UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Cambridge Working Papers in Economics

Cambridge Working Papers in Economics: 1889

DISTRIBUTIONAL EFFECTS OF LOCAL MINIMUM WAGE HIKES: A SPATIAL JOB SEARCH APPROACH

Weilong Zhang

3 December 2018

This paper develops and estimates a spatial general equilibrium job search model to study the effects of local and universal (federal) minimum wage policies. In the model, firms post vacancies in multiple locations. Workers, who are heterogeneous in terms of location and education types, engage in random search and can migrate or commute in response to job offers. I estimate the model by combining multiple databases including the American Community Survey (ACS) and Quarterly Workforce Indicators (QWI). The estimated model is used to analyze how minimum wage policies affect employment, wages, job postings, vacancies, migration/ commuting, and welfare. Empirical results show that minimum wage increases in local county lead to an exit of low type (education<12 years) workers and an influx of high type workers (education_12 years), which generates negative externalities for workers in neighboring areas. I use the model to simulate the effects of a range of minimum wages. Minimum wage increases up to $14/hour increase the welfare of high type workers but lower welfare of low type workers, expanding inequality. Increases in excess of $14/hour decrease welfare for all workers. I further evaluate two counterfactual policies: restricting labor mobility and preempting local minimum wage laws. For a certain range of minimum wages, both policies have negative impacts on the welfare of high type workers, but beneficial effects for low type workers.

Distributional Effects of Local Minimum Wage Hikes: A Spatial Job Search Approach*

Weilong Zhang†

December 3, 2018

Abstract

This paper develops and estimates a spatial general equilibrium job search model to study the effects of local and universal (federal) minimum wage policies. In the model, firms post vacancies in multiple locations. Workers, who are heterogeneous in terms of location and education types, engage in random search and can migrate or commute in response to job offers. I estimate the model by combining multiple databases including the American Community Survey (ACS) and Quarterly Workforce Indicators (QWI). The estimated model is used to analyze how minimum wage policies affect employment, wages, job postings, vacancies, migration/commuting, and welfare. Empirical results show that minimum wage increases in local county lead to an exit of low type (education$<$12 years) workers and an influx of high type workers (education$\geq$12 years), which generates negative externalities for workers in neighboring areas. I use the model to simulate the effects of a range of minimum wages. Minimum wage increases up to \$14/hour increase the welfare of high type workers but lower welfare of low type workers, expanding inequality. Increases in excess of \$14/hour decrease welfare for all workers. I further evaluate two counterfactual policies: restricting labor mobility and preempting local minimum wage laws. For a certain range of minimum wages, both policies have negative impacts on the welfare of high type workers, but beneficial effects for low type workers.

Keywords: spatial equilibrium, local minimum wage policy, labor relocation
JEL Code: J61, J63, J64, J68; R12, R13.

*I am truly indebted to Petra Todd, Chris Flinn and Sarah Moshary for their generous support and encouragement. I wish to thank Andrew Shephard, Camilo Garcia-Jimeno, Hanming Fang, Iourii Manovskii, Zhesheng Qiu, Hongxun Ruan, Qing Gong, Mallick Hossain, Amanda Chuan, David Neumark, Jean-Marc Robin, Luca Flabbi, Kai Liu, Shiko Maruyama and Shenggang Liu. I also seminar and conference participants at UNC, University of Cambridge, Royal Holloway University of London, Tsinghua University, CUHK, UTS, University of Melboune, SED Mexico city 2018, SOLE Toronto 2018, EALE conference Lyon 2018 and EEA-ESEM Cologne 2018 for their helpful comments. I specially thank Erin Wolcott for kindly sharing Help Wanted Online (HWOL) monthly data with me. All errors and omissions are mine.

†Faculty of Economics, Austin Robinson Building, Sidgwick Avenue, Cambridge, UK. CB3 9DD. Email: weilong.zhang@econ.cam.ac.uk

1 Introduction

Traditional minimum wage studies estimate local labor market employment and wage effects by comparing a group that experienced the minimum wage change to a similar group in nearby region that did not experience a change.[1] This approach can be problematic when local minimum wage changes are large, because substantial local minimum wage increases likely induce labor mobility and have spillover effects on neighboring areas.[2] A full accounting of minimum wage effects must take into account workers from all affected areas.[3] Furthermore, when faced with higher labor costs, firms may substitute lower productivity workers with higher productivity ones to keep profitable (Horton, 2017). Therefore, some workers may benefit from minimum wage increases, while others are adversely affected. This paper studies the distributional and welfare effects of local and universal (federal) minimum wage policies taking into account worker heterogeneity, spatial mobility, and minimum wages of varying magnitudes.

To this end, I develop a spatial general equilibrium model that extends Flinn (2006) to a spatial search context. The economy consists of two adjacent regions, similar to the cross-border contiguous county pairs in Dube et al. (2016). Workers are differentiated by their types and locations.[4] They receive job offers from local firms and from firms in a neighboring county. Workers accept a local offer if its value exceeds the value of unemployment. When considering offers from neighboring regions, workers require extra compensation to offset migration/commuting costs. Firms decide in which counties to post vacancies, where the number of vacancies is determined by a free entry condition. Given the assumption of random search, heterogeneous workers in all locations are contacted by firms at identical rates. An individual's productivity when meeting a firm is determined by his/her type and an idiosyncratic random matching quality. The bargained wage is determined by a surplus

[1] There is an ongoing debate concerning the effect of minimum wages on employment. See Card and Krueger (1994, 2000); Dube et al. (2007, 2010, 2016); Neumark (2001); Neumark et al. (2014a,b); Jardim et al. (2017).

[2] Recent studies have documented increased labor mobility induced by minimum wage changes, especially for low skilled workers (Monras, 2015; McKinnish, 2017).

[3] As of September 2017, 39 counties and cities have passed new minimum wage laws according to the UC Berkeley Labor Center. 23 out of 39 cities/counties have passed minimum wages of $15 or more, while the current federal minimum wage remains at $7.25. See http://laborcenter.berkeley.edu/minimum-wage-living-wage-resources/inventory-of-us-city-and-county-minimum-wage-ordinances/for more details.

[4] Ideally, type could be a summary statistic to rank workers expected productivity. I empirically use educational attainment as a proxy for worker types. Low type represents high school dropouts while high type represents high school graduates or more. According to 2015 the Current Population Survey (CPS), 5.8 percent workers are paid an hourly rate at or below federal minimum wage for the low type group, while this rate drops to 2.9 percent for the high type group.

division rule, subject to the minimum wage constraint, which left-truncates the original wage distribution (Flinn, 2006). The new wage structure is a continuous distribution with a mass point at the minimum wage level.

I estimate this spatial job search model using a Simulated Method of Moments estimator that combines county-level data moments from various sources. The migration and commuting flows are obtained from the American Community Survey (ACS). Local labor market conditions (hiring rates, separation rates and employment rates) are obtained from Quarterly Workforce Indicators (QWI) survey. The payroll share of firms' expenditures, and the ratio of job postings to workers come from the Economics Wide Key Statistics (EWKS) and the Conference Board Help Wanted Online (HWOL).

This model provides a framework to access the effects of minimum wage increases of a range of magnitude. Previous studies have focused on the most disadvantaged workers, without considering the welfare consequences for high type workers. To study the impacts of minimum wage increases for heterogeneous workers, my model incorporates four important effects. First, conditional on being employed, workers receive a higher wage from the same matches (the "*wage enhancement effect*"). Second, a minimum wage increase also causes a disemployment effect, because it dissolves marginally acceptable matches (the "*disemployment effect*"). Low type workers are more likely to be the marginally hired worker. Third, when firms are mandatory to pay workers more, they receive a smaller fraction of the surplus from same matches (the "*share reduction effect*"). Fourth, the probability of filling the vacancy with a high productivity worker increases in the higher minimum wage county but decreases in its neighboring county (the "*worker relocation effect*"). The incentive for firms to post vacancies is reduced in both counties, but especially in the county that does not change its minimum wage, due to negative spillover effects.

My analysis yields three main results. First, local minimum wage hikes have contrasting impacts on differentiated workers, expanding the inequality between low type and high type workers. Low type workers are adversely affected by higher minimum wages, primarily due to greater *disemployment effect*. For high type workers, the *wage enhancement effect* dominates the *disemployment effect* when the minimum wage level is less than \$14, above which the countervailing *disemployment effect* start to dominate. Therefore, the welfare of high type workers displays a hump shape with a peak at \$14/hour. When simulating the welfare difference of a range of minimum wages, the inequality between high and low type workers grows as the local minimum wage increases and reaches its peak at \$15.

Second, I use the estimated model to evaluate two policies: restricting labor mobility and preempting local minimum wage laws.[5] For a range of minimum wage values, I find that the welfare of high type workers is negatively impacted, but both policies have beneficial effects for low type workers. In the experiment of restricting spatial labor mobility, the low type workers in neighboring counties prefer two labor markets to be isolated when local minimum wage increases are large (above $10), because the cost of lost working opportunities is fully compensated by the benefit of eliminating spillover externalities. In the experiment of preempting local minimum wage laws, I compare local minimum wage changes with universal minimum wage changes. I find that low type workers prefer universal minimum wage hikes over local minimum wage hikes when the minimum wage change is moderate (below $14.5). The benefit of reducing spillover externalities outweighs the cost of a larger disemployment effect.

Third, I find the disemployment effect of a minimum wage increase is underestimated if one ignores labor mobility. On one hand, low educated workers tend to move away in response to a minimum wage increase and thus "disappear" from the "treated" county. On the other hand, they "reappear" in the neighboring area, contaminating the control group. I obtain with the model an estimate of the elasticity of employment with respect to the minimum wage equal to -0.073; ignoring labor mobility cuts this value in half to -0.034. The bias is most severe for counties with higher fractions of mobile workers.

My paper contributes to four broad strands of the literature. First, it is the first paper highlighting the negative spillover effects created by local minimum wage policies. There are a few recent papers documenting worker migration/commuting decisions are responsive to local minimum wage changes (Monras, 2015; McKinnish, 2017). However, this is the first paper linking labor flows with negative externalities for neighboring area workers. The insight that local policies may create externalities in the neighboring area through policy-induced migration is also discussed in the fiscal-federalism literature. For example, Serrato and Zidar (2016) studies the incidence of state corporate taxes on the welfare of workers, landowners and firm owners. In their model, a state tax cut reduces the tax liability and the cost of capital, attracting more establishments to move in. Cohen et al. (2011) studies the effects of marginal tax rates on migration decisions in the U.S., while Young and Varner (2011) and Moretti and Wilson (2017) focus on the geographic locations of top earners. Although policy-induced migration has already drawn significant attentions in the tax competition literature, my paper is the first application in the minimum wage context.

[5]The minimum wage preemption laws prohibit cities from enacting their own minimum wage laws. As of July 6, 2017, 25 states have passed such laws. See http://www.nelp.org/publication/fighting-preemption-local-minimum-wage-laws/ for a more comprehensive policy review.

My paper also contributes to the structural minimum wage literature. It extends Flinn (2006) by allowing for location-specific minimum wages and spatial mobility. Previous minimum wage studies usually assume one universal minimum wage for the whole labor market. (Eckstein and Wolpin, 1990; Van den Berg and Ridder, 1998; Flinn, 2006; Mabli and Flinn, 2007; Eckstein et al., 2011; Flinn and Mullins, 2015; Flinn et al., 2017) By extending the framework to multiple connected sub-markets, my model is able to incorporate geographical minimum wage variation for identification and evaluate the externalities of local minimum wage laws. The spatial search framework in my paper is similar to that of Meghir et al. (2015), which develops an equilibrium wage-posting model with formal and informal sectors. Their paper focuses on firm heterogeneity while I focus on worker heterogeneity. Other relevant spatial equilibrium frameworks include Coen-Pirani (2010); Baum-Snow and Pavan (2012); Kennan and Walker (2011); Schmutz and Sidibe (2016). By embedding local minimum wage policy into a spatial equilibrium model, my model allows examination of the effects of minimum wages on labor mobility, local employment, migration, wages and welfare.

This paper also explores the methodological implications for minimum wage studies that use adjacent counties as the control group. Starting with Card and Krueger (1994), cross-border comparisons became a popular method of studying the employment effects of minimum wage increases. For example, Dube et al. (2007, 2010, 2016) generalize this strategy to all contiguous county pairs and find small disemployment effects, consistent with Card and Krueger (1994). Although the cross-border design is persuasive, because of the geographic proximity between the treatment and control areas, there are concerns about the assumption that adjacent counties are unaffected, particularly when the minimum wage discrepancy between counties is large. I find that ignoring labor mobility leads to an underestimation of disemployment effects for two reasons. First, the unemployed workers move out of the "treated" area when they can not find jobs, and second, they move into neighboring areas, contaminating the control group.

Lastly, this paper contributes to the recent local labor market policy literature, emphasizing the potential externalities caused by place-based policies.[6] I show that low type workers, who are the intended beneficiaries of minimum wage policies, are actually worse-off after minimum wage increases. The estimates of moving cost confirms that taking the neighboring job is costly in general, which is consistent with the finding in Manning and Petrongolo (2017). While their paper argues that the probability of a random distant (at least 5km

[6] See Glaeser et al. (2008) and Enrico (2011) for reviews. Other recent papers include Kline (2010); Busso et al. (2013); Kline and Moretti (2013)

away) job being preferred to random local (less than 5km away) job is only 19% based on data from UK. Using county-level U.S. data, I find a slightly higher probability of 22.2%.

The structure of the paper is as follows. The next section presents a spatial job search equilibrium model. Section 3 describes the multiple data sources I will use to estimate the model. Section 4 discusses the identification and estimation strategy. Section 5 present the estimation results. Section 6 discusses the counterfactual experiments. Section 7 concludes.

2 Model

I develop a dynamic spatial search model where individuals live and work in one of the paired counties (j, j'). A job seeker in one county may receive either a local offer or a neighboring offer at certain rates. When a worker meets a firm in county j, they bargain over the wage subject to the minimum wage policy in county j. Local minimum wage changes would potentially affect labor market conditions in the neighboring county due to labor mobility.

2.1 Framework

I consider a continuous time model, where infinitely lived, risk neutral workers maximize their expected utility (income) with discount rate ρ. The economy consists of two adjacent local markets, a pair of counties (j, j'). The economy has a fixed number of potential workers with different types a. $N(a, j)$ represents the number of workers with type a in county j. Type is discrete, taking n different values $a \in A = \{a_1, ..., a_n\}$.[7] The number of workers for each type is exogenous. However, their working and living status are determined by the endogenous job searching process. $U(a, j)$, $L(a, j)$, and (a, j) represent the number of unemployed workers, local workers, and mobile workers with type a in county j. I focus on job search and labor mobility behavior in the steady state.

