way of assessing a bulk of empirical results from individual studies at a time for the purpose of condensing their findings into a single statistic. Initially, the method has gained a toehold in medical, psychological, and educational research to be later adopted also by economists following the publication of Stanley and Jarrell (1989). Specifically, Stanley and Jarrell (1989) elaborate on a quantitative approach to literature review, the meta-regression analysis (MRA), that allows us to explore the origins of the variation in research results in the same manner the primary studies analyze the examined phenomena. Besides that, meta-analysis can help us uncover and correct patterns stemming from the incentives faced by scientists during the publication process. The rest of the section outlines a fragment of the meta-analytical tool-set used further on in the thesis for studying the effect changes in the minimum wage rate have on the price level in the economy.

4.1 Publication selectivity

People respond to incentives. Gregory Mankiew lists the thesis as number four among the fundamental principles of economics. Academic publications and research in general are not exempt from the rule. In a system, where journals pre-select research papers for publication based on conformity with commonly accepted theories and statistical significance of reported results, publication selectivity is likely to be abundant.

Card and Krueger (1995b) point out that scientists, whose career depends on how many of their studies are published in prestigious academic journals, sometimes incline to specification searches in order to satisfy the demand for statistical significance. In other words, they tend to estimate their models repeatedly using different specifications, until they obtain estimates large enough to compensate for their large standard errors (Havranek and Irsova, 2011). Such practices will project in the meta-dataset by a disproportionately strong presence of estimates significant precisely on the 5% significance level which is typically used in economics and other social sciences as the threshold for statistical significance.

On top of all the bias caused by the journals' selection of papers, scientists themselves may reinforce the distortion by self-censorship through not submitting articles which are in disagreement with the standard theory or contain only insignificant results. The expectation that such articles would be rejected for publication anyway leads many scientists to discard otherwise standard research results. Such behavior was established in meta-analytical literature under the term "file drawer" problem (Rosenthal, 1979) .

Meta-analysis allows us to identify the magnitude of the true underlying effect by eliminating the distortive effects of publication selection in academic literature.

4.1.1 Graphical approach to publication selectivity

Funnel graph

Presumably also for its simplicity, the funnel graph ranks among the most widely used statistical tools designed for detecting discrepancies in meta-analytical datasets. The name "funnel" relates to the expected shape of the graph, which in the ideal case should resemble the upturned kitchen utensil.

Generally speaking, it is a simple scatter plot of the investigated effect on the horizontal axis versus a measure of its precision on the vertical one. The two most utilized measures of precision are the inverse of standard errors of the estimates and the square root of the number of observations in the original study. If there is an underlying true effect, published estimates should vary randomly and symmetrically around it, unless the plot is skewed by the presence of publication selectivity. Estimates reported in studies with lower number of observations and/or lower precision should be relatively more spread out at the base of the graph, whereas the precise estimates should form the slender peak of the inverted funnel together creating the typical shape of the graph. In case the data is biased, the graph will be unbalanced in favor of one or the other direction (Stanley and Doucouliagos, 2010).

In general, funnel graphs are designed for detection of a particular type of publication selectivity (bias) characteristic by overweighing the funnel in one or the other direction. This is the so called *type I publication bias*, or directional bias, and it arises when journals and researchers become influenced by their expectations of the size and direction of the estimated effect and thus incline to censorship or self-censorship to obtain non-contradictory results (Doucouliagos and Laroche, 2003). The directional type of bias is also the more harmful one as it distorts the expected magnitude of the effect commonly measured as the average of the estimates over the population of studies.

Stanley (2005) describes also some of the limitations of funnel plots. To mention at least the most prevalent one, it should be pointed out that publication bias may not be the only source of possible funnel graph asymmetry. Meta-analysis a priori assumes the existence of a single "true effect"' across the whole body of literature. Even if there is no "true effect"', the empirical estimates should be randomly distributed thus preserving the funnel graph symmetry. However, this assumption is false if the true effect actually varies with different countries, datasets or time periods for which the effects were estimated. In that case, if the population of studies is somewhat skewed (eg. most of the estimates are related to US data), it may happen that the final funnel graph will appear asymmetric for different reasons than the presence of publication selectivity.

Galbraith Plot

Type II publication bias arises when studies with significant results are more likely to get published compared to other studies of similar quality but lack of significant estimates. Generally speaking, it is considered to be a less harmful form of distortion as it causes only excess variation in the estimated effects without actually affecting their average magnitude.

Galbraith plot is a scatter diagram depicting the pairs of standardized estimated effect such as t-values on the vertical axis and their precision on the horizontal axis. The plot is in fact only a modification of the funnel graph, where the axis are switched and the values on the vertical axis are standardized, i.e. transformed to

$$t_i = \frac{effect_{\mathrm{i}} - true\ effect}{Se_i} \tag{4.1.1}$$

As a result, the plotted distribution becomes comparable to standard normal distribution. The true effect must be either estimated, or alternatively, a weighted average of the estimates can be used as its approximation.

Under the condition that the likelihood of publishing results of empirical studies does not depend on their statistical significance, the standardized estimates should be randomly distributed and we would expect only about 5% of the observations to lie out of the 95% confidence band around 0. That is to say, only 5% of the data points will not satisfy the condition

$$\left| \frac{effect_{\mathrm{i}} - true\ effect}{Se_i} \right| \leq 1.96 \tag{4.1.2}$$

4.1.2 Formal approach to publication selectivity

Despite the ease with which the graphical analysis can hint at the presence of certain patterns in the data, visual analysis of figures remains very subjective and therefore insufficient as evidence for claiming the existence of either type of publication bias. In regard of that, it should be treated only as an indication of potential data problems and any findings should be validated in objective statistical tests.

An effective way of testing for the presence of publication selectivity directly follows from the graphical analysis of a funnel plot. Equation 4.1.3 works on the presumption that estimates of the observed effect are randomly distributed

around the "true effect"' β unless they are systematically discarded on the basis of being insignificant or in contradiction with the theoretical expectations. If the later is true, the following regression will unveil a correlation between the reported estimates and their standard errors (Card and Krueger, 1995b; Havranek, Irsova, and Janda, 2012; Stanley, Doucouliagos, and Jarrell, 2008).

$$b_j = \beta + \beta_0 Se_j + u_j, \qquad j = 1, ..., n \tag{4.1.3}$$
$$u_j \mid Se_j \sim N(0, \delta^2)$$

Here b_j denotes the j-th collected estimate, Se_j is its corresponding standard deviation, β represents the underlying effect, and u_j is the remaining disturbance term. The connection to the funnel plot follows from using the inverse of the precision values Se_j and rotating the axes to display the size of the effect on the vertical one.

To make matters more complicated, the fact that the estimates of b_j and their standard errors generally come from different empirical studies casts doubts on the OLS assumption of homoscedasticity of the disturbance term u_j. Luckily, a quick-fix solution can be easily applied by estimating the equation in the weighted least squares (WLS) form using the evident measure of heteroscedasticity Se_j as weights. The WLS estimation leads to the equation

$$\frac{b_j}{Se_j} = t_j = \beta_0 + \beta \frac{1}{Se_j} + \nu_j, \qquad j = 1, ..., n \tag{4.1.4}$$
$$\nu_j \mid Se_j \sim N(0, \sigma^2)$$

with t_j being the t-values of the reported estimates. In the specification 4.1.4, the interpretations of the slope and the intercept coefficients are reversed, with β_0 standing for the publication bias this time. Testing the null hypothesis $H_0 : \beta_0 = 0$ after running an OLS regression of equation 4.1.4 is equivalent to testing the asymmetry of the funnel graph and is commonly known under the name Funnel asymmetry test (FAT) .

Besides the FAT test, equation 4.1.4 conveniently incorporates a basic test of the presence of a genuine effect beyond publication selectivity. With its t-test often referred to as the precision effect test (PET) , the coefficient β can be interpreted as an estimate of the authentic effect corrected for the publication selectivity filtered out by the intercept term (Stanley, 2008; Doucouliagos and Stanley, 2009).

Further examination of the statistical properties of the model reveals that heteroscedasticity of the error term, as notorious as it is, is not the only shortcoming in need to be dealt with.

Since the independent variable in equation 4.1.4 itself needs to be estimated and therefore contains errors-in-variable bias, the OLS estimators of β_0 and β turn out to be biased and inconsistent (Wooldridge, 2012). Possible remedies include replacing the $\frac{1}{Se_j}$ in equation 4.1.4 by a variable free of the errors-in-variable bias (eg. degrees of freedom or $\sqrt{n}$), or using instrumental variable (IV) estimation instead of OLS. If the correlation between $\sqrt{n}$ and $\frac{1}{Se_j}$ is sufficiently high and provided $\sqrt{n}$ is not correlated with the estimation error, $\sqrt{n}$ appears to be a convenient IV for $\frac{1}{Se_j}$. An ordinary t-test of the 2SLS IV estimator of β_0, known as the funnel asymmetry instrumental variable estimator (FAIVE) , presents a consistent alternative to FAT. Furthermore, by employing the White's heteroscedasticity-consistent formula for the 2SLS IV estimator, we arrive at a consistent, heteroscedasticity-robust, yet not necessarily unbiased alternative to FAT, the FAIVEHR test (Stanley, 2005) .