2.2 Worker's problem with wage w

A job seeker of type a in county j may receive wage offers from county j or j'. Upon meeting a firm, the productivity is given by

$$y = a\theta$$

[7]For computational tractability, I consider two types: high (a_h) and low (a_l) in the empirical analysis.

where θ is the random matching quality, which is assumed to be an i.i.d. draw from the distribution function $G(\theta)$.[8] I first consider the decisions for an unemployed worker with ability a in location j. At the beginning of each period, the worker experiences a set of location preference shocks $\{\vartheta_{uj}(a,j), \vartheta_{uj'}(a,j)\}$. The worker then chooses to become either a stayer to search jobs in the local labor market or a mover to search jobs the neighboring market. I use notation $d(a,j,\vartheta)$ to summary the dummy location choice, which has the following expression:

$$d(a,j,\vartheta) = \begin{cases} 0(stayer) & \text{if} \quad v_{uj}(a,j) + \vartheta_{uj}(a,j) > v_{uj'}(a,j) + \vartheta_{uj'}(a,j) \\ 1(mover) & \text{if} \quad v_{uj}(a,j) + \vartheta_{uj}(a,j) \leq v_{uj'}(a,j) + \vartheta_{uj'}(a,j) \end{cases}$$

an unemployed worker faces individual-specific preference shocks for locations they search for jobs. I use notation $V_u(a,j)$ to represent the ex-ante unemployed value before the preference shock realized. After the location preference shocks $\vartheta_{uj}(a,j)$ and $\vartheta_{uj'}(a,j)$ realized, the unemployed workers always choose whichever labor market generating a higher value, which is captured by the following equation:

$$V_u(a,j) = E_{\nu_{uj}\nu_{uj'}} \max\left\{ \; v_{uj}(a,j) + \vartheta_{uj}(a,j), \quad v_{uj'}(a,j) + \vartheta_{uj'}(a,j) \; \right\}$$

where $v_{uj}(a,j)$ is the systematic component of the ex-post unemployed values when workers decide to search job in local county j. And $v_{uj'}(a,j)$ is the systematic component when searching jobs in neighboring county j'. Assuming the preference shock $\{\vartheta_{uj}(a,j), \vartheta_{uj'}(a,j)\}$ follow i.i.d. type I extreme value distributions with a location parameter 0 and a common scale parameter σ_a^u, then the choice probability of each location is specified as:

$$P_j(a,j) = \frac{exp(v_{uj}(a,j)/\sigma_a^u)}{exp(v_{uj}(a,j)/\sigma_a^u) + exp(v_{uj'}(a,j)/\sigma_a^u)}, P_{j'}(a,j) = \frac{exp(v_{uj'}(a,j)/\sigma_a^u)}{exp(v_{uj}(a,j)/\sigma_a^u) + exp(v_{uj'}(a,j)/\sigma_a^u)}$$

I will now specify the constant part of the unemployed value under these two searching states:

[8] The assumption that the flow productivity $y_{ij} = a_i\theta_j$ is the multiplicity of a firm type θ_j and a worker type a_i is standard in the literature (Postel-Vinay and Robin (2002); Cahuc et al. (2006)). Following this spirit, the distribution of matching productivity should be location-specific (firm-specific) $G_j(\theta)$. Since labor market conditions in county pairs should be similar, I assume the matching productivity $G_j(\theta)$ is the same for these two counties.

$$\rho v_{uj}(a,j) = \underbrace{ab_j}_{\text{(1) flow value}} + \underbrace{\lambda_j \int_{m_j}^{\infty} \{V_e(w,j) - V_u(a,j)\}^+ dF(w|a,j)}_{\text{(2) option value of accepting a local offer}}$$

$$\rho v_{uj'}(a,j) = \underbrace{ab_j}_{\text{(1) flow value}} + \underbrace{\lambda_{j'} \int_{m_{j'}}^{\infty} \{V_e(w,j') - c(a,j) - V_u(a,j)\}^+ dF(w|a,j'))}_{\text{(3) option value of accepting a neighbouring offer}}$$

The notation $\{x\}^+ \equiv \max\{x, 0\}$. ab_j represents the flow utility of staying unemployed for workers with ability a and in location $j.\lambda_j$ is the job arrival rate in the local market and $\lambda_{j'}$ is the job arrival rate in the neighboring market. m_j and $m_{j'}$ represent the minimum wage level in county j and county j', respectively. Conditioning on location j and ability a, the agent draw a match-specific quality θ and bargaining wage w with firms. I will specify the bargaining process in the next session. The optimal job acceptance decision is maximizing between staying unemployed $V_u(a,j)$ and accepting the new employment status.

While the option values between local market and neighboring market have a very similar structure, the key distinguishable term is the moving cost $c(a,j)$ when accepting the neighboring offer, following similar assumption in Baum-Snow and Pavan (2012) and Schmutz and Sidibe (2016). If $c(a,j) = 0$, the workers in county j and county j' will have exactly the same working opportunities, which means paired counties are essentially one united labor market. If $c(a,j) = +\infty$, the paired counties are totally isolated markets. As pointed out by Schwartz (1973) and Greenwood (1975), this moving cost summarizes both the psychic cost of losing local social connections with family and friends and the physical transportation cost, which depends on the moving distance. The specifically parametric form of the moving cost will be shown in section 4.1.

I assume no on-the-job search. Therefore, the worker who accepts a job with wage w will never voluntarily quit the current job. Thus the existing matches only dissolve with a constant exogenous rate η_j. The value of employment, V_t^e, has the the following form[9]:

(1) $$V_e(w,a,j) = \frac{w + \eta_j V_u(a,j)}{\rho + \eta_j}$$

2.3 Bargaining with a minimum wage constraint

In this section, I specify how the wage between the worker and the firm is determined. I first consider the case without a minimum wage. If a worker with type a meets a firm in

[9]The derivations of equations **??** and 1 are described in Appendix A.1.

location j and draws a matching quality θ in period t, the bargained wage is assumed to be derived from a Nash bargaining solution. The wage $\hat{w}(a, j, \theta)$ maximizes the weighted product of the worker's and firm's net return from the match. To form the match, the worker gives up the value of unemployment $V_u(a, j)$, and the firm gives up the unfilled homogeneous vacancy, which has zero value according to the free entry condition.[10]

$$\hat{w}(a, j, \theta) = \arg\max_{w} \left(V_e(w, a, j) - V_u(a, j)\right)^{\alpha_j} V_f(w, a, \theta, j)^{1-\alpha_j}$$

where location specific bargaining weight α_j is strictly between 0 and 1, representing the relative bargaining strength of the labor side. V_f is the present value of the filled vacancy for the firm. As derived in Appendix A.2, the bargained wage offer function is:

$$\hat{w}(a, j, \theta) = \rho V_u(a, j) + \alpha_j(a\theta - \rho V_u(a, j)) \tag{2}$$

The interpretation of this bargained wage is intuitive. The workers receive their reservation wage $\rho V_u(a, j)$ and a fraction of bargained share α_j of the net surplus of the current match, which is the total production $a\theta$ minus what workers give up $\rho V_u(a, j)$.

Following Flinn (2006), the introduction of a minimum wage in area j is treated as a "side constraint" to the original bargaining problem.

$$w(a, j, \theta) = \arg\max_{w \geq m_j} \left(V_e(w, a, j) - V_u(a, j)\right)^{\alpha_j} V_f(w, a, \theta, j)^{1-\alpha_j}$$

The minimum wage constraint $w \geq m_j$ is imposed by local policy maker and applies to all potential job matches. Before considering the case when the minimum wage binds, I solve for the critical value of matching quality where the worker receives exactly the minimum wage based on the original surplus decision rule (Equation 2).

$$\hat{\theta}(a, j) = \frac{m_j - (1 - \alpha_j)\rho V_u(a, j)}{a\alpha_j}$$

If $\hat{\theta}(a, j) \leq \frac{m_j}{a}$, the minimum wage has no effect on the bargained wage because the reservation value is so high that all acceptable matches for workers actually give them wages equal or larger than m_j. (i.e. $a\theta^*(a, j) \geq m_j$). If $\hat{\theta}(a, j) > \frac{m_j}{a}$, the minimum wage is binding when $\theta \in [\frac{m_j}{a}, \hat{\theta})$. The firms in this scenario would pay workers m_j, which is more than the

[10] I do not model different outside options for local workers and mobile workers for two reasons. First, it is unclear whether moving costs are a credible "threat point" for mobile workers because they have to pay the moving cost before they can work in the other county. Second, due to menu costs, it is not economic for firms to make a separate wage offers for mobile workers who are a minority of new hires.

worker's "implicit" reservation wage $\hat{w}(a, j, \theta)$. Although payroll expenditure expands, it is still in firms' best interests to hire these workers, because destroying the jobs would reduce profits to zero. The binding minimum wage creates a wedge between the worker's wage and their "implicit" reservation wage, making the latter unobservable. Following Flinn (2006), I introduce the reservation matching quality $\theta^*(a, j)$, which is the lowest matching quality of a local match that a worker with type a will accept. In other words, the worker is indifferent between accepting a local job with matching quality $\theta^*(a, j)$ and staying unemployed.

$$V^e(\hat{w}(a, j, \theta^*(a, j)), a, j) = V^u(a, j)$$

$$\theta^*(a, j) = \frac{\rho V_u(a, j)}{a}$$

This reservation matching quality would be "implicit" in the case when the minimum wage binds ($m_j > \rho V_u(a, j)$).In this way, I obtain an affine mapping between the cumulative distribution of the matching quality, $G(\theta)$, and the cumulative wage distribution $F(w|a, j)$:

$$f_t(w|a, j) = \begin{cases} \frac{(a\alpha)^{-1} g(\tilde{\theta}(w,a,j))}{\tilde{G}(\frac{m_j}{a})} & w > m_j \\ \frac{\tilde{G}(\hat{\theta}(a,j)) - \tilde{G}(\frac{m_j}{a})}{\tilde{G}(\frac{m_j}{a})} & w = m_j \\ 0 & w < m_j \end{cases} \tag{3}$$

where $f(w|a, j)$ is the probability distribution function(PDF) of $F(w|a, j)$, $g(\theta)$ is the PDF of $G(\theta)$, and $\tilde{G}(\theta) = 1 - G(\theta)$ is the complementary function of the cumulative distribution function $G(\theta)$. $\tilde{\theta}(w, a, j) = \frac{w - (1-\alpha_j)\rho V_u(a,j)}{a\alpha_j}$ denotes the matching quality whose bargained wage is equal to w. The observed wage distribution consists of a point m_j with mass $\frac{G(\hat{\theta}(a,j)) - \tilde{G}(\frac{m_j}{a})}{\tilde{G}(\frac{m_j}{a})}$ and a continuous function (assuming $G(\theta)$ is continuous) when $\theta > \hat{\theta}$. Thus the bargained wage can be summarized as:

$$w(a, j, \theta) = \max\{m_j, \alpha_j a\theta + (1 - \alpha_j)\rho V_u(a, j)\} \tag{4}$$

It is worth to point out that a binding minimum wage affects the wages of all workers, but through different channels. For the workers with matching quality $\theta \in [\frac{m_j}{a}, \hat{\theta}(a, j))$, the minimum wage directly benefits them by boosting their wage to m_j. For workers with even higher matching quality $\theta \in [\hat{\theta}(a, j), \infty)$, the minimum wage changes their value of unemployment $\rho V_u(a, j)$.[11] To summarize, introducing the minimum wage as a side restriction

[11]However, the sign of this change is ambiguous, depending on the trade-off between the increase of

on Nash-bargained wages converts a continuous underlying productivity distribution into a mixed continuous-discrete accepted wage distribution, with a mass point at the minimum wage.

2.4 Migration/commuting trade-off

Next, I characterize the spatial strategies of the workers. To capture the different types of labor mobility observed in the data, I distinguish commuting from migrating by specifying a choice-specific moving cost $cc_h(a, j), h = \{0, 1\}$. The timing is as follows: (1) an offer from neighboring area j' arrives at rate $\lambda_{j'}$. (2) After the matching quality θ is realized, the worker decides to accept/reject the offer based on the trade-off between the wage offer $w(a, j', \theta)$ net of the ex-ante moving cost $c(a, j)$ and the value of unemployment, $V_u(a, j)$. (3) If the worker accepts the offer, the preference shock ε_h is realized and the worker chooses whether to commute or migrate.

A worker with type a continues to receive job offers from the neighboring county at rate $\lambda_{j'}$. The expected moving cost $c(a, j)$, is a function of the worker's type and location-specific characteristics. Following Schmutz and Sidibe (2016), I introduce a *"implicit" mobility compatible indifferent matching quality* $\theta^{**}(a, j)$, fulfilling the following condition:

$$V_u(a, j) + c(a, j) = V_e(\theta^{**}(a, j), a, j')$$

where j represents the worker's place of residence and thus j' will be the worker's place of work. The worker will accept the neighboring offer if and only if the matching quality of the offer exceeds the mobility compatible threshold $\theta \geq \theta^{**}(a, j)$. This match will also be sustainable for firms as long as $\theta \geq \frac{m_{j'}}{a}$. To summarize, the worker whose residence is in county j will accept a neighboring offer if and only if $\theta \geq \max\{\frac{m_{j'}}{a}, \theta^{**}(a, j)\}$.

After accepting the neighboring offer, workers have two alternatives. They can either work as migrants ($h = 1$), pay a lump-sum cost $cc_{h=1}(a, j)$, and become a native worker in county j' or work as commuters ($h = 0$) and pay a recurring commuting cost $cc_{h=0}(a, j)$. I use $cc_h(a, j)$ to represent the lump-sum equivalent cost. The choice-specific moving cost $cc_h(a, j)$ is a function of both the worker's type and physical distance between counties, as well as the cost differences of house renting between paired counties. Its exact parametric form will be discussed in Section 4.1.

In additional to the moving cost $cc_h(a, j)$, workers also receive an unobserved preference

expected income and the reduction of expected working opportunities.

shock ε_{ah}. The workers thus choose their lowest cost mobility option, $h(a,j)$:

$$h(a,j) = \begin{cases} 0 & \text{if } \varepsilon_{ah} - cc_0(a,j) > \varepsilon_{ah} - cc_1(a,j) \\ 1 & \text{if } \varepsilon_{ah} - cc_0(a,j) \leq \varepsilon_{ah} - cc_1(a,j) \end{cases}$$

Assuming the preference shock ε_{ah} follows an i.i.d. type I extreme value distribution with a location parameter 0 and a common scale parameter σ_a^c, then the ex-ante expected cost has the following analytic formula (Rust 1987):

$$\begin{aligned} c(a,j) &= \max\{\varepsilon_{a0} - cc_0(a,j), \varepsilon_{a1} - cc_1(a,j)\} \\ &= \sigma_a^c \log(\textstyle\sum_{h=0}^{1} \exp(-cc_h(a,j))/\sigma_a^c) + \sigma_a^c \gamma \end{aligned}$$

The probability of each choice is specified as:

$$Q_h(a,j) = \frac{exp(-cc_h(a,j)/\sigma_a^c)}{exp(-cc_0(a,j)/\sigma_a^c) + exp(-cc_1(a,j)/\sigma_a^c)} \tag{5}$$

2.5 Worker's optimal strategies

The worker's optimal strategies consists of the search location choices and the sequential job taking strategies in both locations. The local decision is fully described by the implicit reservation matching quality $\theta^*(a,j)$, while the moving decisions are summarized by both the implicit mobility compatible matching quality $\theta^{**}(a,j)$ and migration/commuting choice probability $Q_h(a,j)$.

Proposition 1. *OPTIMAL STRATEGIES*

For unemployed workers of type a in county j, the optimal strategy is:

- *Search in the local labor market with probability* $P_j(a,j)$
  - *Accept any local job with matching quality* θ *higher than* $\max\{\theta^*(a,j), \frac{m_j}{a}\}$
- *Search in the local labor market with probability* $P_{j'}(a,j) = 1 - P_j(a,j)$
  - *Accept any neighboring job with matching quality* θ *higher than* $\max\{\theta^{**}(a,j), \frac{m_{j'}}{a}\}$
    - *Choose to commute with probability* $Q_0(a,j)$
    - *Choose to migrate with probability* $Q_1(a,j)$

In the last part of this section, I describe the fixed point equation system that is used to solve for $\theta^*(a,j)$ and $\theta^{**}(a,j)$. By applying both the reservation matching quality $\theta^*(a,j)$

and mobility compatible matching quality $\theta^{**}(a,j)$ to Equation **??**, I get the following system of equations:[12]

(6)

$$
\begin{aligned}
a\theta^*(a,j) \quad = \quad & \underbrace{ab_j}_{\text{(1) Flow utility}} + \frac{\lambda_j}{\rho+\eta_j}[\underbrace{\boldsymbol{I}\,(\theta^*(a,j) < \frac{m_j}{a})(m_j - a\theta^*(a,j))\left(\tilde{G}(\hat{\theta}(a,j)) - \tilde{G}(\frac{m_j}{a})\right)}_{\text{(2) Local offer with wage } m_j} \\
+ \quad & \underbrace{\int_{\max\{\hat{\theta}(a,j),\theta^*(a,j)\}} a\alpha_j(\theta - \theta^*(a,j))dG(\theta)]}_{\text{(3) Local offer with wage } w_j > m_j} \\
+ \quad & \frac{\lambda_{j'}}{\rho+\eta_{j'}}[\underbrace{\boldsymbol{I}(\theta^{**}(a,j) < \frac{m_{j'}}{a})(m_{j'} - a\theta^*(a,j'))\left(\tilde{G}(\theta^{**}(a,j)) - \tilde{G}(\frac{m_{j'}}{a})\right)}_{\text{(4) Neighbouring offer with wage } m_{j'}} \\
+ \quad & \underbrace{\int_{\max\{\hat{\theta}(a,j'),\theta^{**}(a,j)\}} a\alpha_j(\theta - \theta^*(a,j'))dG(\theta)}_{\text{(5) Neighbouring offer with wage } w_{j'} > m_{j'}} \\
+ \quad & \underbrace{(\rho+\eta_{j'})\,(\frac{a(\theta^*(a,j) - \theta^*(a,j'))}{\rho} + c(a,j))\tilde{G}(\theta^{**}(a,j))]}_{\text{(6) The unemployed value difference between staying/moving}}
\end{aligned}
$$

with

$$
\begin{aligned}
\hat{\theta}(a,j) &= \tfrac{m_j - (1-\alpha_j)a\theta^*(a,j)}{a\alpha_j} \\
\hat{\theta}(a,j') &= \tfrac{m_{j'} - (1-\alpha_{j'})a\theta^*(a,j')}{a\alpha_{j'}} \\
\theta^{**} : V_u(a,j) + c(a,j) &= V_e(\theta^{**}(a,j), a, j')
\end{aligned}
$$

In equation 6, the value of the implicit matching quality $a\theta^*(a,j)$ consists of six components: (1) the instant flow utility ab when unemployed; (2) the expected value associated with a local offer with binding minimum wage m_j; (3) the expected value associated with a local offer with wage $w_j > m_j$; (4) the expected value associated with an acceptable neighboring offer with binding minimum wage $m_{j'}$; (5) the expected value associated with an acceptable neighboring offer with wage $w_{j'} > m_{j'}$; (6) the unemployed utility difference between staying and moving, which includes both the moving cost $c(a,j)$ and the change of the option value of being unemployed $a\theta^*(a,j) - a\theta^*(a,j')$.

The intuition of equation 6 is straightforward. The value difference between accepting the lowest acceptable job and remaining unemployed $a\theta^*(a,j) - ab_j$ reflects an opportunity cost, which is perfectly compensated by the expected premium of finding a better job in the future. This job could either be a local one or a neighboring one after paying the moving cost $c(a,j)$.

[12] The derivation of equation 6 can be found in Appendix A.3

2.6 Endogenous contact rate

In this section I consider how the contact rates $\lambda_j, j = 1, 2$, are determined in general equilibrium. I assume that firms randomly encounter workers with the same probability. This assumption captures the idea that workers applying for the same position may have different productivity but are easily to substitute with each other. I adapt the Mortensen and Pissarides (1994) framework and allow firms to post vacancies K_j in county j with constant marginal cost ψ_j which is open to all workers in both counties. The matching technology is assumed to be constant returns to scale. Let $N = \sum_{a\in A}(U(a, j) + U(a, j'))$ be the number of all job seekers in the economy, where $U(a, j)$ is the number of unemployed workers with type a in county j. If the firms in county j creates K_j vacancies, then the total number of potential matches created in county j, M_j, is given by

$$M_j = N^{\omega_j} K_j^{1-\omega_j}$$

where ω_j is the matching elasticity parameter in market j.

I use a Cobb-Douglas matching function with constant return to scale and total factor productivity equal to 1. It then only requires one parameter ω_j to characterize the heterogeneity of matching functions in each local labor market j.

The contact rate per job in county j, $q_j(k_j)$, can be represented as:

$$q_j(k_j) = k_j^{\omega_j}$$

where $k_j = \frac{N}{K_j}$ captures the market tightness. The correlation between market tightness and job arrival probability λ_j is

$$\lambda_j = k_j(K_j, N)^{\omega_j - 1} \tag{7}$$

It is important to emphasize that although workers in both counties have the exact same opportunities to meet with the same firm, their willingness to accept the same job is different due to moving costs. For workers living in the neighboring county, they are more picky about neighboring jobs because the job premium has to compensate for the additional moving cost.

The total number of matches created by the firms in county j is:

$$\text{Total Hires} = \frac{M_j}{N}\sum_{a\in A}\left(\underbrace{U(a,j)G\left(\max\{\theta^*(a,j),\frac{m_j}{a}\}\right)}_{\text{Local Hires}}+\underbrace{U(a,j')G\left(\max\{\theta^{**}(a,j'),\frac{m_j}{a}\}\right)}_{\text{Neighboring Hires}}\right)$$

The firm's value of a match can be represented as:

$$V_f(\theta,a,j)=\frac{a\theta - w(a,\theta,j)}{\rho+\eta_j} \tag{8}$$

The expected value of creating a vacancy for firms V_v in county j is:

$$V_v = -\psi_j + \frac{k_j(K_j,N)^{\omega_i}}{N}\sum_{a\in A}\left[\underbrace{U(a,j)\int_{\max\{\theta^*(a,j),\frac{m_j}{a}\}}V_f(\theta,a,j)dG(\theta)}_{\text{Profit from local workers}}+\underbrace{U(a,j')\int_{\max\{\theta^{**}(a,j'),\frac{m_j}{a}\}}V_f(\theta,a,j)dG(\theta)}_{\text{Profit from neighboring workers}}\right]$$

Assuming each county has a population of potential entrants with an outside option equal to 0, firms will continue to create vacancies until the expected profit is equal to 0, $V_v = 0$. Under the free entry condition (FEC), the endogenous contact rate is determinate by the following equation

$$\psi_j = \frac{k_j(K_j,N)^{\omega_i}}{N}\sum_{a\in A}\left[U(a,j)\int_{\max\{\theta^*(a,j),\frac{m_j}{a}\}}V_f(\theta,a,j)dG(\theta)+U(a,j')\int_{\max\{\theta^{**}(a,j'),\frac{m_j}{a}\}}V_f(\theta,a,j)dG(\theta)\right] \tag{9}$$

2.7 Definition of a steady-state spatial equilibrium

Let $\theta \in \boldsymbol{R_+}$, $a \in \boldsymbol{A} = \{a_1, a_2, ..., a_n\}$, $j \in \boldsymbol{J} = \{1, 2\}$, and let $\boldsymbol{S_1} = \boldsymbol{R_+} \times \boldsymbol{A} \times \boldsymbol{J}$ and $\boldsymbol{S_2} = \boldsymbol{A} \times \boldsymbol{J}$. Let $\boldsymbol{B(R_+)}$ be the Borel $\sigma-$algebra of $\boldsymbol{R_+}$ and $\boldsymbol{P(A)}$, $\boldsymbol{P(J)}$ the power sets of $\boldsymbol{A}$ and $\boldsymbol{J}$, respectively. Let $\boldsymbol{\aleph} = \boldsymbol{B(R_+)}\times\boldsymbol{P(A)} \times \boldsymbol{P(J)}$, and $\boldsymbol{M}$ be the set of all finite measures over the measurable space $(\boldsymbol{S_1}, \boldsymbol{\aleph})$

Definition 1. A steady-state spatial equilibrium is a set of individual functions for workers $V_u : \boldsymbol{S_1} \to \boldsymbol{R_+}$ and $V_e, \theta^*, \theta^{**}, P_h : \boldsymbol{S_2} \to \boldsymbol{R_+}$, a set of the functions for firms $V_f : \boldsymbol{S_1} \to \boldsymbol{R_+}$ and $\{K_j\}_{j=1,2}$, a set of contact rates $\{\lambda_j\}_{j=1,2}$ and wage rates $w : \boldsymbol{S_1} \to \boldsymbol{R_+}$ and a set of aggregate measures of different working status $U, L, M : \boldsymbol{S_2} \to \boldsymbol{R_+}$, the following conditions hold:

1. Worker's problem: given the contact rate, wage and initial condition, V_u and V_e are the solutions of Eqs. **??** and 1, respectively. The optimal strategies θ^*, θ^{**} are described in

Proposition 1 and $\{P_h\}_{h=0,1}$ are described in Eq. 5. The functions $\{V_u, V_e, \theta^*, \theta^{**}, P_h\}$ are measurable with respect to $\aleph$.

2. Firm's problem: given the contact rate, wage and initial condition, V_f is solved by Eq. 8 and K_j is solved by Eq. 9.

3. The bargained wage: the bargained wage with a minimum wage constraint is defined by Eqs. 3 and 4.

4. Endogenous contact rate (labor market clear): the contact rate λ_j is solved by Eq. 7.

5. The aggregate measures of working status keep constant

$$\begin{aligned}
\underbrace{\lambda_j \left(U(a,j)\tilde{G}(\max\{\theta^*(a,j), \frac{m_j}{a}\}) + U(a,j')P_0(a,j')\tilde{G}(\max\{\theta^{**}(a,j'), \frac{m_j}{a}\})\right)}_{\text{Inflow to L}} &= \underbrace{L(a,j)\eta_j}_{\text{Outflow from L}} \\
\underbrace{U(a,j)\left(\lambda_j \tilde{G}(\max\{\theta^*(a,j), \frac{m_j}{a}\}) + \lambda_{j'}\tilde{G}(\max\{\theta^{**}(a,j), \frac{m_{j'}}{a}\})\right)}_{\text{Outflow from U}} &= \underbrace{L(a,j)\eta_j + M(a,j)\eta_{j'}}_{\text{Inflow into U}} \\
\underbrace{\lambda_j U(a,j')P_1(a,j')\tilde{G}(max\{\theta^{**}(a,j'), \frac{m_j}{a}\}))}_{\text{Inflow to M}} &= \underbrace{M(a,j)\eta_j}_{\text{Outflow from M}}
\end{aligned}$$

3 Data and descriptive statistics

This paper primarily uses two data sets: the Quarterly Workforce Indicators (QWI) for local labor market information and the American Community Survey (ACS) for labor mobility information. QWI provides the number of job stocks and flows, and average earnings by industry, worker demographics, employer age, and size. The QWI comes from the Longitudinal Employer-Household Dynamics (LEHD) linked employer-employee micro data, which are collected through a unique federal-state sharing collaboration between the U.S. Census Bureau and state labor market agencies.[13] Compared to the CPS and JOLTS, the QWI has near-universal worker-employer paired information, covering 96% of all private-sector jobs. Second, QWI provides worker-side demographic information such as age, sex, race/ethnicity, and education.[14] This feature allows me to analyze the demographics of a particular industry or specific local market.[15] Lastly, QWI has labor flow information, including hires, separations, and turnovers, which are important because the direct impacts

[13] Data for Massachusetts, Puerto Rico, and the US Virgin Islands are still under development.

[14] Workers are identified by their Social Security number and linked with a variety of sources, including the 2000 Census, Social Security Administrative records, and individual tax returns to get their demographic information.

[15] While CPS contains similar information based-on household surveys, it generates small sample sizes when analyzing individual industries or areas.

of minimum wage hikes are on job turnovers rather than employment stocks.[16] I focus on 2005-2015 primarily because the states missing from QWI before 2005 are not random - smaller states are under-represented. By 2005, all states except Massachusetts have joined the QWI program.[17]

In addition to QWI data, I also use the ACS from 2005-2015 to identify the commuting and migration flows between different jurisdictions.[18] Commuters are defined as people whose place of work is different from their place of residence, while migrants are defined as those who have changed their place of residence in the past year, according to the ACS. The basic geographic units in the ACS are "Public Use Micro Areas"(PUMAs) which are special non-overlapping areas that partition each state into contiguous geographic units containing between 100,000 to 300,000 residents. There were a total of 2,071 PUMAs in the 2000 census.

3.1 Contiguous border county pairs and their associated geographic minimum wage variations

Following the contiguous county-pair design proposed by Dube et al. (2010, 2016), I divide all counties in the U.S. into two sub-samples: counties that border another state (border counties), and counties that do not (interior counties). Out of 3,124 counties, 1,139 counties are border counties and I construct 1,181 unique pairs.[19] Figure 1 shows the locations of all counties along with their associated minimum wage policies. Between 2005 and 2015, there were 332 minimum wage adjustments (see 13 for details of minimum wage policies). While 78 changes are driven by the federal minimum wage law, the Fair Minimum Wage Act of 2007,[20] the other 164 events were due to state ordinances. Two observations are highlighted on the map. First, border counties frequently adjust their minimum wages. Between 2005 and 2015, all counties (except for those in Iowa) changed their local minimum wage at least three times, which gives me adequate variation to identify the effects of minimum wage hikes. Second, western counties are larger than other counties. Thus, the workers in those counties may suffer higher moving costs when working in a neighboring county.

In a given year, about half of the county pairs have different minimum wages. These differences average about 10%, but there is substantial heterogeneity across years (see Table

[16] See Dube et al. (2010, 2016) for detailed discussions.