One should be aware, that all of the models described above implicitly treat the observations as if they were a product of random sampling - an assumption rather bold and applicable only in case each study is represented in the dataset by a single effect-size or in our case by a single partial correlation. Even if this was true, some studies might still share the same author, primary dataset, or estimation technique all of which further undermine the validity of this assumption by introducing a dependant relationship in the data. On the other hand, discarding part of the effect-sizes reported in each study in favor of a preferred effect-size can be a source of additional bias and leads to loss of valuable information. Nevertheless, even these issues with seemingly opposing solutions can be addressed in a statistically correct manner. Havranek, Irsova, and Janda (2012) suggest using the mixed-effects multi-level model specified below in equation 4.1.5 as one of the possible treatments of the issue.

$$t_{ij} = \beta_0 + \beta \frac{1}{Se_{ij}} + \mu_j + \epsilon_{ij}, \qquad j = 1, ..., n \tag{4.1.5}$$
$$\mu_j \mid Se_{ij} \sim N(0, \psi) \qquad \epsilon_{ij} \mid Se_{ij} \sim N(0, \theta)$$

The subscripts i and j denote estimate and study or author respectively. The specification 4.1.5 decomposes the overall error term from equation 4.1.4 into study-level random effect μ_j and estimate-level disturbances ϵ_{ij}, both of which

are assumed mutually independent and therefore additive in such way that $Var(\nu_{ij}) = \psi + \theta$. Disturbance terms ψ and θ represent the between and within study heterogeneity respectively. However, with ψ approaching zero, the benefit of using the mixed-effects multi-level model instead of OLS diminishes.

The "mixed-effects"' in the name of the model follows from combining the fixed effect in β with the random component μ_j. The method is closely related to the random effect model frequently used for panel data analysis. The use of the maximum likelihood estimator instead of generalized least squares renders it more suitable for meta-analytical research by allowing for more flexibility in nesting multiple random effects (e.g. study, author, geographical affiliation) and by compensating for the imbalance brought by studies containing a substantially higher number of estimates (Havranek, Irsova, and Janda, 2012).

Alternatively, the distortional effects of within study dependance can also be eliminated by using clustered data analysis methods (Doucouliagos and Stanley, 2009). We will proceed by combining the two approaches as a robustness check since a large difference between the estimates obtained through clustered OLS and mixed effects could point towards violation of the exogeneity assumption underlying the mixed effects model.

4.1.3 Correcting for publication selectivity

A natural question arises in case the tests described in section 4.1.2 confirm the presence of publication selectivity - how do we rid the collected data of its distorting effects? Let us now modify the original FAT test adding absolute values to the left side of the equations 4.1.4 and 4.1.5 (Stanley, 2005)

$$\left|\frac{b_j}{Se_j}\right| = |t_j| = \beta_0 + \beta\frac{1}{Se_j} + \nu_j, \qquad j = 1, ..., n \tag{4.1.6}$$

and

$$\left|\frac{b_j}{Se_j}\right| = |t_{ij}| = \beta_0 + \beta\frac{1}{Se_{ij}} + \mu_j + \epsilon_{ij}, \qquad j = 1, ..., n \tag{4.1.7}$$

When statistical significance of the reported results is used by journals as a quality measure irrespective of the direction of the estimated effect, a systematic relationship emerges between the absolute value of the magnitude of the effects and their standard errors. Shifting each observation towards the *true effect* by the

factor $\beta_0 Se_i$ will clear the data of type II publication selectivity. In general, such transformation of the data helps achieve the desirable shapes of both funnel graphs and Galbraith plots.

4.2 Heterogeneity among studies

Needless to say, there is more to meta-analysis than controlling for publication selectivity. It helps us decipher why scientists often reach very different results despite trying to answer the same question. One solution to the puzzle lies in the decomposition of the observed variation between the product of an unexplainable random disturbance term and the part of the variation that can actually be explained by factual attributes of the primary studies.

Besides reflecting the size of the true underlying effect and possibly being distorted by publication selectivity, empirical results are also affected by many other factors such as geographical affiliation of the data used for estimation, underlying theoretical models and their assumptions, controlling for or omitting important variables, empirical approach in general and many others.

Let us start from the heteroscedasticity-adjusted FAT-PET model in equation 4.1.4. By adding further meta-independent variables Z_{jk} to the right hand side of the regression we arrive at an extended model 4.2.1 in which the inherent heterogeneity between our studies is already accounted for (Stanley, Doucouliagos, and Jarrell, 2008).

$$\frac{b_j}{Se_j} = t_j = \beta_0 + \beta \frac{1}{Se_j} + \sum_{k=1}^{K} \frac{\alpha_k Z_{jk}}{Se_j} + v_j, \qquad j = 1, ..., N \tag{4.2.1}$$
$$v_j \mid Se_j \sim N(0, \sigma^2)$$

The downside of OLS (or WLS) estimation is that it weights each estimate equally, which in the likely case that some studies report more estimates than others leads to their overrepresentation. Drawing on the example of Havranek and Irsova (2011) we estimate the regression also in its mixed-effects multilevel specification described in equation 4.2.2, that way attributing the weights across our studies more evenly.

$$t_{ij} = \beta_0 + \beta \frac{1}{Se_j} + \sum_{k=1}^{K} \frac{\alpha_k Z_{jk}}{Se_j} + \mu_j + \epsilon_{ij}, \qquad j = 1, ..., N \tag{4.2.2}$$

$$\mu_j \mid Se_{ij} \sim N(0, \psi) \qquad \epsilon_{ij} \mid Se_{ij} \sim N(0, \theta)$$

The i-th study is denoted by the subscript i and so it follows that its random intercept is given by $\beta + \mu_j$. One thing to keep in mind when modeling heterogeneity using the mixed-effects multilevel model is that the exogeneity assumptions must be fulfilled in order to produce an unbiased estimate of the slope coefficients associated with the meta-independent variables. This means that any unobservable estimate or author/study level characteristics relegated to the error term must not be correlated with either of the observable estimate or author/study level characteristics stored in the independent variables. Additionally, correlation among the meta-independent variables can cause the bias to spill-over among other coefficients in the regression (Nelson and Kennedy, 2009). For that reason any inference should be drawn with extra caution especially when substantial differences from the OLS estimates are detected.

Chapter 5

MRA of the minimum wage effect on price level changes

In the following chapter, we focus on a practical application of the formerly described meta-analytical techniques to study the literature addressing the effect that changes in the minimum wage rate have on the price level movement. We start by using the collected data to test for the presence of publication selectivity and later on enhance our analysis by employing a meta regression analysis to identify more subtle sources of heterogeneity among the reported estimates.

5.1 Graphical analysis of publication selectivity

Figure 5.1 contains a funnel plot for the full dataset with two vertical reference lines depicting distinctive values of the partial correlation displayed on the horizontal axis. The dashed line at value 0.027 corresponds to the weighted average of partial correlation coefficients computed with both the inverse of standard errors and square root of the number of observations in primary studies as weights (the values are identical after rounding up to 3 decimal places). We have placed a second reference line representing zero partial correlation into the picture in order to highlight the considerable asymmetry in the data in which zero to positive values have a relatively stronger presence while the negative partial correlations are considerably under-represented.

Even though the transformation of our data into partial correlations is unavoidable to ensure the comparability of all collected estimates, it is also the culprit of another limitation evident in the figure 5.1. The size of partial cor-

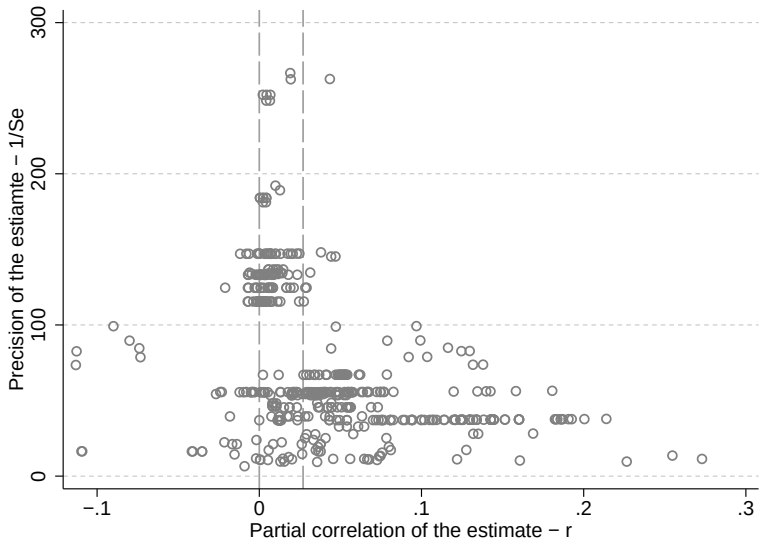

Figure 5.1: Funnel graph: full sample

relation and its standard error is strongly driven by the size of the primary dataset, which often does not vary for most of the estimates reported within one study. This in combination with the relative imbalance in the representation of certain studies in our meta-dataset causes the clustering of data points into horizontally aligned lines.