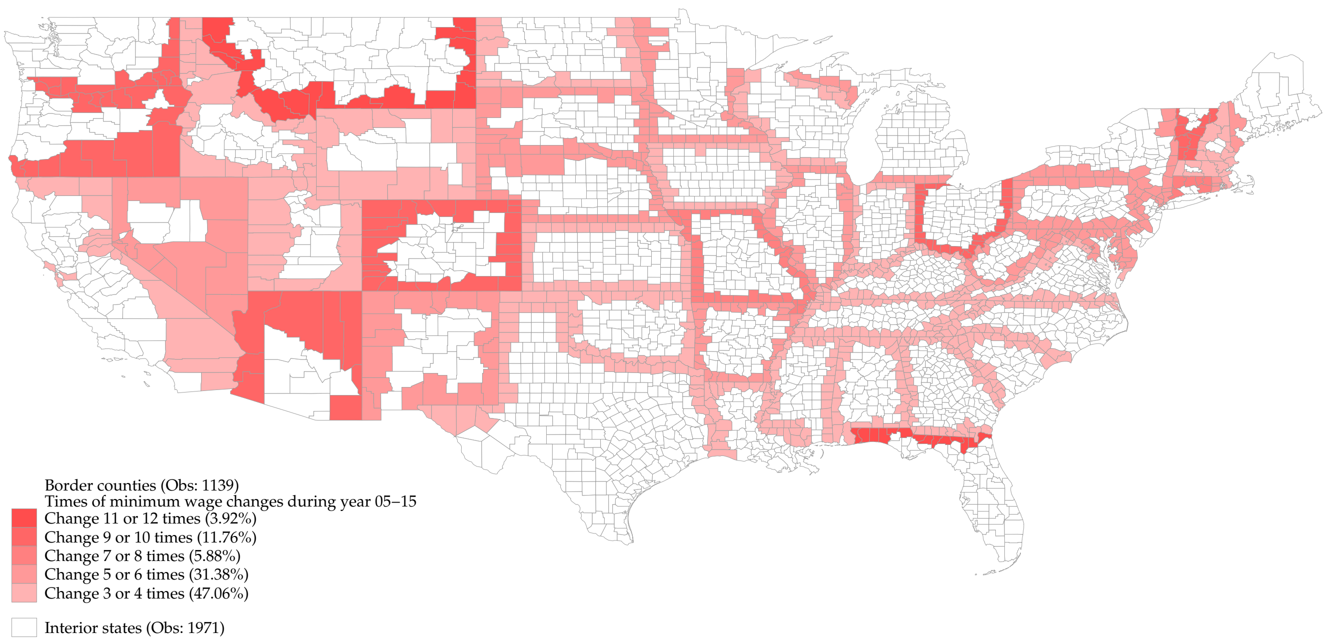
[17] Massachusetts does not join the QWI until 2010.

[18] I combine the 2005-2007, 2008-2010, and 2011-2015 ACS.

[19] Counties may border more than one county in the adjacent state, resulting in more pairs than border counties.

[20] The Act raised the federal minimum wage in three stages: to \$5.85 60 days after enactment (2007-07-24), to \$6.55 one year after that (2008-07-24), then finally to \$7.25 one year after that (2009-07-24).

Figure 1: Frequency of Minimum Wage Adjustments for Border Counties (2005-2015)



1). Overall, the substantial variation between county minimum wages is useful for identifying the effect of minimum wage hikes.

3.2 Migration and commuting flows

I use the American Community Survey (ACS) Public Use Microdata Sample (PUMS) data between 2005-2015 to distinguish commuters and migrants. Each respondent provides information about their place of residence one year ago, their current residence, and their current working address. To perform policy analysis, I convert PUMAs into pseudo-counties using the Michigan Population Studies Center PUMA-to-County crosswalk.[21]

To construct a sample of workers most sensitive to minimum wage changes, I restrict my sample to individuals between 16 and 30 that live in the continental U.S. and are not currently in the military. I divide this sample into two groups based on education: the low educated group (high school dropouts group) and the high educated group (the high school graduates and above). These restrictions are commonly used in the literature because young people and least-educated people are more likely to be minimum wage workers (Deere et

[21] I do this for two reasons. First, since PUMAs are population-based, they are not natural jurisdictions for local policy analysis. In urban areas, a single county may contain multiple PUMAs. For example, Los Angeles County, California is comprised of 35 PUMAs. Likewise, a PUMA will consist of several counties in less population areas. Second, I want to match the ACS to county-based statistics from the QWI. See Appendix C.2 and http://www.psc.isr.umich.edu/dis/census/Features/puma2cnty/for details.

Table 1: Differences in County Pair Minimum Wages (2005-2015)

Year	Share of pairs with minimum wages differential	Percent difference in minimum wages
2005	27.6%	18.6%
2006	33.6%	19.1%
2007	66.0%	15.6%
2008	63.7%	11.1%
2009	52.2%	8.7%
2010	31.8%	5.8%
2011	36.2%	6.0%
2012	37.8%	7.7%
2013	44.1%	7.4%
2014	49.0%	8.6%
2015	68.5%	9.4%
Average	46.4%	10.7%

al. (1995); Burkhauser et al. (2000); Neumark (2001)). If the minimum wage effect is not significant for this group, then it is unlikely to be significant for other groups.

Local governments prioritize their residents over residents of neighboring counties and as a result, I carefully distinguish between migrants (who have moved out of a county) and commuters (who might work in neighboring counties). Descriptive statistics for both commuting outflows to other states and migration inflows from other states are provided in Table 2.[22] Migrants are defined as individuals whose county of residence last year differs from their current county of residence. Commuters are defined as workers whose state of work differs from their state of residence. The rate (a value between 0 and 1) represents the share of commuters in the labor force. All statistics are on county-level and are grouped by whether they are border or interior counties. Border counties have higher migration and commuting rates, likely because commuting and moving costs are lower (See Table 2)

I further estimates some regression models to explore how migration flows and commuting flows respond to the local minimum wage hikes. The results suggest that low educated workers tend to move away from rather than move towards counties with minimum wage increases, either by commuting or migration. In contrast, the high educated workers, who are served as the control group, are less responsive to the minimum wage changes. And

[22]The other two potential measures of labor mobility patterns are commuting inflows and migrating outflows. They are in principle able to be calculated by summarizing all workers who migrate from/commute into the targeted PUMA in the sample. However, this calculation suffers from serious measurement error because the migrants from the particular PUMA and the commuters working in the particular PUMA are a small minority in other PUMAs and thus unlikely to be sampled.

Table 2: Summary Statistics of Migrants and Commuters (2005-2015)

		Interior counties		Border counties		Difference	
		Count	Rate	Count	Rate	Count	Rate
ALL workers							
Migrants	Mean	231	0.040	266	0.051	35.0	0.011
	S.D.	749	0.038	829	0.047	(8.96)	(0.0005)
Commuters	Mean	44.9	0.019	210	0.066	165	0.047
	S.D.	138	0.078	718	0.127	(6.48)	(0.0013)
Low educated group							
Migrants	Mean	28.5	0.024	31.5	0.030	3.00	0.006
	S.D.	95.4	0.033	87.2	0.040	(1.01)	(0.0004)
Commuters	Mean	4.49	0.021	20.1	0.047	15.6	0.026
	S.D.	20.2	0.090	74.7	0.118	(0.681)	(0.0012)
High educated group							
Migrants	Mean	203	0.045	235	0.058	22.0	0.013
	S.D.	674	0.043	770	0.053	(8.25)	(0.0005)
Commuters	Mean	40.4	0.031	189	0.070	149	0.039
	S.D.	125	0.097	656	0.130	(5.92)	(0.0013)
Observation		22,033		12,518			

*Data Source: ACS. Note: All statistics are reported at the county level. The count of migrations reports the number of individuals in each county whose place of residence last year differs from the place this year. The rate (a value between 0 and 1) is the percent of migrants in the local population. The count of commuters is the total number of workers whose state of work differs from the state of current residence. The rate (a value between 0 and 1) represents the percent of these commuters among the people who are currently in the labor force. Difference is border minus interior. * for 10%. ** for 5%, and *** for 1%.*

Table 3: County-Level Labor Market Summary Statistics (2005-2015)

	Interior counties		Border counties	
	Mean	SD	Mean	SD
Monthly earnings	1932	739	1930	739
Employment	14883	54878	13045	45968
Separation rates	0.299	0.111	0.301	0.103
Hire rates	0.326	0.171	0.326	0.128
Labor force participation rate				
All	0.618	0.199	0.623	0.197
High educated	0.701	0.222	0.704	0.219
Low educated	0.394	0.161	0.399	0.162

Note: All statistics are quarterly and from Quarterly Workforce Indicators except labor force participation, which is from the American Community Survey. Monthly earnings are in nominal dollars.

these mobility patterns are robust to the following sensitivity analysis: (1) use alternative migration flows based on addresses on the income tax returns provided by the Internal Revenue Service (IRS); (2) using only the minimum wage changes caused by federal minimum wage laws; (3) restricting to county pairs whose centriods are within 75 kilometers. The detailed regression results are reported in Table 11 in Appendix B.1.

3.3 Local labor market outcomes

From the QWI, I extract four quarterly variables: average monthly earnings, employment, hire rates, and separation rates. To make the QWI sample comparable to the ACS sample, I restrict worker's age to be between 19-34.[23] Labor force participation is extracted independently from the ACS. Overall, border and interior counties are similar across labor market statistics (Table 3).

In Appendix B.1, I estimates a regression models following Dube et al. (2007) and Dube et al. (2016) to examine the magnitude of disemployment in response to minimum wage increases. When using a common time fixed effect in column (1) in Table 11, the estimated disemployment elasticity is -0.068. However, this disemployment effect shrinks to -0.039 in column (2) when I replace the common time fixed effect with a pair-specific time fixed effect as the control. I attribute this change to the existence of spatial spillover effect. After the local county increases its own minimum wage, unemployed workers may seek their jobs in the

[23]The division of age groups in QWI are 19-21, 22-24, 25-34, 35-44, 45-54, 55-64, and 65-99. To match with the selected ACS sample whose ages are between 16-30, I combine the first four age spans 14-18, 19-21, 22-24, and 25-34.

neighboring county (either by migration or by commuting), which causes disemployment in the neighboring county. As a result, this spillover effect generates a common trend between counties in one pair. When this pair-specific co-movement is controlled by the pair-specific time effect, the estimates of local disemployment effect become less substantial.

4 Estimation strategy

4.1 Parametrization

To estimate the model, I need to make parametric assumptions for the types and moving costs. To be consistent with the data, I assume workers are of two types: a_h and a_l. High type workers are workers with high school diplomas and above while low type workers are high school dropouts. The proportion of these two types of workers are p_h and p_l.

I assume moving costs depend on a linear combination of worker's type a, the physical distance $d_{jj'}$ as well as the amenity difference $\gamma_j - \gamma_{j'}$ between the two counties.

$$cc_h(a, j) = \begin{cases} \beta_{0j} + \beta_{0d} d_{jj'} + \beta_{0a} I(a = a_h) + \beta_{0\gamma}(\gamma_j - \gamma_{j'}) & \text{if } h = 0 \\ \beta_{1j} + \beta_{1d} d_{jj'} + \beta_{1a} I(a = a_h) + \beta_{1\gamma}(\gamma_j - \gamma_{j'}) & \text{if } h = 1 \end{cases} \tag{10}$$

Equation 10 follows the standard gravity equation for migration. β_{hj} measures the relative openness of labor market j, which is county-specific and differs by the mobility choice h. The different impacts of distance on migrants and commuters are captured by β_{0d} and β_{1d}.[24] I also assume the moving costs to be differ by types a. The coefficientsβ_{0a} and β_{1a} represent the additional costs paid by high type workers. Lastly, I attribute the asymmetry between the cost $cc_h(a, j)$ and the cost $cc_h(a, j')$ to different housing rental prices γ_j and $\gamma_{j'}$, which are proxies of local living cost.

Assuming parametric distribution for matching quality is necessary for identification purposes. As Flinn and Heckman (1982), only a certain class of distributions satisfies the "recovery condition" necessary for identification. Following Flinn (2006) and Flinn and Mullins (2015), I assume the matching quality distribution $G(\theta)$ follows a log-normal distribution. Given the above assumptions, the economy is characterized by the vector S which combines

[24] While the distance between centroids is only a proxy for the real commuting time between two counties, some evidence shows the correlation between these two measures is quite high (Phibbs and Luft (1995);Boscoe et al. (2012)).

a set of general parameters and a set of county-specific parameters.

$$\begin{aligned} S &= \{\rho, \mu_G, \sigma_G, a_h, a_l, \beta_{0d}, \beta_{0a}, \beta_{0r}, \beta_{1d}, \beta_{1a}, \beta_{1\gamma}, \sigma_0, \sigma_1\} && \text{General} \\ &\bigcup \{m_j(n), b_j(n), \eta_j(n), \psi_j(n), \alpha_j(n), \omega_j(n), \gamma_j(n), d_{jj'}(n), \beta_{0j}(n), \beta_{1j}(n), p_h(n), p_l(n)\}_{(j,n)\in\{1,2\}\times N} && \text{County} \end{aligned}$$

The county-pair specific parameters are unique for every $n \in N$, while the general parameters are shared by all counties. Although the general parameters simplify the estimation, the model remains computationally demanding if the county-pair specific parameters are recovered non-parametrically. For tractability, I impose random coefficient assumptions the unobserved county-specific variables $s_j(n) \in \{b_j(n), \psi_j(n), \beta_{0j}(n), \beta_{1j}(n)\}$.[25] Given the close connection between the paired counties, I draw $s_1(n)$ and $s_2(n)$ from a multivariate normal distribution modeled for each $s_j(n) \in \{b_j(n), \psi_j(n), \beta_{0j}(n), \beta_{1j}(n)\}$:

$$\begin{pmatrix} x_{s1} \\ x_{s2} \end{pmatrix} \sim N\left(\begin{bmatrix} \mu_s \\ \mu_s \end{bmatrix}, \begin{bmatrix} \sigma_{s1}^2 & \rho_s\sigma_{s1}\sigma_{s2} \\ \rho_s\sigma_{s1}\sigma_{s2} & \sigma_{s2}^2 \end{bmatrix} \right)$$

where the correlation ρ_s captures the similarity between these two counties. The random variables $s_j(n), j = 1, 2$ are the mapping from the $n - th$ draw of the following one-to-one mapping F (which is 6×1),

$$\begin{aligned} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} &\sim N\left(\begin{bmatrix} \mu_b \\ \mu_b \end{bmatrix}, \begin{bmatrix} \sigma_b^2 & \rho_b\sigma_b^2 \\ \rho_b\sigma_b^2 & \sigma_b^2 \end{bmatrix} \right) \\ \begin{pmatrix} \log\psi_1 \\ \log\psi_2 \end{pmatrix} &\sim N\left(\begin{bmatrix} \mu_\psi \\ \mu_\psi \end{bmatrix}, \begin{bmatrix} \sigma_\psi^2 & \rho_\psi\sigma_\psi^2 \\ \rho_\psi\sigma_\psi^2 & \sigma_\psi^2 \end{bmatrix} \right) \\ \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} &\sim N\left(\begin{bmatrix} \mu_{\beta0} \\ \mu_{\beta1} \end{bmatrix}, \begin{bmatrix} \sigma_{\beta0}^2 & \rho_\beta\sigma_{\beta0}\sigma_{\beta1} \\ \rho_\beta\sigma_{\beta0}\sigma_{\beta1} & \sigma_{\beta1}^2 \end{bmatrix} \right) \end{aligned}$$

Thus, the joint distributions of these six variables are fully characterized by 11 parameters: 4 means, μ_s; 4 variances, σ_s^2; and 3 correlations, ρ_s. These parameters $(\mu_s, \sigma_s, \rho_s : s \in \{b, \psi, \beta_0, \beta_1\})$, in addition to those general parameters $\{\rho, \mu_G, \sigma_G, a_h, a_l, \beta_{0d}, \beta_{0a}, \beta_{0r}, \beta_{1d}, \beta_{1a}, \beta_{1\gamma}, \sigma_0, \sigma_1\}$, constitute the primitive parameters Ω of the model.

[25] The other county-specific parameters $\{m_j(n), \alpha_j(n), \gamma_j(n), \eta_j(n), d_{jj'}(n), p_h(n), p_l(n)\}_{j=1,2}$ are directly observed in the data.

4.2 The method of simulated moments

My model is estimated by the method of simulated moments (MSM). When combining moments from multiple databases, MSM is a more natural estimation approach than maximum likelihood estimation (MLE).

Given Ω, I draw the unobserved variables $\left\{b_j^r, \psi_j^r, \beta_{0j}^r, \beta_{1j}^r\right\}_{j=1,2}$ R times from the distributions of F for each county pair n. Combined with other observed county-level variables $\{m_j(n), \alpha_j(n), \gamma_j(n), \eta_j(n), d_{jj'}(n), p_h(n), p_l(n)\}_{j=1,2}$ and general parameters $\{\rho, \mu_G, \sigma_G, a_h, a_l, \beta_{0d}, \beta_{0a}, \beta_{0r}, \beta_{1d}, \beta_{1a}, \beta_{1\gamma}, \sigma_0, \sigma_1\}$, I then compute the vector of moments $\tilde{M}_{N,R}(\Omega)$ from the simulation. Model parameters are estimated by minimizing the weighted difference between those simulated moments $\tilde{M}_{N,R}(\Omega)$ and the actual data moments M_N, using the following quadratic distance function

$$\hat{\Omega}_{N,R,W} = \arg\min_{\Omega} \Big((M_N - \tilde{M}_{N,R}(\Omega))' W_N (M_N - \tilde{M}_{N,R}(\Omega)) \Big)$$

where M_N denotes the data moments for all county pairs in the data set, and $\tilde{M}_{N,R}(\Omega)$ represents the simulated moment evaluated at Ω based on R simulations of N county pairs. W_N is a symmetric, positive-definite weight matrix constructed using the resampling method of Del Boca et al. (2014). In particular, the resampled moment vector $M_N^g, g = 1, ..., Q$ is calculated by bootstrapping the original data Q times.[26] Then, the weight matrix is the inverse of the covariance matrix of M_N:

$$W_n = Q^{-1} \left(\sum_{g=1}^{Q} (M_N^g - M_N)(M_N^g - M_N) \right)^{-1}$$

Del Boca et al. (2014) show the consistency of this type of estimator for large simulations, $plim_{R\to\infty} \tilde{M}_{N,R}(\Omega_0) = M_N(\Omega_0)$.[27] Given identification and these regularity conditions,

$$plim_{N\to\infty} plim_{R\to\infty} \hat{\Omega}_{N,R,W} = \Omega \text{ for any positive definite } W$$

[26] In practice, I set Q equal to 200.

[27] Compared with directly calculating the optimal weighting matrix, this method simplifies computation significantly. Altonji and Segal (1996) discuss that gains from using an optimal weighting matrix may be limited.

4.3 Identification and selection of moments

My model is not nonparametrically identified, for reasons related to those given in Flinn and Heckman (1982) and Flinn (2006). However, it is useful to briefly discuss the identification in the model of Flinn (2006) given its close relationship with this paper. The model in Flinn (2006) can be regarded as a special case of my model when there is only one type of worker ($a_l = a_h$), one pair of counties and no labor mobility ($M(a, j) = 0$). The only job search decision for the worker is θ^*. Even in this specific case, the model is still unidentified because accepted wage and duration information is not enough for nonparametric identification. He further shows that a center class of parametric distributional assumption G, referred to as the "recoverability condition", is required.[28] In my model, given the assumed log-normal distribution of matching quality, all parameters are identified except for the set of discount factor and unemployment utility(ρ, b) because those parameters enter into the likelihood function through the critical value θ^*. Parameters b, η, G, λ will be identified given a fixed value of discount factorρ. Although I use the moments-based estimator rather than the likelihood-based estimator in Flinn (2006), their identification argument can be carried over in this paper given the same log-normal distribution assumption of θ and ex-ante fixed value of ρ.[29]

This paper extends Flinn (2006) in two dimensions by incorporating multiple worker types and multiple connected markets. As a result, instead of one critical value θ^*, individuals make two optimal decisions: accept local offer if $\theta \geq \theta^*(a, j)$ and accept neighboring offer if $\theta \geq \theta^{**}(a, j)$. Now I focus my attention on the log-wage distribution in one local county j. There are four different group of workers: high type natives, low type natives, high type movers, and low type movers. Given Equation 2, the log-wage distribution of local workers and the distribution of mobile workers only differ in the truncated values of their distributions. Besides the truncated log normal distribution, there is also a mass point as the left end at value m_j when the minimum wage is binding. As a result, the log wage distribution R should be a left truncated normal distribution with a potential mass point at its left end. We use R_0 to represent the distribution for natives and R_1 to represent the

[28] A comprehensive discussion about this "recoverability condition" can be found in Flinn and Heckman (1982).

[29] The identification depends on the proper selection of moments to characterize the wage distribution. I discuss this in Table 4.

distribution for movers.

$$\begin{aligned} \text{Natives} \quad \log w(\theta,a,j) &\sim \quad R_0(\log w; a, \mu_\theta, \sigma_\theta, \alpha_j, m_j, \theta^*) \\ \text{Movers} \quad \log w(\theta,a,j) &\sim \quad R_1(\log w; a, \mu_\theta, \sigma_\theta, \alpha_j, m_j, \theta^*, \theta^{**}) \end{aligned}$$

Since the fractions of local workers $L(a,j)$ and Mobile workers $M(a,j)$ are observed for the four groups of workers, it is straightforward to verify that the parameters $\mu_\theta, \sigma_\theta, a_l, a_h, \alpha_j, \theta^*$ are identified directly. To identify θ^{**}, one additional support condition $\theta^{**}(a,j) > \frac{m_j}{a}$ should be satisfied. Otherwise, the mobile worker's wage distribution R_1 would be identical to local workers' wage distribution, leaving $\theta^{**}(a,j)$ unidentified.

Therefore, I use the fraction of movers Fr, to help identify $\theta^{**}(a,j)$ as well as the moving cost term $c(a,j)$. First of all, I note that

$$Fr(a,j) = \frac{\tilde{G}(\max\{\theta^{**}(a,j), \frac{m_{j'}}{a}\})}{\tilde{G}(\max\{\theta^{*}(a,j), \frac{m_j}{a}\})}$$

Given that $\tilde{G}$ and $\theta^*(a,j)$ are already identified, the critical value $\theta^{**}(a,j)$ is identified directly from the observed $Fr(a,j)$.[30] Moving costs can then be backed out from the following one-to-one mapping:

$$c(a,j) = \frac{\alpha_{j'} a(\theta^{**}(a,j) - \theta^*(a,j'))}{\rho + \eta_{j'}} + \frac{a(\theta^*(a,j') - \theta^*(a,j))}{\rho}$$

Given the identified $c(a,j)$ and observed migration/commuting choices $P_0(a,j)$ and $P_1(a,j)$, the choice-specific moving cost $cc_0(a,j)$ and $cc_1(a,j)$ are identified by the logit assumption of equation 10.

Althrough the bargaining power α_j can be identified from R_0 and R_1, Flinn (2006) uses a Monte Carlo experiment to show its practical power is tenuous. Because of this, I use the average payroll share of firms' expenditures from the Economy Wide Key Statistics (EWKS), which is the U.S. government's official five-year measure of American business and the economy. This payroll share is calculated at the county level and provides cross-sectional variation of the labor share α_j.

The identification of the vacancy cost ψ_j follows from Equation 9 as long as the matching technology ω_j is known. Flinn (2006) uses multiple cross sections with different minimum wages to identify ω_j based on the assumption that the economy is in a steady-state in both

[30] $\theta^{**}(a,j)$ is potentially not identified when $Fr(a,j) > 1$, which means the number of movers are larger than the number is local workers. However, this situation rarely happens empirically.

Table 4: Selection of Moments

Empirical moments	County j Mean	County j S.D.	County j' Mean	County j' S.D.	Identified Parameters
Moments from mean and S.D. in county pair $p(j, j')$					
Employment rate (high type)	0.881	0.083	0.886	0.078	$\mu_b, \sigma_b, \mu_\psi, \sigma_\psi$
Employment rate (low type)	0.785	0.127	0.761	0.127	$\mu_b, \sigma_b, \mu_\psi, \sigma_\psi$
Average hourly wage (high type)	13.63	2.47	-	-	$\mu_b, \sigma_b, a_h, \mu_G, \sigma_G$
Average hourly wage (low type)	9.23	2.57	-	-	$\mu_b, \sigma_b, a_l, \mu_G, \sigma_G$
Proportion of migrants (high type)	0.073	0.050	0.070	0.046	$\mu_{\beta 0}, \sigma_{\beta 0}, \beta_{0a}$
Proportion of migrants (low type)	0.042	0.037	0.037	0.039	$\mu_{\beta 0}, \sigma_{\beta 0}, \beta_{0a}$
Proportion of commuters (high type)	0.113	0.127	0.096	0.106	$\mu_{\beta 1}, \sigma_{\beta 1}, \beta_{1a}$
Proportion of commuters (low type)	0.084	0.102	0.072	0.091	$\mu_{\beta 1}, \sigma_{\beta 1}, \beta_{1a}$
Correlation between migrants and commuters	0.630	-	0.523	-	ρ_β
Correlation between migrants and distance	0.149	-	0.014	-	β_{0d}
Correlation between commuters and distance	0.008	-	-0.168	-	β_{1d}
Correlation between migrants and rent cost	-0.103	-	-0.056	-	$\beta_{0\gamma}$
Correlation between commuters and rent cost	-0.116	-	-0.110	-	$\beta_{1\gamma}$
Correlation between migrants and rent cost					
Correlation between employments (high type)	0.318	-	-	-	ρ_b, ρ_ψ
Correlation between employments (low type)	0.211	-	-	-	ρ_b, ρ_ψ
Correlation between separation rate	0.599	-	-	-	ρ_ψ
Correlation between wage rate	0.498	-	-	-	ρ_b, ρ_ψ
Moments directly measure parameter values					
Separation rates in county j(quarterly)	0.353	0.130	0.358	0.132	η_j
Bargaining power α_j in county j	0.311	0.044	0.310	0.043	α_j
Matching technology ω_j in state $s(j)$	1.36	0.385	1.41	0.406	ω_j
Centroid distance $d_{jj'}$ between j and j'	66.6	45.9	66.6	45.9	$d_{jj'}$
The median rent cost (local amenity γ_j) in j	683	168	683	178	γ_j

Note: (i) For details about the construction of the empirical moments, see Appendix C. (ii) County j represents the county which increases its minimum wage, while county j' is the county keeps the minimum wage fixed.

measurements and the vacancy cost is constant.[31] In this paper, I use a market tightness index (job demand/labor supply) constructed from the Conference Board Help Wanted On-Line (HWOL) data, which is widely used in the macroeconomic literature as a direct measure of matching technology that does not impose any additional assumptions.[32][33]

Table 4 summarizes the empirical moments used to identify the model parameters.

[31] See the discussion of Condition C-Coherency in Flinn (2006) for more details.

[32] Beginning in 2005, HWOL provides a monthly series that covers the universe of vacancies advertised on about 16,000 online job boards and online newspaper editions. While HWOL only collects the job openings advertised online, its pattern is quite similar with the general pattern measured by JOLTS, especially before 2013. A detailed comparison between HWOL and JOLTS can be found in Şahin et al. (2014).

[33] See Petrongolo and Pissarides (2001) for a survey of these studies.

Table 5: Model fit

Empirical moments	County 1		County 2	
	Data	Sim	Data	Sim
Employment rate (high type)	0.883	0.829	0.888	0.827
Employment rate (low type)	0.754	0.789	0.765	0.786
Hire rate	0.375	0.354	0.361	0.348
Average hourly wage (high type)	13.630	13.385	-	-
Average hourly wage (low type)	9.230	9.156	9.230	9.156
Proportion of migrants (high type)	0.074	0.075	0.069	0.070
Proportion of migrants (low type)	0.043	0.047	0.038	0.045
Proportion of commuters (high type)	0.109	0.114	0.094	0.107
Proportion of commuters (low type)	0.082	0.143	0.071	0.131
Correlation between migrants and commuters	0.612	0.695	0.510	0.732
Correlation between migrants and distance	0.123	0.079	0.066	0.031
Correlation between commuters and distance	-0.079	-0.011	-0.155	-0.071
Correlation between migrants and rent cost	-0.101	-0.069	-0.063	-0.103
Correlation between migrants and rent cost	-0.099	-0.029	-0.098	0.011

4.4 Model fit

My model reproduces many features of the data (Table 5). It predicts a higher employment rate and higher average hourly wage for high-type workers compared with those for low-type workers. The fraction of migrants for both low-type workers and high-type workers are also well matched. Although the fraction of high-type commuters is almost perfectly predicted, the fraction of low-type commuters is over-predicted. My model also correctly predicts the correlation between labor mobility patterns and the geographic characteristics (rent prices and physical distance between paired counties). My model replicates the negative correlation between mobility rates and housing prices. I observe low numbers of migrants and commuters in counties with relatively high rental prices. On the other hand, both simulation and data find a positive correlation between migration and distance but a negative correlation between commuting and distance.

5 Estimation results

In this section, I present the parameter estimates and discuss their magnitudes. I then compare the elasticities of migrants and comments with respect to minimum wage changes predicted by the model with the elasticities estimated in the previous regressions. Finally, I quantify the downward bias in the estimation of disemployment effect when ignoring labor

Table 6: Parameter estimates

General parameters				
Parameters	Notation	Mean μ	S.D. σ	Corr. τ
Matching quality	θ	1.963	0.162	-
Unemployed flow utility	b	-23.8	0.123	0.949
Vacancy cost	ψ	428	211	0.196
High type productivity	a_h	3.106	-	-
Low type productivity	a_l	1.406	-	-
Commuting cost	β_0	48.5	1.217	0.458
Migration cost	β_1	78.4	9.709	
Coefficients in equation $cc_h(a,j)$				
	Commuting $(h=0)$		Migration $(h=1)$	
Additional cost for high type	β_{0a}	2.222	β_{1a}	-4.927
Coefficient for different local amenity	$\beta_{0\gamma}$	2.884	$\beta_{1\gamma}$	7.709
Coefficient for different distance	β_{0d}	0.697	β_{1d}	-2.051
Scale of preference shock (low type)	σ_l	15.0		
Scale of preference shock (high type)	σ_h	25.0		

mobility.

5.1 Understanding the model estimates

Table 6 provides model estimates for both the general parameters and the parameters in the moving equation (Equation 10). The estimated value of unemployment, $(b_j, b_{j'})$, is relatively homogeneous across counties. However, the vacancy cost ψ displays considerable heterogeneity across counties. Its mean value is 428, which is equivalent to \$68,480 if the filled worker is required to work 160 hours/month. Furthermore, the large standard error suggests substantial spatial diversification in vacancy costs. In addition, I find the productivity of high educated workers is on average significantly higher than the productivity of low educated workers ($a_h = 3.106$ vs. $a_l = 1.406$). When comparing mobility costs, migrating is more costly ($\beta_1 = 78.4$) than commuting ($\beta_0 = 48.5$), which explains why the fraction of commuters is on average larger than the fraction of migrants.