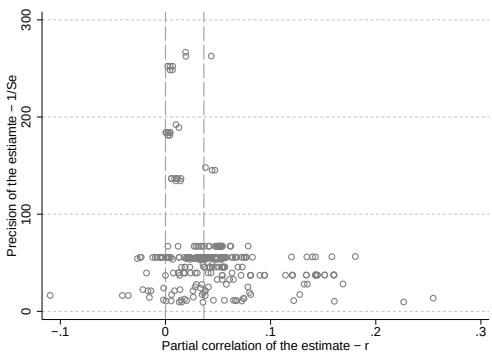

Figure 5.2: Funnel graph: published estimates

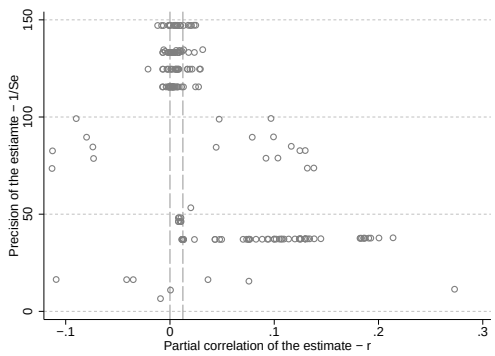

Figure 5.3: Funnel graph: unpublished estimates

Figures 5.2 and 5.3 in which we have made a distinction between published and unpublished studies[1] point to several notable trends. Namely, the overall as well as the individual precision of the partial correlations collected from

[1]Two reference lines were inserted in each of the funnel plots - one corresponding to zero partial correlation and second to the weighted group averages of 0.0365 and 0.01924 respectively, both computed with $\frac{1}{Se_i}$ as weights.

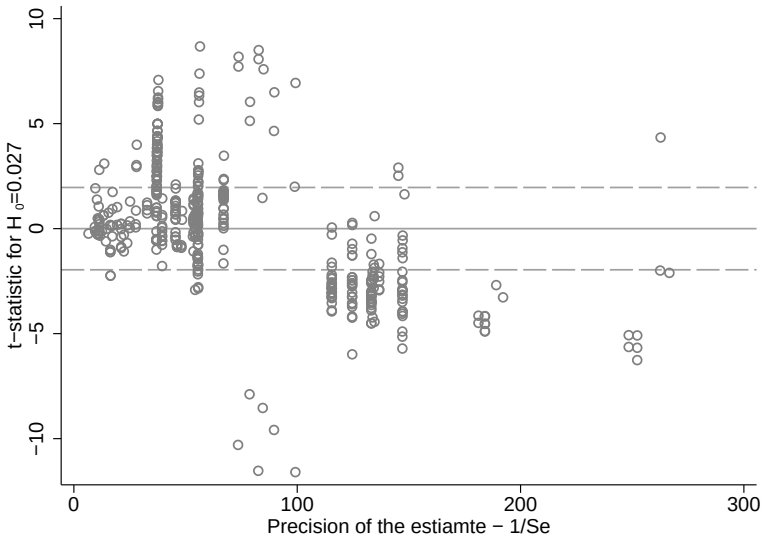

Figure 5.4: Galbraith plot: type II publication bias

unpublished studies are considerably lower and, even more interestingly, the initial asymmetry of the funnel at first sight appears to be less prevalent in the unpublished studies. Despite that, it remains unlikely that further analysis would uncover a significant difference between the two groups on account of the skewness caused by a single unpublished study authored by Lemos (2003). The estimates reported in the study form a horizontal cluster to the right of the reference line with precision nearing the value 37.

In the context of our meta-analysis, assuming that the genuine partial correlation equals to the overall weighted average of 0.027, we would expect no more than 5% of the transformed t-statistics to exceed ± 1.96 in the absence of *type II publication bias*. However, the Galbraith plot of our full sample included in figure 5.4 shows that the proportion of data points outside the outlined confidence band is substantially higher. For consistency and also due to the relatively marked presence of two distinct data clusters in the overall figure, we have again constructed separate figures for published (figure 5.11) and unpublished (figure 5.12) studies[2]. In the case of published studies we find that 78 out of 270 t-statistics exceed the critical value associated with $\alpha = 0.05$. As a result, the null hypothesis of the Chi-square goodness of fit test, which describes the expected distribution of the inspected variable in the absence of *type II publication selection*, is firmly rejected ($\chi^2 = 324.39$, p-value < 0.0001). Likewise, the distribution of t-

[2]Weighted averages, computed with $\frac{1}{Se_i}$ as weights, equaling to 0.0365 and 0.01924 for published and unpublished studies respectively, have been used to obtain the normalized data.

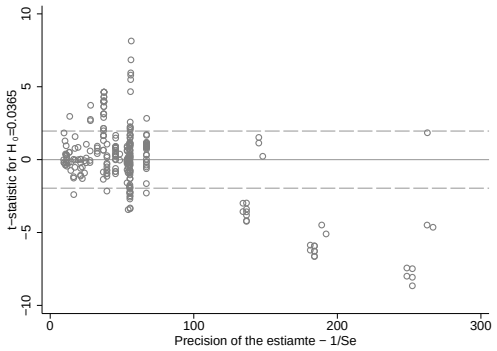

Figure 5.5: Galbraith plot: published studies

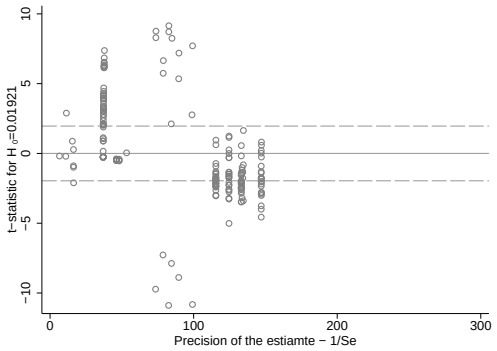

Figure 5.6: Galbraith plot: unpublished studies

statistics collected from unpublished studies points towards the presence of *type II publication selection* with 155 out of the 199 t-statistics exceeding the critical value for $\alpha = 0.05$. The Chi-square goodness of fit test rejects its null hypothesis even stronger ($\chi^2 = 1167.47$, p-value < 0.0001) in this case.

5.2 Formal analysis of publication selectivity

Table 5.9 displays the output of the linear regression from equation 4.1.4 based on the computed partial correlations and their standard errors. Additionally, we have interpreted the data as a longitudinal dataset and estimated the same model with study-level fixed effects. The results are reported in columns 4-6 of table 5.9.As a robustness check, the ordinary least squares and fixed effects estimations reported in the first six columns of the table were complemented by instrumental variable regression in which we have used the variable $\sqrt{nobs}$[3] as an instrument for $\frac{1}{Se_i}$. Finally, we have run separate regressions for the complete dataset, as well as for the published and unpublished studies, all of which provided robust statistical evidence for the asymmetry of their corresponding funnel plots.

[3]Number of observations in the primary study

Table 5.1: Tests for publication selection bias

Variable	**OLS**			**FE**			**FAIVEHR**		
	All	Published	Unpublished	All	Published	Unpublished	All	Published	Unpublished
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β_0 (selectivity)	2.4411***	1.8617**	3.8095*	2.0684***	1.2979***	3.9929***	2.449***	1.867***	3.8226***
	(4.49)	(3.98)	(4.12)	(4.60)	(3.79)	(3.57)	(11.91)	(8.75)	(8.38)
$1/Se_i$ (effect)	-0.0061	0.0052	-0.0225**	-0.001	0.0147**	-0.0245*	-0.0062*	0.0051	-0.0226***
	(-0.93)	(0.88)	(-3.06)	(-0.17)	(2.66)	(-2.04)	(-2.49)	(1.47)	(-6.46)
N	469	270	199	469	270	199	469	270	199
R^2	0.0147	0.0172	0.1057	0.0147	0.0172	0.1057	0.0147	0.0172	0.1057

Significance levels: †:10% *:5% **: 1% ***: 0.1%
Dependant variable: t-statistic of the partial correlation of the effect of minimum wage on inflation
t-statistics of the estimates in parenthesis (study level clustered errors)
OLS - ordinary least squares estimation
FE - study-level fixed effects estimation
FAIVEHR - instrumental variable regression $\sqrt{nobs}$
Source: own analysis

The null hypothesis of both the FAT and FAIVEHR tests ($H_0 : \beta_0 = 0$) were strongly rejected across all sub-samples. In comparison, the results of the PET test lack consistence. An authentic robust link between minimum wage and price level changes significantly different from zero was detected only when estimated on the sub-sample of unpublished studies. However, with regard to the guidelines for the use of partial correlation coefficients in meta-analysis published by Doucouliagos (2011), the small size of the coefficient reduces its practical significance to zero. The guidelines combine the data used in 41 meta-analyses focused on various fields to identify the distribution of the size of empirical effect reported in economic literature. A general classification is proposed within the study, according to which partial correlations sized in absolute value between 0.07 and 0.17 fall within the *small* category. Partial correlations larger than 0.33 in absolute values are considered *large* and, naturally, anything between 0.17 and 0.33 is considered *medium*. The rest of the results in table 5.9 irrespective of the data subset and estimation method suggest that the true underlying relationship is in fact too small to be considered significant both practically and statistically.