The lower panel in Table 6 reports the the determinants of choice-specific moving costs $cc_h(a, j)$. The positive sign of β_{0a} and negative sign of β_{1a} indicate that, compared to low educated workers, high educated workers are more likely to migrate when accepting the job offers from a neighboring county. These two coefficients rationalize the observation that when looking at commuting behavior, 40% of high educated workers are commuters whereas only 34% of low educated workers are commuters. The next two coefficients, $\beta_{0\gamma}$ and $\beta_{1\gamma}$

Table 7: Moving costs and neighboring county preference

	Lump-sum ex-ante moving cost (unit: \$)				Indifferent opportunity			
	Low educated		*High educated*		*Low educated*		*High educated*	
	County j	County j'	County j	County j'	County j	County j'	County j	County j'
10th	7,749	7,773	7,648	7,626	0.010	0.008	0.022	0.014
25th	8,400	8,286	8,602	8,459	0.047	0.036	0.092	0.067
Median	8,794	8,760	9,262	9,210	0.222	0.181	0.229	0.211
75th	9,194	9,109	9,818	9,741	0.880	0.823	0.561	0.495
90th	9,691	9,544	10,410	10,365	1.023	1.018	0.997	0.974
Mean	8,693	8,683	9,098	9,067	0.407	0.373	0.367	0.336
SD	1,046	931	1,330	1,242	0.405	0.402	0.351	0.333

Note: the dollar value of ex-ante moving cost $c(a, j)$ is estimated based on a representative full time worker whose working time is 160 hours/month.

link the moving cost with the local housing rental price, which is regarded as a proxy of local amenities . The positive values of $\beta_{0\gamma}$ and $\beta_{1\gamma}$ mean that high housing costs are associated with high moving costs. Workers are less likely to take neighboring jobs when the housing price in neighboring county is high. Even accepting the neighboring jobs, the mobile workers are less likely to choose migration as their preferred moving option. The coefficients β_{0d} and β_{1d} capture the correction between physical distance and moving cost. The positive sign of β_{0d} and negative sign of β_{1d} reflects the pattern that more mobile workers would choose migration over commuting as county pairs are farther apart. Lastly, the scale parameters for low-educated workers is smaller than that for high educated workers.

Table 7 reports the distributions of moving costs. The left panel displays summary statistics of the ex-ante moving costs $c(a, j)$ for workers differentiated in their types and locations. According to my estimates, the ex-ante moving cost is on average \$8,700 for low type workers and \$9,100 for high type workers. These costs are summary statistics of relocation costs, housing market transaction costs (for migrants only) and psychic costs.(Schwartz (1973), Greenwood (1975)) The estimated moving costs are lower than previous ones reported in the literature. For example, Kennan and Walker (2011) estimate a moving cost value of \$312,000 for an average movers across states in the US. And Schmutz and Sidibe (2016) find the average moving cost among French cities is around €15,000. The moving costs in my paper are lower for two reasons: first, I focus on the migration/commuting flows between two contiguous counties. The geographic proximity could greatly reduce the potential moving costs. Secondly. I focus on younger workers who are more likely to be affected by minimum wages. The opportunity costs of moving for those workers are relatively low.

Moreover, the moving cost can be equivalently measured using the openness of the local labor market. The right panel illustrates this idea and calculates the possibility that a random job from a neighboring county is preferred to a random job from the local county. The probability distribution for low type workers are less diversified than that for high type workers. The probability of preferring a neighboring county ranges from 0.010 (the 10th percentile) to 1.023 (the 90th percentile) for low educated workers in county j; this probability shrinks to a range of 0.020 (the 10th percentile) to 0.997 (the 90th percentile) for high educated workers. I also find this distribution is right skewed. In the median county, the probability for a low-skilled worker to receive a preferred job from neighboring county is 22.2%. This effect is comparable to the results of Manning and Petrongolo (2017). Using UK data, they find that the probability of a random job 5km distant being preferred to random local job is only 19%.

5.2 Out-of-sample validation: comparing model-based predictions with regression results

In this section, I use the model to predict the minimum wage elasticities of commuters and migrants and then compare the predicted elasticities to actual elasticities estimated by regression 13. This comparison is treated as an extra out-of-sample validation since the elasticities of commuters and migrants with respect to minimum wage are not used as targeted moments when estimating the baseline model.

Given county specific parameters and local minimum wage levels, the model allows me to calculate the fraction of migrants and commuters in each county . Specifically, the fraction of migrants in county j given minimum wage pair $MI(a, j; m_j, m_{j'})$ is expressed as the number of migrants from county j' to county j, divided by the sum of local hires in county j, and total mobile hires from county j', i.e.

$$MI(a, j; m_j, m_{j'}) = \frac{P_1(a, j')U(a, j')\tilde{G}(\max\{\theta^{**}(a, j'), \frac{m_j}{a}\})}{U(a, j)\tilde{G}(\max\{\theta^{*}(a, j'), \frac{m_j}{a}\}) + U(a, j')\tilde{G}(\max\{\theta^{**}(a, j'), \frac{m_j}{a}\})}$$

Meanwhile, the fraction of commuters in county j given the local minimum wage pair $CM(a, j; m_j, m_{j'})$ is given by the total number of commuters from county j, divided by the sum of local hires in county j' and all mobile workers (both commuters and migrants) from county j, i.e.

$$CM(a, j; m_j, m_{j'}) = \frac{P_0(a, j)U(a, j)\tilde{G}(\max\{\theta^{**}(a, j), \frac{m_{j'}}{a}\})}{U(a, j')\tilde{G}(\max\{\theta^{*}(a, j'), \frac{m_{j'}}{a}\}) + U(a, j)\tilde{G}(\max\{\theta^{**}(a, j), \frac{m_{j'}}{a}\})}$$

When the minimum wage in county j increases from m_j to $m_j + \Delta m_j$ but the minimum wage in county j' remains unchanged, I calculate new fractions of commuters $CM(a, j; m_j + \Delta m_j, m_{j'})$ and migrants $MI(a, j; m_j, m_{j'})$ in the new steady-state. The percentage changes in labor mobility are defined as:

$$\begin{aligned}
\Delta \log MI(a, j) &= \log(MI(a, j; m_j + \Delta m_j, m_{j'}) - \log MI(a, j; m_j, m_{j'}) \\
\Delta \log CM(a, j') &= \log(CM(a, j'; m_j + \Delta m_j, m_{j'}) - \log CM(a, j'; m_j, m_{j'})
\end{aligned}$$

Using data on minimum wage changes, I predict $\Delta \log MI(a, j)$ and $\Delta \log CM(a, j')$. Figure 2 displays the distributions of $\Delta \log MI(a, j)$ and $\Delta \log CM(a, j')$ for different types. First, all distributions show substantial heterogeneity across county-pairs, suggesting local markets are diversified. Minimum wage hikes decrease the chance of finding a job but increase the expected wages once hired. When the cost exceeds the benefit, the local labor market becomes less attractive, and workers either move away or stop moving in. The mean value of $\Delta \log CM(low, j)$ is positive (0.034) whereas the average value of $\Delta \log MI(low, j)$ is negative (-0.034), both of which indicate that low-skilled workers are more likely to leave areas with higher minimum wages in the majority of county pairs. Second, the distributions for low-skill workers are more dispersed than those for high-skilled workers. This is in line with the observation that low-skill workers are more responsive to minimum wage changes.

Next, I check the out-of-sample validation by comparing the model generated $\Delta \log MI(a, j)$ and $\Delta \log CM(a, j')$ with the data. In the sample, the percentage changes of migrants and commuters are directly calculated by comparing the fractions of mobile workers before minimum wage changes with those after minimum wage changes. Then I run the following regression to compute the minimum wage elasticity from model predictions ("Model-based elasticity β_1^*") and from data observations ("Data-based elasticity β_1") separately:

$$\begin{aligned}
\Delta \log MI(a, j) &= \beta_1^* \Delta logMW_j + \Delta\epsilon_0 \\
\Delta \log CM(a, j) &= \beta_1^* \Delta logMW_j + \Delta\epsilon_1
\end{aligned} \tag{11}$$

The regression results based on the data were previously calculated in Table 10 since regression 11 is a simplified version of Equation 13 that ignores the county fixed effect and restricts the observational period. Table 8 shows that the model-based β_1^* and the data-based β_1 are comparable. For low educated mobile workers, both estimates suggest that they exit counties with minimum wage hikes. And the magnitudes of both elasticities are very similar (within a 90% confidence interval). In addition, both estimates find the elasticities for low educated

Figure 2: The distribution of $\Delta \log MI(a, j)$ and $\Delta \log CM(a, j')$ after minimum wage hikes

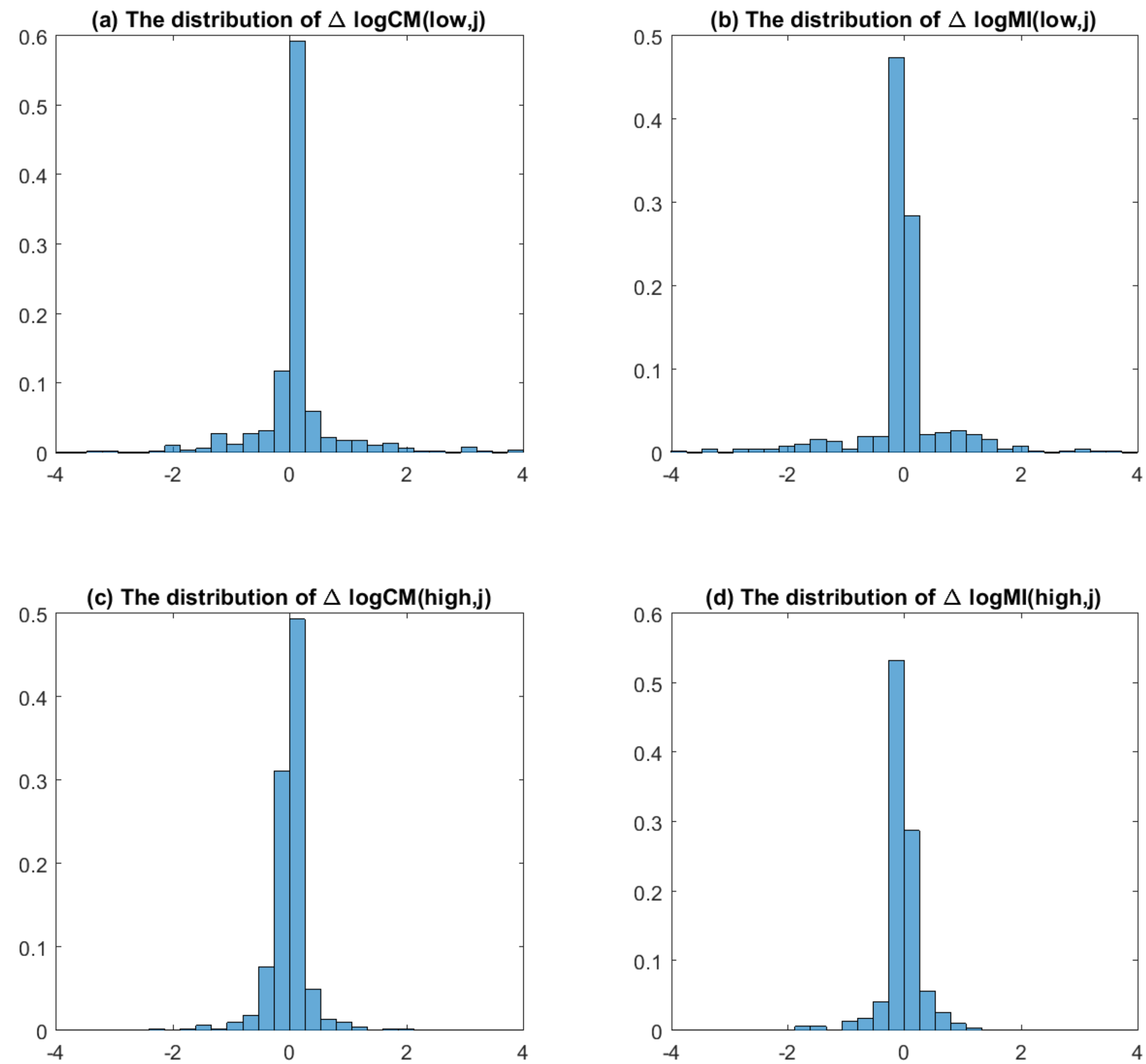


Table 8: The comparison between model predictions and regression results

	Model-based (β_1^*)	Data-based (β_1)
Low skilled Commuters	0.741***	0.458**
	(0.234)	(0.215)
Low skilled Migrants	-0.590**	-0.589***
	(0.260)	(0.160)
High skilled Commuters	-0.282***	0.263**
	(0.082)	(0.133)
High skilled Migrants	-0.081	-0.101
	(0.080)	(0.112)

Note: The regression column is directly from Table 10. Standard errors are displayed in parentheses.
** for 10%. ** for 5%, and *** for 1%.*

workers(absolute value) are larger than the elasticities for high skilled workers. This is consistent with the intuition that low educated workers are more responsive to the minimum wage adjustments.

The model-based elasticity for more highly educated commuters is less consistent with data-based elasticity. This discrepancy can be attributed to the distinction between short- and long- run effects. Althrough the data-based β_1 captures the immediate response after the minimum wage change, the model-based β_1^* demonstrates cumulative changes between two steady states.

This distinction between the short- and long-run effects is also emphasized in Sorkin (2015). He argues that the reduced-from effects are essentially uninformative about the true long-run elasticity. In my case, the key reason is the sorting of workers provides additional feedback effects in the long run. When a local county increases its minimum wage, the fraction of low type workers decreases in local county but increases in the neighboring county. As the average worker quality improves in the local market, firms have more incentive to post vacancies in local county rather than in neighboring county. However, this feedback effect is hard to be observed in the short run since the adjustment of local worker quality is slow. I will further explore this mechanism in Section 6.

5.3 Quantifying the underestimation of disemployment effects when ignoring labor mobility

Starting with Card and Krueger (1994), cross-border comparisons became a common method of studying the disemployment effects of the minimum wage. Dube et al. (2010) and Dube et al. (2016) generalize this strategy to all county pairs and find limited disemployment effects, which is consistent with Card and Krueger (1994). Although the cross-border design allows one to assume similarity between the treated area and control area, it may be problematic. As pointed out by Neumark et al. (2014b), "spillover effects can certainly contaminate the control observations. If workers displaced by the minimum wage find jobs on the other side of the border, employment will expand in the control areas". Based on my model, I quantitatively evaluate two sources of the underestimation of disemployment effects. First, unemployed workers who leave are "missing" from the treated county. Second, they may "reappear" in the neighboring county, contaminating the control group.

To evaluate the first channel, I compare the disemployment effect from two different minimum wage increases. In case 1, both counties increase their minimum wage by the same percentage $(m_j, m_{j'}) \to (m_j + \Delta m_j, m_{j'} + \Delta m_{j'})$. In case 2, only one county increases its minimum wage $(m_j, m_{j'}) \to (m_j + \Delta m_j, m_{j'})$. In case 1, the geographical minimum wage differences are more compressed since the minimum wage increases in both counties rather than increase only in one local county. Therefore, the opportunity to arbitrage relative minimum wage differences are largely eliminated in case 1 compared with case 2. The disemployment effect caused by minimum wage hikes is defined as the change of the log employment rate under the steady-state before minimum wage change and the new steady-state after the minimum wage change:

$$
\begin{aligned}
&\text{Case 1:} \quad \Delta \log Emp_j = \quad \log Emp_j(m_j + \Delta m_j, m_{j'} + \Delta m_{j'}) \quad - \log Emp_j(m_j, m_{j'}) \\
&\text{Case 2:} \quad \Delta \log Emp_j = \qquad \log Emp_j(m_j + \Delta m_j, m_{j'}) \qquad - \log Emp_j(m_j, m_{j'})
\end{aligned}
$$

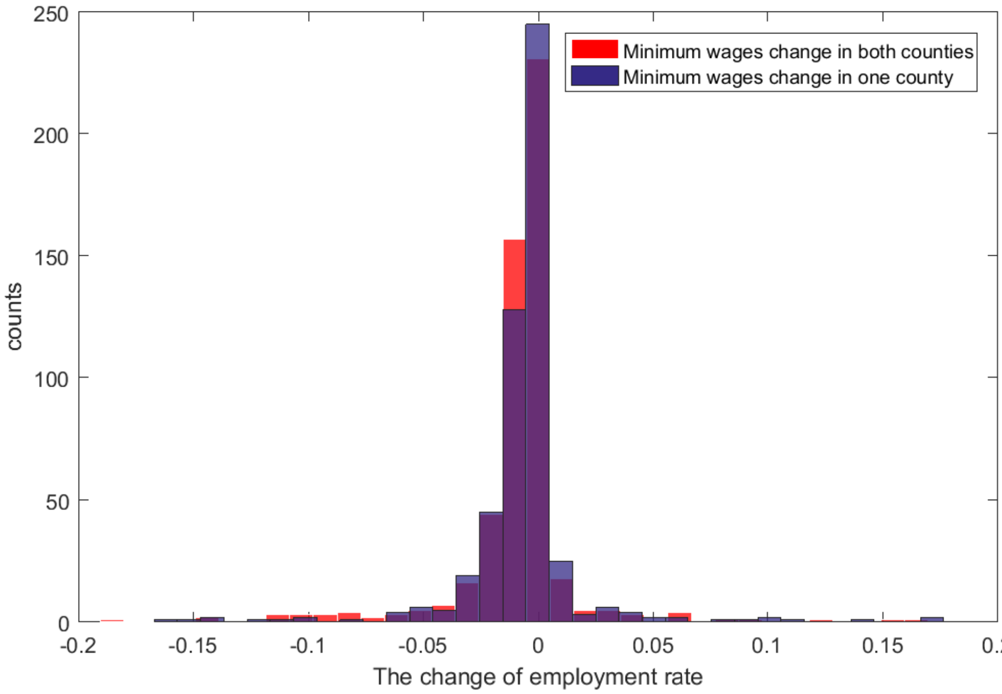
Figure 3 compares the distribution of $\Delta \log Emp_j$ under case 1 and case 2. The average value of $\Delta \log Emp_j$ in case 1 is more negative than that in case 2 while the distribution of $\Delta \log Emp_j$ in case 1 (red histogram) is more right-skewed than in case 2 (blue histogram). 86.9% of counties in case 1 experience negative employment changes due to minimum wage hikes compared to only 82.0% in case 2. This comparison confirms that the spillover effect actually attenuates the disemployment effect.

Next, I calculate the minimum wage elasticity of employment by estimating the following regressions:

$$
\begin{aligned}
\text{Case 1:} \quad \Delta \log Emp_j &= \beta_1 \Delta log MW_j + \Delta \epsilon_1 \\
\text{Case 2:} \quad \Delta \log Emp_j &= \beta_2 \Delta log MW_j + \Delta \epsilon_2
\end{aligned}
$$

To include the potential bias caused by the contamination of control group, I recalculate the disemployment elasticity in case 2 using the neighboring county as the control group. This

Figure 3: The disemployment effect under different minimum wage hikes



calculation mimics the diff-in-diff approach:

$$\text{Case 3:} \quad \Delta \log Emp_j - \Delta \log Emp_{j'} \quad = \quad \beta_3 \Delta log MW_j + \Delta \epsilon_3$$

Table 9 reports the minimum wage elasticity of employment in all three cases. "Case 1" reports the elasticity of employment when both counties increase their minimum wages by the same proportion. "Case 2" reports the elasticity of employment when only the local county increases its minimum wage. Finally, "Case 3" displays the alternative elasticity if the neighboring county is used as the control group. Workers in case 1 have less incentive to arbitrage the minimum wage difference between two counties compared with their incentive in case 2. Therefore, changes in labor mobility after minimum wage hikes in case 1 is smaller than changes in case 2. Consequently, I observe a larger disemployment effect in case 1 (-0.0733) compared with case 2 (-0.0421). Furthermore, when using the neighboring county as the control group, the disemployment effect continues to shrink from -0.0421 to -0.0341. This shares the same pattern with the different disemployment effect estimated in Table 11. In Table 11, the minimum wage elasticity of employment changed from -0.068 to -0.039 after controlling for pair-specific time trends instead of a common time trend. Dube et al. (2016) argue that this change is driven by spatial heterogeneity. My findings suggest that

Table 9: Elasticity of employment with respect to the minimum wage

	Case 1 (β_1)	Case 2 (β_2)	Case 3 (β_3)
Whole Sample	-0.0733***	-0.0421***	-0.0341***
	(0.0069)	(0.0075)	(0.0096)
Below bottom quartile	-0.0957***	-0.0445***	-0.0153
of moving cost	(0.0163)	(0.0136)	(0.0190)

*Note: "Case 1" reports the elasticity of employment when both county increase their minimum wages by the same proportion. "Case 2" reports the elasticity of employment when only the local county increase its minimum wage. "Case 3" displays an alternative elasticity if the neighboring county is wrongly picked as the control group. Standard errors are displayed in parentheses. * for 10%. ** for 5%, and *** for 1%.*

such changes are driven by labor mobility rather than by spatial heterogeneity. This result highlights the concern that neighboring counties, despite their geographic proximity, may not be the appropriate control group due to the contamination caused by labor mobility.

If labor mobility is causing underestimation of the disemployment effect, then the bias should be larger for counties with lower moving costs. To verify this conjecture, I conduct an additional placebo test for a sub-sample of counties whose moving costs are in the bottom quartile. My estimates, reported in the second row of Table 9, are in line with this conjecture. First, the difference of the elasticities between case 1 and case 2 becomes larger when using the restricted sample. The main reason is that the disemployment effect in case 1 is larger (-0.0957) compared with the previous effect (-0.0733) using the full sample. Second, using the neighboring county as the control group creates more severe downward bias. Although the elasticities in case 2 are robust to different sub-samples, it goes down sharply to -0.0153 in case 3 when using the neighboring county as the control group. To summarize, ignoring labor mobility and potential spillover effects cause the disemployment effect to be underestimated.

6 Policy experiments

In this section, I use the estimated model to examine the distributional impacts of local minimum wage hikes. There are (at least) two criteria to evaluate the welfare consequences of the minimum wage polices. The first natural welfare candidate is the value of unemployment $V_u(a, j)$, which can also be interpreted as the ex-ante welfare of heterogeneous workers with different types a and locations j. This is my primary measure because my goal is to understand the distributional effects for heterogeneous workers under minimum wage hikes. A second welfare criteria is defined for the local government, which is of particular interest when considering the total spillovers of local minimum wage policy to the neighboring county. Following Flinn (2006), I assume that the minimum wage is the only policy instrument available to the local government and the welfare function of local government defined as follows:

$$
\begin{aligned}
W_j(m_j) \quad = \quad & \textstyle\sum_{a\in\{a_l,a_h\}}[\underbrace{L(a,j)\bar{V}_e(\theta,a,j,\theta^{*}(a,j))}_{\text{(1) Local employed workers}} + \underbrace{MI(a,j)\left(\bar{V}_e(\theta,a,j,\theta^{**}(a,j')) - c(a,j')\right)}_{\text{(2) Migrants from neighbouring county}} \\
+ \quad & \underbrace{CM(a,j')\left(\bar{V}_e(\theta,a,j',\theta^{**}(a,j')) - c(a,j)\right)}_{\text{(3) Commuters to the neighbouring county}} + \underbrace{U(a,j)V_u(a,j)}_{\text{(4) Unemployed workers}} \\
+ \quad & \underbrace{E(a,j)\bar{V}_f(a,j)}_{\text{(5) Revernue from filled vacancies}} \quad] - \quad \underbrace{K_j\psi_j}_{\text{(6) Total cost of vacancies}}
\end{aligned}
$$

where (1) $L(a,j)$ is the population of local employed workers with $\bar{V}_e(\theta,a,j,\theta^{*}(a,j))$ denoting their average welfare. (2) $MI(a,j)$ is the population of migrants who move from county j', with $\bar{V}_e(\theta,a,j,\theta^{**}(a,j')) - c(a,j')$ as their average net welfare. (3) $CM(a,j')$ is the population of migrants who commute to work in county j', with $\bar{V}_e(\theta,a,j',\theta^{**}(a,j)) - c(a,j)$ as their average net welfare. (4) $U(a,j)$ is the population of local unemployed workers (all unemployed workers have same welfare level $V_u(a,j)$). On the demand side of the market, while there are K_j vacancies in county j, only $M_j = \sum_a E(a,j)$ are filled with workers and generate positive revenue. The free entry condition guarantees that the revenue generated from the filled vacancy is equal to the total cost of posted vacancies in the steady state. Thus the total contribution of terms (5) and (6) is equal to 0.

To understand the distributional effects of local minimum wage hikes, it is important to recognize the different forces at play. Assume county 1 changes its minimum wage while county 2 keeps its minimum wage unchanged. The direct effect in county 1 depends on the trade-off between the decrease in working opportunities (*"disemployment effect"*) and the increase in expected income (*"wage enhancement effect"*). Because the productivity distribution of high type workers first-order stochastically dominates that of low type workers , the working opportunity of the high type is less hurt by the same minimum wage increase compared with that of low type workers. As a result, low type workers have stronger incentives to move out of the country to avoid welfare losses caused by the minimum wage hike. Besides the direct effects, there is an additional general equilibrium effect through the change in a firm's incentive to post vacancies. First, the share of matching surplus decreases when firms are constrained by a higher minimum wage (*"share reduction effect"*). Secondly, due to the assumption of random search, firms are unable to screen workers' type when they post vacancies. Thus, vacancies (per capita) will be negatively correlated with the proportion of low type workers in their local county. Worker sorting decreases the composition of high type in county 2. As a result, the local workers in county 2 suffer additional welfare losses because of the decrease in hiring probability (*"composition changing effect"*). The additional force is the moving costs which generates welfare differences between the same type workers in different locations. Compared with local workers, mobile workers have to pay additional moving costs to work in the same job, *ceteris paribus*. This friction is traditionally referred to as the *lock-in effect*.

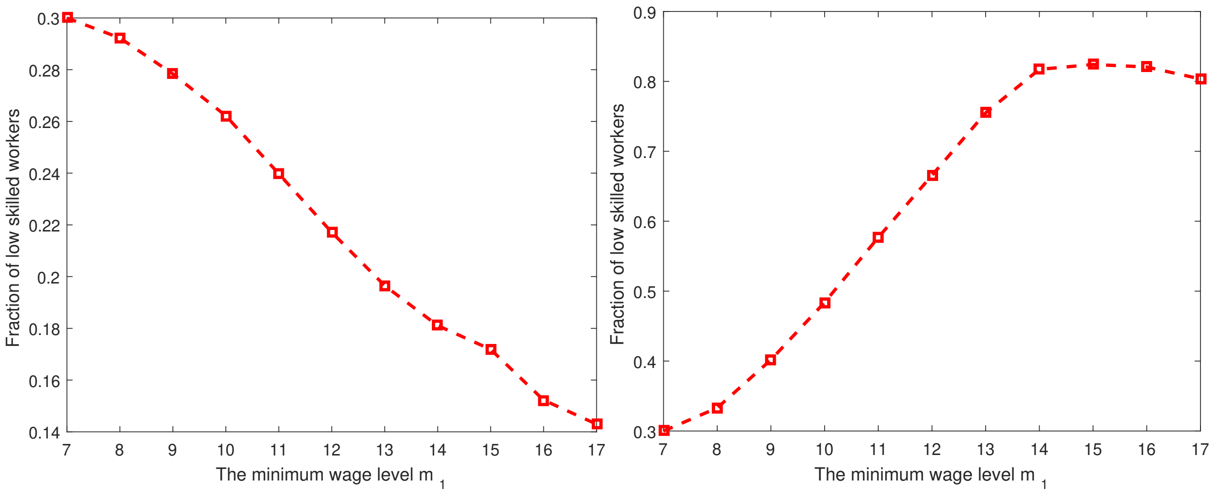
After understanding the distributional effects of heterogeneous workers, I conduct two counterfactual policies aiming for the reduction of spillover externalities. In the first policy experiment, I completely restrict labor mobility between counties by increasing the moving cost to infinity. This experiment captures the extreme case when no labor mobility is allowed. In the second policy experiment, the central government preempts the local minimum laws. In other word, the two paired counties follow universal minimum wage hikes rather than setting up their local minimum wages. In reality, the preemption of local minimum wage laws is a popular policy intervention for state legislation to avoid "patchwork" of wage levels within a state. So far, 27 states have passed such laws.

6.1 The distributional effect of local minimum wage hikes

In this section I explore how the welfare of workers (differentiated by their type a and location j) changes with respect to local minimum wage changes in county 1. To better exclude the effect of local minimum wage hikes from other disturbances such as geographic asymmetry, I consider symmetric county pairs where the geographic parameters in both counties take the mean values of the distributional estimates. The distributional effects depend critically on the magnitude of local minimum wage increases. I assume the initial hourly minimum wage in both counties is \$7 and consider welfare changes when increasing the minimum wage in county 1 to an amount between \$7 and \$17. Most of my results are presented in graphical form. I will first report the change in local economic conditions (e.g. contact rates, the composition of heterogeneous workers). Then, I will compute welfare changes with respect to changes in minimum wages for different workers. Lastly, I show welfare changes of local governments with changes in minimum wages.

Figure 4 display changes in worker composition in both counties under different minimum wage increases. As the local minimum wage in county 1 increases from \$7 to \$17, the fraction of low type worker in county 1 monotonically decreases to 0.15 while the fraction of low type workers in county 2 has a hump shape with a peak of 0.8 when m_1 = \$14. These two patterns suggest that local minimum wage policy serves as a worker selection device. By

Figure 4: Worker composition under different minimum wages

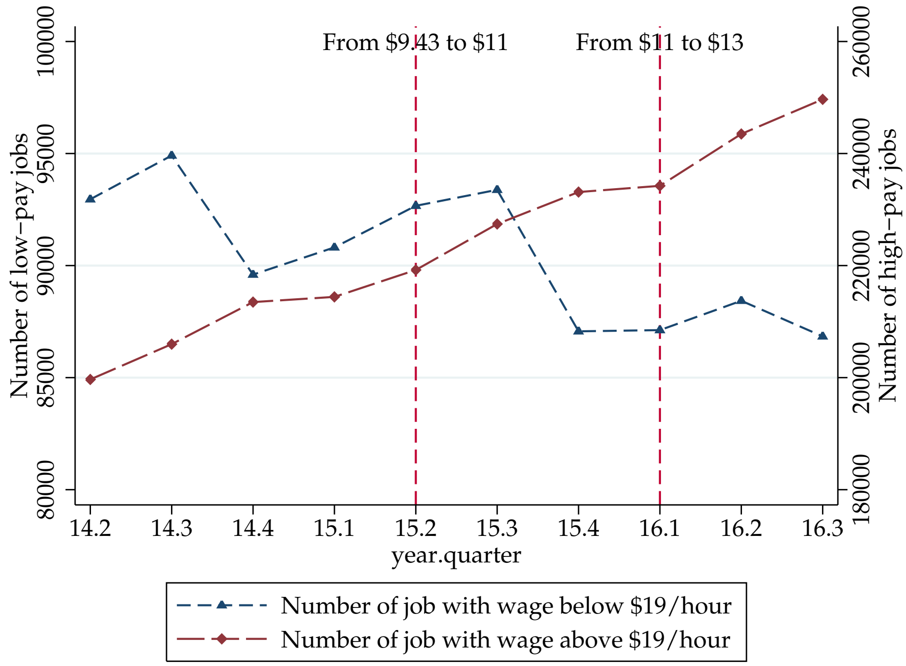


setting a higher minimum wage, the local government extracts high type workers from the neighboring county while also dumping low type workers on the neighboring county. The prediction in my model is consistent with the real changes happening in Seattle after its city-level minimum wage increases from \$9.43/hour to \$13/hour.(Jardim et al. (2017)) As showed in figure 5, the number of low-pay job ($wage < 19/hour$) decreases while the number of high-pay job ($wage >$ \$$19/hour$) increases.