A summary of the results of the mixed effects multi-level regressions based on specification 4.1.5 is reported in table 5.2. Since multiple studies in our dataset share the same author or authors, we have estimated the model twice for each data subset allowing for study and author-level random effects separately. The rejection of the null hypothesis of the likelihood ratio test ($H_0 : \rho = 0$) serves as an indication of a substantial between-study heterogeneity, the occurrence of which implies that the OLS regression is in fact misspecified and offers less reliable results when compared to mixed effects multi-level model. The differences in the estimated results are however not large and do not include significant directional changes in the estimated relations. The consistency between the OLS and mixed-effect multilevel models provides us with assurance that the exogeneity assumptions backing the later are not seriously violated.

Likewise the OLS and FE regression, neither the mixed effects multi-level model signals the existence of a linkage between the movements in minimum wage and price levels. Out of all used specifications, the only estimate of the relationship significantly different from zero, though on the 1% level, was obtained from the sub-group of published studies when study-level clustering was applied. However, the practical strength of the relationship estimated to be 0.0128 renders the result negligible as in the OLS and FAIVEHR specifications.

Despite the drop in statistical significance, the MRA analysis reported in ta-

Table 5.2: Mixed effects multi-level model

Variable	**Study level ME**			**Author level ME**		
	All	Published	Unpublished	All	Published	Unpublished
	(1)	(2)	(3)	(4)	(5)	(6)
β_0 (selectivity)	1.5459***	0.8985†	2.7617***	1.206**	1.1804*	2.3185*
	(3.52)	(1.75)	(3.45)	(2.70)	(2.32)	(2.47)
$1/Se_i$ (effect)	0.0021	0.0128**	-0.0162†	0.0007	0.0029	-0.0172†
	(0.48)	(2.91)	(-1.87)	(0.27)	(1.26)	(-1.74)
♯ groups	23	13	10	12	9	7
N	469	270	199	469	270	199
Correlation (ρ)	0.2786	0.4593	0.1556	0.2115	0.3717	0.2244
χ^2	70.03***	77.25***	6.44**	69.76***	50.31***	9.63***

Significance levels: †:10% *:5% **: 1% ***: 0.1%
Dependant variable: t-statistic of the partial correlation of the effect of minimum wage on inflation
t-statistics of the estimates in parenthesis
Correlation ρ: within group correlation
χ^2 : LRT test
Source: own analysis

ble 5.2 identifies an upward publication selectivity significant at least at 10% level across all estimated regressions. Rather surprisingly, its distortive effects are the strongest among the group of unpublished studies, where the intercept term encompassing the publication selectivity exceeds two in absolute values irrespective of the selected clustering. Such an extent of publication bias is considered "severe"' and hints at frequent usage of specification searching. This statement is supported by a simple thought experiment - if only positively significant partial correlations appeared in the literature despite the true partial correlation being zero, the detected publication bias would be roughly equal to two, the benchmark t-statistic of statistical significance. This way, publication selectivity in the literature can skew the population of reported estimates to the extent that on average we start observing a significant effect even though in reality there is none at all (Doucouliagos and Stanley, 2008; Havranek, Irsova, and Janda, 2012).

5.3 Correcting for publication selectivity

Figure 5.7 plots the corrected and unadjusted partial correlations of our complete dataset, marked with cross and circle respectively. To correct the data we shrunk all reported partial correlations towards the intercept 0.0007 from regression (4) in table 5.2 by the 1.206 product of their standard errors. Both

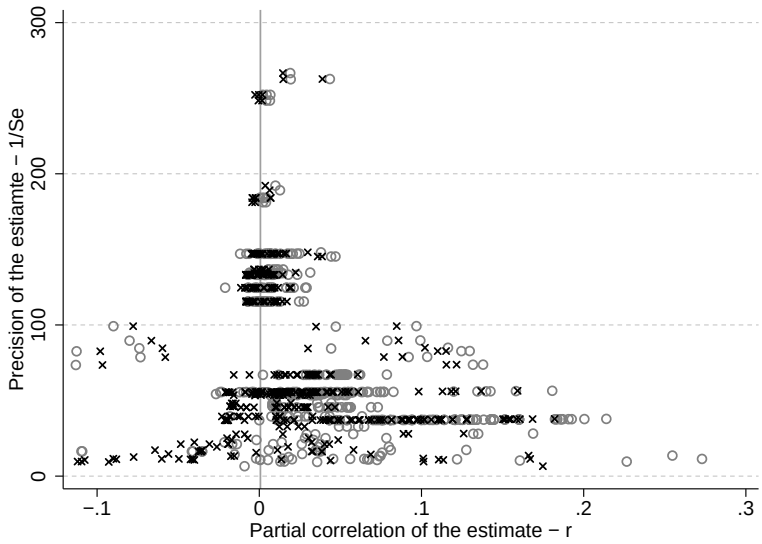

Figure 5.7: Funnel plot: type I publication bias corrected

values were taken from the preferred mixed-effects multilevel regression estimated on the full dataset. The positive skewness of the middle part of the plot becomes much less evident after the adjustment is made contributing to a more funnel-like shape. We also note a higher occurrence of slightly negative partial correlations, which given the size of the underlying effect is in line with our expectations. Nevertheless, even the corrected plot appears to be heavier towards the left especially at the base of the plot where the less-precise partial correlations are displayed.

In the ideal case, shrinking the partial correlations in absolute values by the appropriate multiple of its standard error should rid the data of any systematic bias including *type II publication selectivity*. The scatter diagram in figure 5.8 plotting the Galbraith plot of both adjusted and original data however suggests just how far from the ideal our simple MRA is. Even though we see an improvement in the proportion of data points lying outside the confidence band around zero, the acceptable 5% benchmark remains out of reach.

One possible explanation as to why the publication selectivity displays such level of persistence is that the explanatory power of our MRA analysis does not suffice for capturing even the most essential patterns in our data. The way the Galbraith plot in figure 5.8 was constructed, we assume that there is only one true relationship between shifts in minimum wage rate and prices equal to 0.0007, not taking into account that the possible existence of two or more different true effects depending on structural characteristics of the primary data

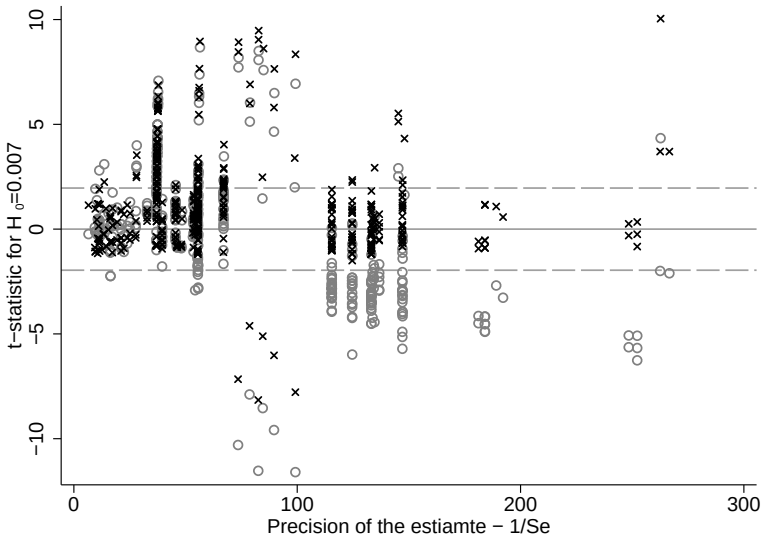

Figure 5.8: Galbraith plot: type II publication bias corrected

might be just as real.

In light of that, we dedicate the following section to designing a more complex MRA analysis focused on explaining the heterogeneity, structural or induced by the statistical methods applied to obtain the empirical estimates.

5.4 Meta-regression analysis of heterogeneity among studies

As the first step before we actually design our MRA model, correlations between available meta-independent variables should be considered carefully to rule out or at least limit collinearity and spurious correlation from the regression. In that regard, we have inspected the correlation matrix[4] of all meta-independent variables paying special attention to variables that displayed mutual correlation of 0.8 and higher.

Some of the dummy variables, or more precisely their combinations were perfectly collinear by design (e.g. time-series versus panel data dummy). In other cases, the implications of excluding any of the highly correlated variables, were carefully considered. As a result, several controls were dropped from the

[4]The matrix is not reported in the thesis but will be provided upon request

regression[5] that way creating a control group of estimates characterized mainly by being based on panel data, using WLS or MLE estimation techniques, and by measuring the price level movements on a broad consumption bundle as opposed to studies focused on item-level prices.

The way we design our model, we augment the basic FAT-PET regression with all variables that passed the correlation analysis. We go through the process twice, once for the mixed-effects and once for the basic OLS specification with error clustering. For each estimated model, we test the joint significance of the least significant variables in the regression. All jointly insignificant variables are excluded in one step. The only variable never included in the F-test irrespective of its p-value is the underlying effect control $\frac{1}{Se_i}$. Bearing in mind that the purpose of our augmented MRA is to shed light on the study-level specifics that affect the reported estimates as such, all meta-independent variables in the regression were divided by the standard error of the partial correlations in conformity with the models in equations 4.2.1 and 4.2.2.

The final results of the augmented model are reported in table 5.3. We use both OLS and mixed-efects multilevel approaches as a robustness check. Nevertheless, given the correlation structure of our dataset, fulfilling the exogeneity assumptions of the ME model appears doubtful at best. For that reason, if inconsistencies between the two models are revealed, the OLS specification is treated as superior. Moreover, under the current conditions with several researchers authoring more than one study in our sample, author-level clustering of studies is preferred and as such reported. However, additional regression results including the original model and models with study-level clustered errors are attached as part of Appendix C.

Perhaps most importantly, the augmented model has robustly confirmed the presence of a strong publication selectivity in our data. Both models also detect a slight but significant trace of a genuine relationship between increases in minimum wage rate and prices. There is, however, no improvement in terms of the practical size of the link which remains negligible when the classification by Doucouliagos (2011) is considered. Apart from that, ten meta-independent variables consistently produce significant coefficients irrespective of the actual specification suggesting that the design of empirical studies is systematically reflected in their results.

[5] *repeccit, panel, price_index_data, avdate, loc_usa, loc_OECD, loc_sa, loc_au, genconsbungle, general_eq, short, long, lag, lead, WLS, MLE*

Table 5.3: MRA full results (author level clustering)

Variable	ME		OLS	
β_0 (selectivity)	0.9103*	(2.38)	0.9465*	(2.89)
$1/Se_i$ (effect)	0.043†	(1.85)	0.0387†	(1.88)
Quality characteristics				
Published	-		0.0095	(1.47)
Google scholar citations	-0.0000†	(-1.73)	-0.0000***	(-3.98)
Repec impact factor	0.0014	(1.19)	0.0017	(1.66)
Data characteristics				
Time series data	-		0.0371†	(2.10)
Item-level data	0.022*	(2.27)	0.02614**	(3.22)
Estimation characteristics				
OLS	-0.0873***	(-4.63)	-0.2901***	(-8.16)
Robust regression	-0.0977***	(-4.00)	-0.3053***	(-7.76)
Median regression	-0.0706***	(-2.89)	-0.2782***	(-7.08)
IV estimation	-0.111***	(-4.20)	-0.3148***	(-8.49)
Nature of dependant variable				
Food	0.0107*	(2.09)	0.0108***	(17.18)
Restaurants	-0.0063	(-1.47)	-0.0062***	(-6.14)
Poor	-0.0065	(-1.36)	-0.0038***	(-5.40)
Rich	-0.006	(-1.19)	-	
Model				
Imperfect competition	-		-0.0148**	(-3.16)
VAR	-0.1252	(-1.19)	-0.2532***	(-10.32)
Conditioning variables				
Combination	-0.0038	(-1.06)	-0.0045	(-1.64)
# of lead months	0.0019†	(1.60)	0.002	(1.51)
City area control	0.0053	(1.35)	0.0061	(1.21)
Unemployment change control	-		-0.0068	(-1.14)
Seasonality controls	-		-0.0004	(-0.11)
Mean reversion	0.0169*	(2.32)	0.0128†	(2.03)
Input price shock	0.0037	(0.91)	0.0061**	(3.06)
Real factors				
Dataset span	0.0012†	(1.87)	0.0016***	(4.78)
Average year	-0.0016*	(-2.13)	-0.0011	(-0.99)
EU	-		0.1549***	(4.38)
North America	-0.0012†	(-0.07)	0.1898***	(5.70)
N	469		469	
Authors	12		12	

Significance levels: †:10% *:5% **: 1% ***: 0.1%
Dependant variable: *t*-statistic of the partial correlation of the effect of minimum wage on inflation
t-statistics (z-statistics for ME) of the estimates in parenthesis (author level clustered errors)
ME - mixed effects multi-level model
OLS - ordinary lease squares estimation
Source: own analysis

In line with our initial expectations, both OLS and ME models detect a small yet significant increase in reported partial correlations for studies employing *item-level* data. In other words, overall price level measured by an aggregated price index is not as easily susceptible to reacting to minimum wage rate fluctuations as prices of individual consumer goods are. The self-evident explanation is that price indices consist of representative bundles of consumer goods not all of which are produced in industries where cheap labor forms a significant part of the work force. Such firms are then not forced to markedly reflect the potential increase in labor costs into their final prices. Knowing that, researchers incline to using item-level prices in cheap-labor intensive industries in hope of obtaining significant and publication-worthy estimates.

In our model, the variable *food* controls for the possibly non-random choice of industries. What the inclusion of the dummy in effect allows us to do is to separate the two influences. The obtained results show that using food-industry prices further slightly but significantly reinforces the observed relationship. The food-related control for *restaurant* prices produces a negative coefficient. This may seem startling at first, however, it should be noted that the majority of the food-related estimates comes either from *restaurant* or *fast-food* businesses. What the negative coefficient in fact does, is point to a lower responsiveness of restaurant prices compared to fast-food industry, where the proportion of minimum-wage-earning employees is likely to be even higher.

Regarding publication characteristics, neither the number of *Google scholar* citations, nor the *impact factor* of publication outlet appear to be substantially linked to the size of partial correlations derived from the study's empirical results. On the other hand, there seems to be a systematic difference in the size of the effect reported in studies that were actually accepted for *publication* and those left by their authors in the "file drawer"'. The finding is in line with the inference drawn from the graphical analysis in section 5.1, where partial correlations from published studies markedly overweighed the funnel plot in the positive direction.

Finally, besides the statistical properties of the estimation models, geographical affiliation appears to be an important determinant of the size of the reported effect. Both EU and North American datasets show a markedly higher responsiveness of goods markets' to changes in labor market policies compared to the control group. This suggests a degree of proportionality between the sensitivity rate and the maturity of the markets and economy.

Conclusion

The key objective of the diploma thesis was to synthesize and evaluate existing empirical research focused on the price effect of minimum wage changes. To the best of our knowledge, all preceding literature surveys with similar orientation were of a narrative nature thus making this study the very first quantitative literature review proposing a systematic approach to the matter.

An absolutely essential part of the work consisted in gathering a sufficiently large and comprehensive dataset with estimates of the analyzed effect. I carefully read through over 40 empirical studies out of which 23 were selected for the final meta-dataset. In the context of meta-analysis, data collection typically counts among the most time consuming tasks. Our study was no exception to the rule, particularly since in this case updating or expanding an already existing meta-analytical dataset was not an option. Instead, we created an entirely new data matrix counting 469 observations of the effect and additional 46 qualitative variables describing the individual estimates. Depending on the methodology used within the studies, we found inconsistency in the interpretation of reported estimates. To resolve the issue and preserve their comparability, all reported estimates were transformed into partial correlations.

The ambitions of the meta-analysis within the study were two-fold. At first, the focus was on detecting and treating for the publication selectivity in the literature sample. Depending on how skewed the incentives of researchers are, the issue has a potential to significantly distort our understanding of the story told by empirical research, especially when the overall effect reported in the literature is measured by arithmetic or weighted means.

Using various meta-analytical methods, both graphical and statistical, we find no conclusive evidence of the existence of a link between minimum wage hikes and the price level changes. The estimated strength of the relationship corresponds to a partial correlation coefficient between 0.0029 and 0.0147 for pub-

lished and -0.0245 and -0.0162 for unpublished studies, where only the later coefficients are consistently significant across different specifications of the model. Even though the small size of the coefficients reduces their practical significance to zero, the negative estimates of the true partial correlation among unpublished studies give an idea of how distortive a measure arithmetic and weighted means (amounting to 0.042 and 0.027 respectively) are.

Besides the basic OLS specification, the size of the underlying population value and the publication selectivity tests were estimated on a heteroscedasticity robust model. Moreover, we interpreted the dataset as longitudinal data and estimated the same model with study-level fixed effects. Finally, mixed-effects multilevel model able to account for heteroscedasticity as well as between-study heterogeneity was employed. All models have robustly confirmed the presence of publication selectivity in the literature. While the unpublished studies in fact appear to be the most distorted according to our tests in terms of significance as well as the size of the bias, it should be taken into account that the skewness of the population of estimates reported in unpublished studies can be in a large part accounted to a single study by Lemos (2003).

The second aim of our thesis was to construct an explanatory MRA model that helped us uncover some of the sources of the heterogeneity among the estimated results. In that respect, our augmented mixed-effects and OLS models found that reported price level changes were statistically most significantly linked to the selected estimation methods, empirical models, and geographical affiliation of the underlying dataset.

Overall, our findings are not too different from those outlined by Lemos (2006b). No robust link of a relevant size between minimum wage changes and price increases was discovered. The implication to macroeconomic policy-making is that adjustments of the minimum wage level may be used by the governments as a valid labor market instrument capable of increasing the income of lower social classes without the concern of producing inflationary pressures that could easily erase its socially beneficial purpose.

The benefits brought by using meta-analysis as a quantitative alternative to traditional literature surveys were clearly outlined several times in the thesis. To assess the method objectively, however, it should not be withhold from the reader that same as for any other method, many factors enter the equation for reliability of obtained outcomes. In particular, the inference drawn from our models critically hinges on the parameters of the underlying meta-dataset and

the, I dare say, meticulousness with which it was assembled. During the data collection process, I had to deal with a considerable number of data issues reinforced by the low number of studies in the particular research domain. This essentially prevented me from using the widely advised solution to generally any data issue - dropping inconvenient studies from the sample. The problems I have encountered ranged from the need to impute unreported p-values to transforming the whole dataset into partial correlations, all of which introduce a level of distortion into the meta-dataset.

Last but not least, I would like to stress the essentiality of individual empirical research itself as without it, and the issues I have encountered are a sound testament to this, the power of literature surveys remains limited.

Bibliography

Aaronson, D. (2001): "Price pass-through and the minimum wage," *Review of Economics and statistics*, 83(1), 158–169.

Aaronson, D., E. French, and J. MacDonald (2008): "The minimum wage, restaurant prices, and labor market structure," *Journal of human resources*, 43(3), 688–720.

Andreica, M. E., L. Aparaschivei, A. Cristescu, and N. Cataniciu (2010): "Models of the minimum wage impact upon employment, wages and prices: the Romanian case," in *Proceedings of the 11th WSEAS International Conference MCBE*, vol. 10, pp. 104–109.

Basker, E., and M. Khan (2013): "Does the Minimum Wage Bite into Fast-Food Prices?," *Available at SSRN 2326659*.

Borjas, G. J. (2005): *Labor Economics*. McGraw-Hill New York.

Brown, C. (1999): "Minimum wages, employment, and the distribution of income," *Handbook of labor economics*, 3, 2130.

Card, D., and A. B. Krueger (1995a): *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton University Press.

——— (1995b): "Time-series minimum-wage studies: a meta-analysis," *American Economic Review*, 85(2), 238–243.

Cipra, T. (2008): *Finanční ekonometrie*. Ekopress.

Cuong, N. V. (2011): "Do Minimum Wage Increases Cause Inflation?," *ASEAN Economic Bulletin*, 28(3).

Dickey, D. A., and W. A. Fuller (1979): "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American statistical association*, 74(366a), 427–431.

Doucouliagos, C., and P. Laroche (2003): "What Do Unions Do to Productivity? A Meta-Analysis," *Industrial Relations: A Journal of Economy and Society*, 42(4), 650–691.

Doucouliagos, H. (2011): "How Large is Large? Preliminary and relative guidelines for interpreting partial correlations in economics," Deakin Univerity Australia.

Doucouliagos, H., and T. D. Stanley (2008): "Theory Competition and Selectivity: Are all economic facts greatly exaggerated?," Discussion paper, Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.

Doucouliagos, H., and T. D. Stanley (2009): "Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis," *British Journal of Industrial Relation*, 47(2), 406–428.

Draca, M., S. Machin, and J. Van Reenen (2005): *The impact of the national minimum wage on profits and prices: report for the Low Pay Commission*. Centre for Economic Performance, London School of Economics and Political Science: University College London.

Dube, A., S. Naidu, and M. Reich (2006): "Economic Effects of a Citywide Minimum Wage, The," *Indus. & Lab. Rel. Rev.*, 60, 522.

Engle, R. F., and C. W. Granger (1987): "Co-integration and error correction: representation, estimation, and testing," *Econometrica: journal of the Econometric Society*, pp. 251–276.

Falconer, R. T. (1979): "The minimum wage: a perspective," .

Fels, A., and T. V. Hoa (1981): "Causal Relationships in Australian Wage Inflation and Minimum Award Rates," *Economic Record*, 57(1), 23–34.

Fougère, D., E. Gautier, and H. Le Bihan (2010): "Restaurant prices and the minimum wage," *Journal of Money, Credit and Banking*, 42(7), 1199–1234.

Frye, J., and R. J. Gordon (1981): "Government Intervention in the Inflation Process: The Econometrics of" Self-Inflicted Wounds"," .

Glass, G. V. (1976): "Primary, secondary, and meta-analysis of research," *Educational researcher*, pp. 3–8.

Granger, C. W. (1969): "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438.

Greenberg, D. H., C. Michalopoulos, and P. K. Robins (2003): "Meta-Analysis of Government-Sponsored Training Programs, A," *Indus. & Lab. Rel. Rev.*, 57, 31.

Hadi, A. S. (1992): "Identifying multiple outliers in multivariate data," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 761–771.

Havranek, T., and Z. Irsova (2011): "Estimating vertical spillovers from FDI: Why results vary and what the true effect is," *Journal of International Economics*, 85(2), 234–244.

Havranek, T., Z. Irsova, and K. Janda (2012): "Demand for gasoline is more price-inelastic than commonly thought," *Energy Economics*, 34(1), 201–207.

Johansen, S. (1991): "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models," *Econometrica: Journal of the Econometric Society*, pp. 1551–1580.

Khan, M. T. (2013): "Essays on retail pricing," Ph.D. thesis, University of Missouri–Columbia.

Lee, C., and B. O'Roark (1999): "The impact of minimum wage increases on food and kindred products prices: An analysis of price pass-through," Discussion paper, United States Department of Agriculture, Economic Research Service.

Lee, C., G. Schluter, and B. O'Roark (2000): "Minimum wage and food prices: an analysis of price pass-through effects," *The International Food and Agribusiness Management Review*, 3(1), 111–128.

Lee, C., G. E. Schluter, and B. O'Roark (2000): *How much would increasing the minimum wage affect food prices?* Economic Research Service, US Department of Agriculture.

Lemos (2003): "The Effect of Minimum Wages on Prices in Brazil," Discussion Paper 04-01, University College London.

Lemos, S. (2004a): "Do Minimum Wage Price Effects Hurt the Poor More?," *Revista Economica*, 50(1), 67–83.

——— (2004b): "The Effects of the Minimum Wage on Wages, Employment and Prices," .

Lemos, S. (2005): "Minimum Wage Effects on Wages, Employment and Prices: Implications for Poverty Alleviation in Brazil," .

Lemos, S. (2006a): "Anticipated Effects of the Minimum Wage on Prices," *Applied Economics*, 38(3), 325–337.

Lemos, S. (2006b): "A Survey of the Effects of the Minimum Wage on Prices.," *Journal of Economic Surveys*, 22(1), 187–212.

L'horty, Y., and C. Rault (2004): "Inflation, minimum wage and other wages: an econometric study on French macroeconomic data," *Applied Economics*, 36(4), 277–290.

MacDonald, J. M., and D. Aaronson (2000): "How do retail prices react to minimum wage increases?(No. WP-00-20)," Federal Reserve Bank of Chicago.

——— (2006): "How firms construct price changes: evidence from restaurant responses to increased minimum wages," *American Journal of Agricultural Economics*, 88(2), 292–307.

Machin, S., A. Manning, and L. Rahman (2003): "Where the minimum wage bites hard: Introduction of minimum wages to a low wage sector," *Journal of the European Economic Association*, 1(1), 154–180.

MaCurdy, T., and F. McIntyre (2001): "Winners and losers of federal and state minimum wages," *Employment Policies Institute Foundation Paper*, 6(01).

MaCurdy, T., and M. O'Brien-Strain (2000): "'Increasing the Minimum Wage: California's Winners and Losers.'," Discussion paper, Working Paper. San Francisco: Public Policy Institute of California.

MaCurdy, T. E., and M. O'Brien-Strain (1997): *Who Will be Affected by Welfare Reform in California?* Public Policy Instit. of CA.

Nelson, J. P., and P. E. Kennedy (2009): "The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment," *Environmental and Resource Economics*, 42(3), 345–377.

Powell, L. A. (2007): "Approximating variance of demographic parameters using the delta method: a reference for avian biologists," *The Condor*, 109(4), 949–954.

Rosenthal, R. (1979): "The file drawer problem and tolerance for null results," *Psychological bulletin*, 86(3), 638.

Shin, Y., and P. Schmidt (1992): "The KPSS stationarity test as a unit root test," *Economics Letters*, 38(4), 387–392.

Stanley, T., and H. Doucouliagos (2010): "Picture this: a simple graph that reveals much ado about research," *Journal of Economic Surveys*, 24(1), 170–191.

Stanley, T. D. (2005): "Beyond publication bias," *Journal of Economic Surveys*, 19(3), 309–345.

Stanley, T. D. (2008): "Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection," *Oxford Bulletin of Economics and statistics*, 70(1), 103–127.

Stanley, T. D., C. Doucouliagos, and S. B. Jarrell (2008): "Meta-regression analysis as the socio-economics of economics research," *The Journal of Socio-Economics*, 37(1), 276–292.

Stanley, T. D., and H. Doucouliagos (2012): *Meta-regression analysis in economics and business*. Routledge Abingdon, Oxon ; New York.

Stanley, T. D., and S. B. Jarrell (1989): "Meta-Regression analysis: A quantitative method of literature surveys," *Journal of Economic Surveys*, 3(2), 161–170.

Stiegler, G. J. (1946): "The economics of minimum wage legislation," *he American Economic Review*, 36(3), 358–365.

Wadsworth, J. (2010): "Did the National Minimum Wage Affect UK Prices?*," *Fiscal Studies*, 31(1), 81–120.

Werner, T., F. L. Sell, and D. C. Reinisch (2013): "Price effects of minimum wages: Evidence from the construction sector in East and West Germany," Discussion paper, Volkswirtschaftliche Diskussionsbeiträge, Universität der Bundeswehr München, Fachgruppe für Volkswirtschaftslehre.

Wilson, D. M., W. W. Beach, R. A. Rector, and R. S. Hederman (1999): "How Congress's Tax Bill Would Affect Families, the Economy, and the Federal Budget," Center for data analysis report 99-06, The Heritage Center for Data Analysis.

Wolff, E. N., and M. I. Nadiri (1980): *A simulation model of the effects of an increase in the minimum wage on employment, output and the price level*. Department of Economics, New York University.

Wooldridge, J. (2012): *Introductory econometrics: A modern approach*. Cengage Learning.

Appendix A: List of studies

Table 5.4: List of studies

ID study	Author (Year)	Country	No. of estaimtes
1	Andreaica et al. (2010)	Romania	1
2	Aaronson (2001)	USA	164
3	Aaronson et al. (2008)	USA/Canada	18
4	Basker and Khan (2013)	USA	6
5	Card and Krueger (1993)	USA	3
6	Card and Krueger (1995)	USA	4
7	Draca et al. (2005)	UK	1
8	Dube et al. (2006)	USA	10
9	Fels and Hoa (1981)	Australia	2
10	Fougère et al. (2010)	France	12
11	Frye and Gordon (1981)	USA	1
12	Katz and Krueger (1992)	USA	4
13	Khan (2013)	USA	118
14	Lemos (2003)	Brazil	51
15	Lemos (2004)	Brazil	18
16	Lemos (2004)	Brazil	4
17	Lemos (2005)	Brazil	1
18	Lemos (2006)	Brazil	11
19	L'horty and Raul (2004)	France	2
20	Machin et al. (2003)	UK	4
21	MacDonald and Aaronson (2000)	USA	11
22	MacDonald and Aaronson (2006)	USA	11
23	Powers (2009)	USA	12

Appendix B: Detail description of data-set

Table 5.5: Detail description of data-set

Variable	(Definition)	Dummy	(Mean)	(Std. error)
Effect size and its precision				
r	partial correlation coefficient	no	0.042	(0.055)
$\frac{1}{Se}$	precision measure, used to test the genuine effect (PET)	no	73.03	(50.247)
$\frac{r}{Se}$	dependent variable of the FAT-PET	no	1.995	(2.529)
Quality characteristics				
repecimp	size of impact factor of publication outlet, RePEc	no	9.704	(7.923)
repeccit	number of citations on RePEc	no	8.615	(8.617)
schcit	number of citation on Google scholar	no	69.908	(212.887)
published	=1, if study published	yes	0.576	(0.495)
Data characteristics				
panel	=1, if study uses panel data	yes	0.959	(0.197)
timeseries	=1, if study uses time-series data	yes	0.041	(0,197)
own_data	=1, if author collected data himself	yes	0.079	(0.27)
item_price_data	=1, if study uses item prices	yes	0.493	(0.5)
price_index_data	=1, if study uses a price index	yes	0.507	(0.5)
Estimation characteristics				
OLS	=1, if study uses OLS regression	yes	0.665	(0.472)
robust	=1, if study uses robust regression	yes	0.047	(0.212)
median	=1, if study uses median regression	yes	0.038	(0.192)
IV	=1, if study uses instrumental variable regression	yes	0.058	(0.233)
WLS	=1, if study uses WLS regression	yes	0.181	(0.386)
MLE	=1, if study uses MLE regression	yes	0.009	(0.092)

Table continues on the next page

Table 5.6: Detail description of data-set

Variable	(Definition)	Dummy	(Mean)	(Std. error)
Specifics of the dependent variable				
food	=1, if effect related to food prices	yes	0.77	(0.421)
restaurants	=1, if effect related to restaurant prices	yes	0.222	(0.416)
fastfood	=1, if effect related to fast-food prices	yes	0.484	(0.5)
genconsbungle	=1, if effect related to the price of a general consumption bundle	yes	0.194	(0.396)
poor	=1, if effect related to inflation perception of the poor	yes	0.075	(0.263)
rich	=1, if effect related to inflation perception of the rich	yes	0.068	(0.252)
Underlying model				
imperfect	=1, if regression based on imperfect competition model	yes	0.087	(0.283)
general_eq	=1, if regression based on general equilibrium model	yes	0.094	(0.292)
DID	=1, if difference-in-difference approach used	yes	0.041	(0.197)
VAR	=1, if study uses vector autoregressive model (VAR)	yes	0.011	(0.103)
Real factors: time and geographical affiliation				
date_length	data set span in months	no	12.251	(6.448)
avdate	average year of the data, mean year subtracted	no	0	(6.821)
loc_usa	=1, if US data used	yes	0.699	(0.459)
loc_oecd	=1, if OECD member state data used	yes	0.817	(0.387)
loc_eu	=1, if EU member data used	yes	0.043	(0.202)
loc_sa	=1, if South American data used	yes	0.181	(0.386)
loc_na	=1, if North American data used	yes	0.772	(0.42)
loc_au	=1, if Australian data used	yes	0.004	(0.065)
Controls in primary regressions				
short	=1, if estimate relates to contemporaneous effect	yes	0.356	(0.479)
long	=1, if estimate relates to longer-term effect	yes	0.644	(0.479)
lag	=1, if estimate is or includes lagged effect	yes	0.488	(0.505)
lead	=1, if estimate is or includes leading effect	yes	0.418	(0.494)
#lag	no. of lagged months covered by the estimate	no	1.5946	(2.226)
#lead	no. of leading months covered by the estimate	no	1.222	(1.542)
combination	=1, if delta method was used to produce the estimate	yes	0.235	(0.424)
cityareactrl	=1, if city area control present in primary regression	yes	0.727	(0.446)
unempchangectrl	=1, if unemployment change control present in primary regression	yes	0.096	(0.295)
seasonalityctrl	=1, if seasonality control present in primary regression	yes	0.797	(0.402)
meanrevertingctrl	=1, if mean reversion control present in primary regression	yes	0.156	(0.363)
inputpriceshockctrl	=1, if input-price-shock control present in primary regression	yes	0.292	(0.445)

Source: own analysis

Appendix C: Corrected graphical analysis

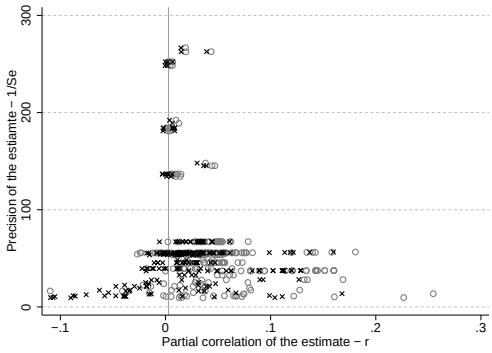

Figure 5.9: Corrected funnel plot: published studies

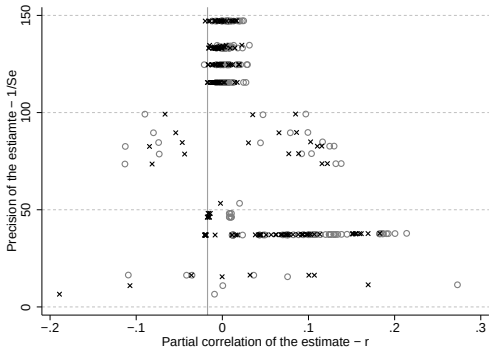

Figure 5.10: Corrected funnel plot: unpublished studies

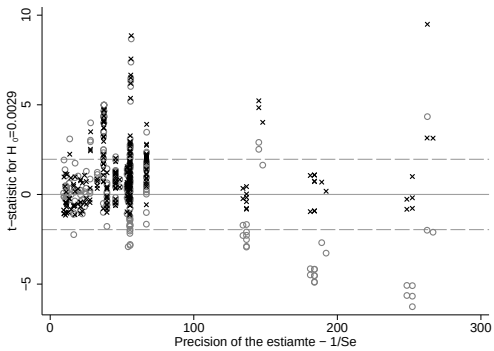

Figure 5.11: Corrected Galbraith plot: published studies

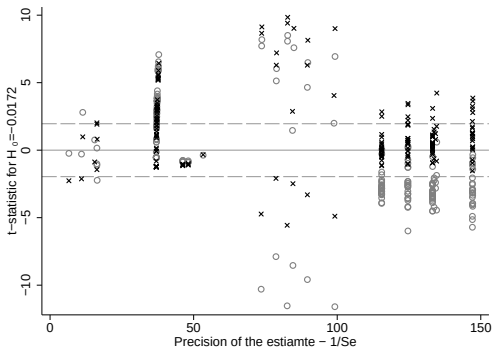

Figure 5.12: Corrected Galbraith plot: unpublished studies

Appendix D: MRA full results

Table 5.7: MRA full results (study level clustering)

Variable	1-ME		2-ME		1-OLS		2-OLS	
β_0 (selectivity)	1.1189	(1.56)	0.7602	(1.44)	0.8341	(1.34)	0.8905	(1.60)
$1/Se_i$ (effect)	0.0186	(0.44)	0.0469†	(1.66)	0.043	(1.24)	0.0644***	(3.85)
Quality characteristics								
Published	0.0147	(0.75)	-		0.0074	(0.83)	-	
Google scholar citations	-0.0000	(-1.12)	-0.0000	(-1.56)	-0.0000†	(-1.92)	-0.0000*	(-2.38)
Repec impact factor	0.0025	(1.09)	0.0017	(0.98)	0.0017	(0.99)	-	
Data characteristics								
Time series data	0.0386	(0.42)	-		0.0416	(1.11)	0.0203	(0.66)
Collected own data	-0.0233	(-0.31)	-		0.0064	(0.15)	0.0119	(0.31)
Item-level data	0.0332†	(1.82)	0.0257*	(1.97)	0.026†	(1.88)	0.0134***	(4.20)
Estimation characteristics								
OLS	-0.2933	(-1.13)	-0.0631*	(-2.03)	-0.2935***	(-6.00)	-0.2859***	(-6.59)
Robust regression	-0.3145	(-1.20)	-0.0798*	(-2.11)	-0.3087***	(-5.79)	-0.2879***	(-6.79)
Median regression	-0.0528	(-1.09)	-0.0528	(-1.39)	-0.2815***	(-5.27)	-0.2607***	(-6.15)
IV estimation	-0.0737	(-1.19)	-0.0737*	(-1.99)	-0.3157***	(-6.58)	-0.3103***	(-7.30)
Nature of dependant variable								
Food	0.0103†	(1.79)	-		0.0098***	(3.96)	0.0096***	(18.82)
Restaurants	-0.0068	(-1.38)	-0.0041	(-0.87)	-0.0057**	(-3.02)	-0.0054*	(-2.56)
Fastfood	0.0003	(0.08)	0.0019	(0.59)	0.0011	(0.39)	-	
Poor	-0.0075	(-1.50)	-0.008†	(-1.65)	-0.0059	(-0.93)	-	
Rich	-0.0065	(-1.28)	-0.0071	(-1.43)	-0.0051	(-0.64)	-	
Model								
Imperfect competition	-0.0133	(-0.69)	-		-.0124	(-1.31)	-0.0239	(-1.61)
Difference in difference	0.0289	(0.44)	-		0.0103	(0.32)	-	
VAR	-0.2828	(-1.31)	-0.1152	(-0.93)	-0.254***	(-6.59)	-0.2822***	(-10.23)
Conditioning variables								
Combination	-0.0044	(-1.19)	-0.0028	(-0.76)	-0.0039	(-0.83)	-	
# of lagged months	0.0000	(-0.05)	-		-0.0000	(-0.19)	0.0000	(0.58)
# of lead months	0.0019	(1.59)	0.002†	(1.66)	0.0019	(1.18)	0.0011	(1.12)
City area control	0.0000	(1.63)	0.0044	(1.10)	0.0062	(1.38)	0.0056*	(2.05)
Unemployment change control	-.0073	(-0.76)	-0.0059	(-0.68)	-0.006†	(-1.81)	-0.0069	(-1.37)
Seasonality controls	-.0011	(-0.17)	-		-.0014	(-0.32)	-	
Mean reversion	0.0068	(0.52)	-		0.0143	(1.18)	0.0136*	(2.06)
Input price shock	0.0051	(0.95)	0.0066	(1.54)	.0048	(1.45)	0.0078	(1.65)

Table continues on the next page

Table 5.8: MRA full results (study level clustering)

Variable	**1-ME**		**2-ME**		**1-OLS**		**2-OLS**	
Real factors								
Dataset span	0.0017†	(1.71)	0.0008	(1.13)	0.0016†	(1.87)	0.0011***	(3.03)
Average year	-0.0019	(-0.99)	-0.0031*	(-2.38)	-0.0013	(-1.13)	-0.0027***	(-5.87)
EU	0.1764	(0.72)	-		0.1545***	(4.35)	0.1776***	(6.20)
North America	0.2024	(0.78)	-0.0089	(-0.37)	0.1925***	(4.39)	0.1962***	(5.05)
N	469		469		469		469	
Studies	23		23		23		23	

Significance levels: †:10% *:5% **: 1% ***: 0.1%

Dependant variable: *t*-statistic of the partial correlation of the effect of minimum wage on inflation

t-statistics (z-statistics for ME) of the estimates in parenthesis (study level clustered errors)

ME - mixed effects multi-level model

OLS - ordinary lease squares estimation

Source: own analysis

Table 5.9: MRA full results (author level clustering)

Variable	1-ME		2-ME		1-OLS		2-OLS	
β_0 (selectivity)	0.834	(1.50)	0.9103*	(2.38)	0.8341	(1.23)	0.9465*	(2.89)
$1/Se_i$ (effect)	0.043	(1.25)	0.043†	(1.85)	0.0429	(1.52)	0.0387†	(1.88)
Quality characteristics								
Published	0.0075	(0.50)	-		0.0075	(1.09)	0.0095	(1.47)
Google scholar citations	-0.0000	(-1.42)	-0.0000†	(-1.73)	-0.0000***	(-4.74)	-0.0000***	(-3.98)
Repec impact factor	0.0018	(1.00)	0.0014	(1.19)	0.0018†	(1.84)	0.0017	(1.66)
Data characteristics								
Time series data	0.0416	(0.60)	-		0.0416†	(1.88)	0.0371†	(2.10)
Collected own data	0.0063	(0.12)	-		0.0063	(0.18)	-	
Item-level data	0.026†	(1.77)	0.022*	(2.27)	0.026**	(3.70)	0.02614**	(3.22)
Estimation characteristics								
OLS	-0.2935	(-1.34)	-0.0873***	(-4.63)	-0.2934***	(-6.71)	-0.2901***	(-8.16)
Robust regression	-0.3087	(-1.40)	-0.0977***	(-4.00)	-0.3087***	(-7.10)	-0.3053***	(-7.76)
Median regression	-0.2815	(-1.28)	-0.0706***	(-2.89)	-0.282***	(-6.47)	-0.2782***	(-7.08)
IV estimation	-0.3157	(-1.43)	-0.111***	(-4.20)	-0.316***	(-7.26)	-0.3148***	(-8.49)
Nature of dependant variable								
Food	0.01†	(1.68)	0.0107*	(2.09)	0.01***	(3.92)	0.0108***	(17.18)
Restaurants	-0.0057	(-1.15)	-0.0063	(-1.47)	-0.006***	(-5.98)	-0.0062***	(-6.14)
Fast food	0.0011	(0.30)	-		0.001	(0.54)	-	
Poor	-0.0059	(-1.16)	-0.0065	(-1.36)	-0.006	(-1.22)	-0.0038***	(-5.40)
Rich	-.0051	(-0.98)	-0.006	(-1.19)	-0.006	(-0.49)	-	
Model								
Imperfect competition	-0.0124	(-0.68)	-		-0.012**	(-2.21)	-0.0148**	(-3.16)
Difference in difference	0.0103	(0.19)	-		0.01	(0.41)	-	
VAR	-0.2539	(-1.42)	-0.1252	(-1.19)	-0.254***	(-11.44)	-0.2532***	(-10.32)
Conditioning variables								
Combination	-0.0039	(-1.05)	-0.0038	(-1.06)	-0.004	(-1.07)	-0.0045	(-1.64)
# of lagged months	-0.0000	(-0.13)	-		0	(-0.41)	-	
# of lead months	0.0019	(1.53)	0.0019†	(1.60)	0.002	(1.41)	0.002	(1.51)
City area control	0.0062	(1.42)	0.0053	(1.35)	0.006	(1.22)	0.0061	(1.21)
Unemployment change control	-0.006	(-0.63)	-		-0.006	(-1.45)	-0.0068	(-1.14)
Seasonality controls	-0.0014	(-0.21)	-		-0.001	(-0.57)	-0.0004	(-0.11)
Mean reversion	0.0143	(1.23)	0.0169*	(2.32)	0.014**	(2.05)	0.0128†	(2.03)
Input price shock	.0048	(0.89)	0.0037	(0.91)	0.005***	(3.85)	0.0061**	(3.06)
Real factors								
Dataset span	0.0016†	(1.80)	0.0012	(1.87)	0.002***	(4.41)	0.0016***	(4.78)
Average year	-0.0013	(-0.81)	-0.0016	(-2.13)	-0.001	(-0.96)	-0.0011	(-0.99)
EU	0.1545	(0.74)	-		0.155***	(4.30)	0.1549***	(4.38)
North America	0.1925	(0.89)	-0.0012†	(-0.07)	0.192***	(4.94)	0.1898***	(5.70)
N	469		469		469		469	
Authors	12		12		12		12	

Significance levels: †:10% *:5% **: 1% ***: 0.1%

Dependant variable: *t*-statistic of the partial correlation of the effect of minimum wage on inflation

t-statistics (z-statistics for ME) of the estimates in parenthesis (author level clustered errors)

ME - mixed effects multi-level model

OLS - ordinary lease squares estimation

Source: own analysis