Next I consider the changes of firms' incentive to post vacancies. Figure 6 displays changes of contact rates in both counties and suggests two channels of changing the profit of posted vacancies. First, for the same match, firms get less value per vacancy when the minimum wage is higher. A higher minimum wage decreases both the probability that a given match is acceptable and makes the sustainable match less profitable. This channel explains why contact rates in both counties experience a downward change when minimum wage in county 1 increases. Second, the sorting of workers increases the concentration of high types in county 1 but decreases their concentration in county 2 because firms tend to post relatively fewer vacancies in the county with higher fraction of low type workers. The second channel explains why the contact rate in county 2 is systematically lower than that in county 1. Furthermore, the fraction of low type in county 2 reaches its peak at \$14 and starts to decrease after that, which explains the rebound of the contact rate in county 2 when $m_1 \geq \$15$.

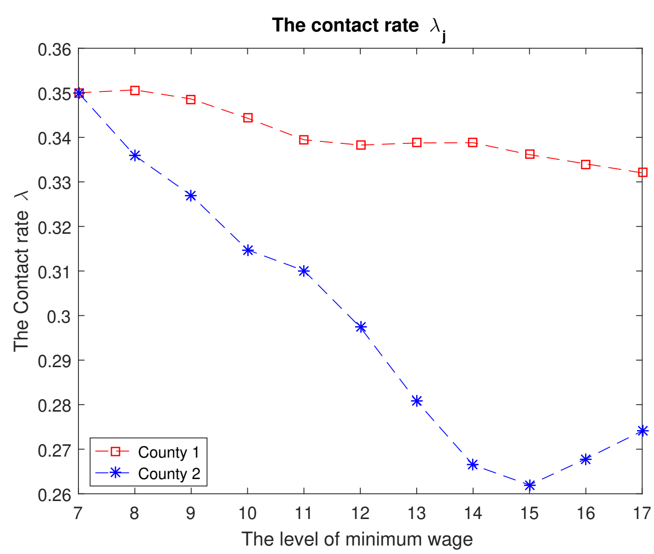
The most crucial results are the distributional effects of local minimum wage policies on heterogeneous workers. This heterogeneity is not well explored in the previous literature because workers are often considered to be ex-ante identical (e.g. Flinn (2006)). Let $V_u(a, j; m_1, m_2)$ be the ex-ante welfare for a worker with type a and in location j when m_1

Figure 5: Changes in Seattle jobs after increasing minimum wage from \$9.43 to \$13



Data source: administrative employment records from the Washington Employment Security Department, reported in Jardim et al. (2017), table 3.

Figure 6: Contact rates under different minimum wages



is set at m and m_2 is set at \$7, then the change of welfare is defined as

$$\Delta W_0(a, j; m, 7) = V_u(a, j; m, 7) - V_u(a, j; 7, 7)$$

Figure 7 shows the results. The top left panel displays welfare changes of low type workers in both county 1 (the blue line) and county 2 (the red line). The low type is severely harmed by higher minimum wages. As noted previously, this is driven by a combination of two effects. First, the higher m_1 rules out previously acceptable wages. Second, the higher minimum wage policy in county 1 pushes low type workers to county 2, diminishing their probability to be hired. The top right panel displays the welfare changes of high type workers in both county 1 (the blue line) and county 2 (the red line). The hump shape in high type welfare shows the existence of countervailing effects. Although raising the minimum wage increases workers' welfare by increasing the return of a match, previously acceptable matches become unacceptable. The latter effect dominates the previous effect when local minimum wage in county 1 exceeds \$14.

The lower panel of Figure 7 reports the change of inequality between high type and low type as minimum wage increase in county 1. Because the welfare of the low type is a convex curve whereas the welfare of the high type is a concave curve, the inequality curve expands and then reaches its peak when $m_1 = 15$. This result reveals that local minimum wage policy could actually increase inequality between high and low type workers, completely opposite of the intended policy effect.

Lastly, the welfare difference between same type workers in two counties indicates the "lock-in" effects due to the existence of moving costs, I will continue to explore this effect in the next section.

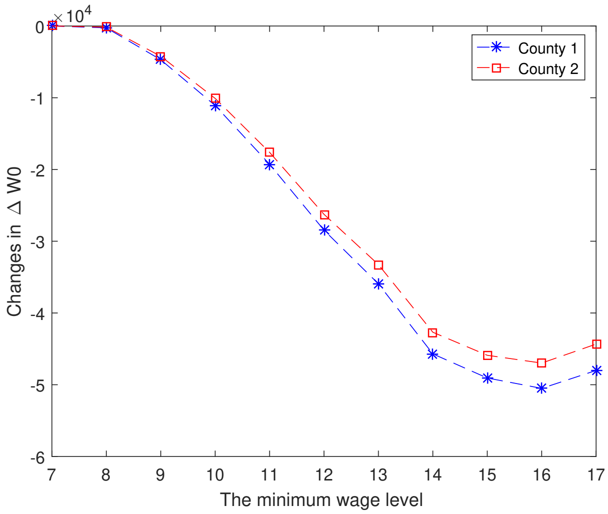
Figure 8 plots the change of total welfare in each county with respect to a change in the local minimum wage. The total welfare in county 1 has a single peak at $m_1 = 8$, while the total welfare in county 2 declines until $m_1 = 16$. An increase in m_1 almost always harms the total welfare in county 2. Put another way, the increases in local minimum wages generate negative externalities to neighboring counties.

6.2 Restricting labor mobility between counties

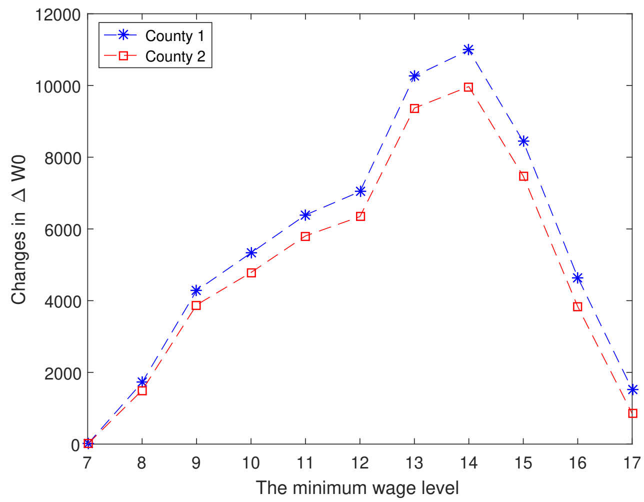
In this session, I want to reduce spillover externalities by completely blocking labor mobility between counties. This can be treated as an extreme way of implementing mandatory local hiring requirements. For example, the public infrastructure projects in San Francisco

Figure 7: Welfare changes across heterogeneous workers under different minimum wage increases

(a) Welfare changes - low type



(b) Welfare changes - high type



(c) Welfare inequality between high type and low type

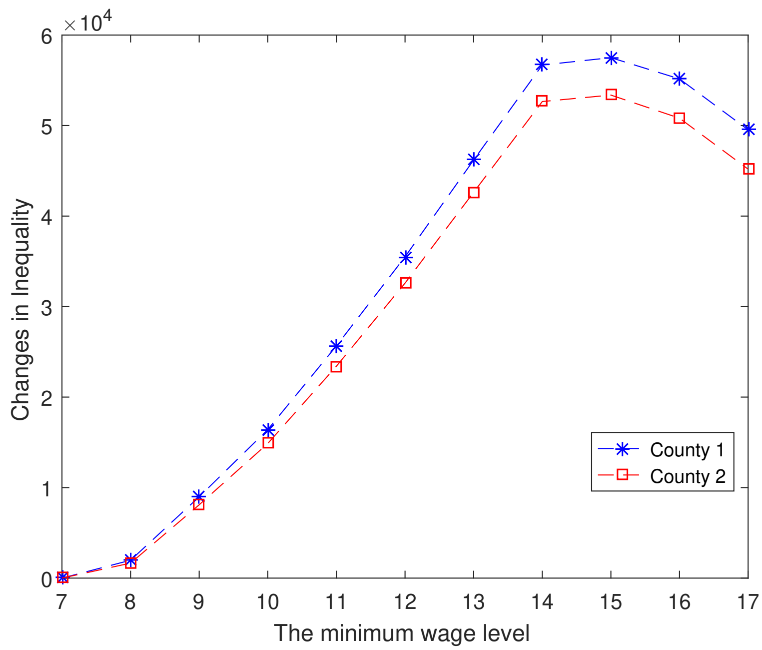
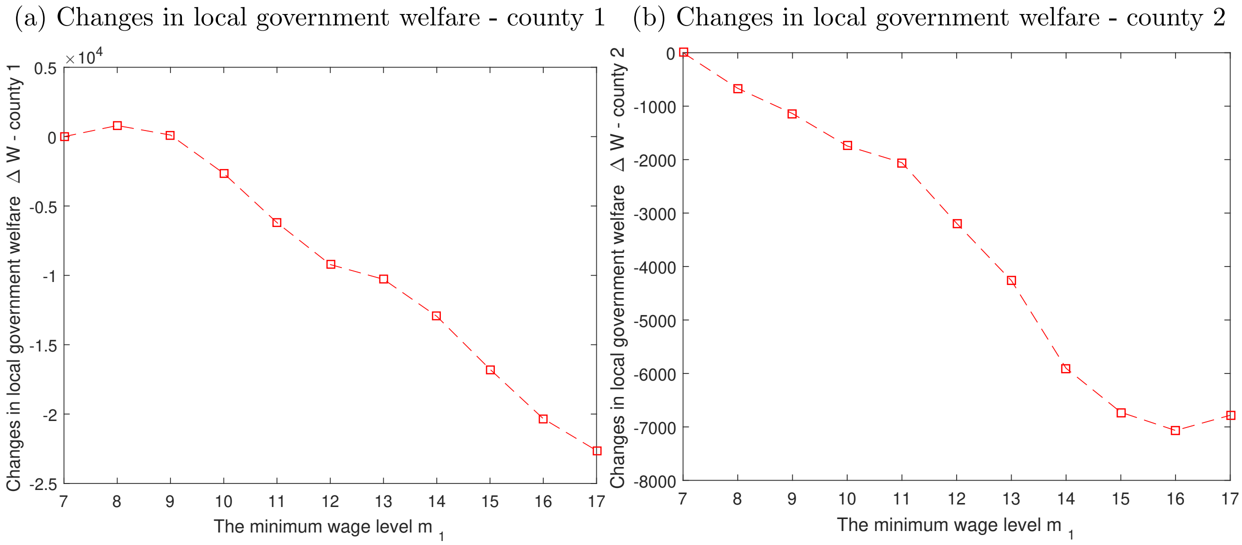


Figure 8: Changes in local government welfare as minimum wage changes in county 1



require that at least 50% of their job hours to go to San Francisco residents. I achieve this moving barrier in the model by setting the moving cost to be infinite ($c(a, j) = +\infty$) so that the two labor markets are totally disconnected, which is referred as "Autarky case".

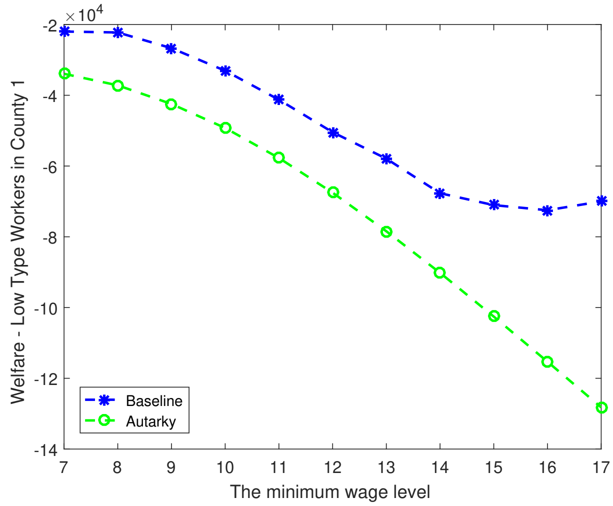
Figure 9 compares the ex-ante welfare across different types of workers in the "Baseline" and "Autarky" cases. In the "Autarky" case, a minimum wage increase in county 1 has no effect on the workers in county 2, because these two labor markets are totally segregated. Therefore, the welfare of worker in county 2 (green line) is a horizontal line in the "Autarky" case. The welfare in the "Baseline" case and in the "Autarky" case differ because of two effects. First, workers in the "Baseline" case have additional working opportunities from the neighboring county, which generate welfare gains for all types of workers. Secondly, the sorting of workers discourages firms from posting vacancies in county 2. This reduction of contact rates in county 2 has a negative effect on all workers, but particularly on lower type workers, because they are more concentrated in county 2 when m_1 increases. Taken together, welfare increases for everyone except for the low skill worker in county 2. When $m_1 > 10$, they would prefer to stay in the "Autarky" case to avoid the negative spillover effects.

6.3 Preempting local minimum wage laws (universal (federal) minimum wages vs. local minimum wages)

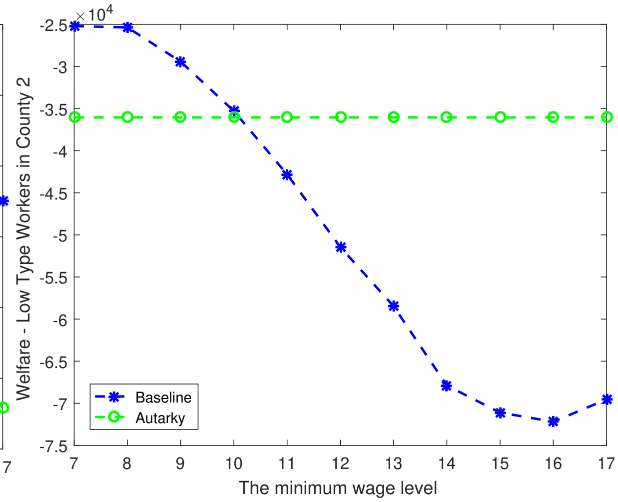
In this section, I perform a second counterfactual experiment: preempting local minimum wage laws. In reality, the preemption of local minimum wage laws is a popular

Figure 9: Change in worker welfare both in "Baseline" case and "Autarky" case

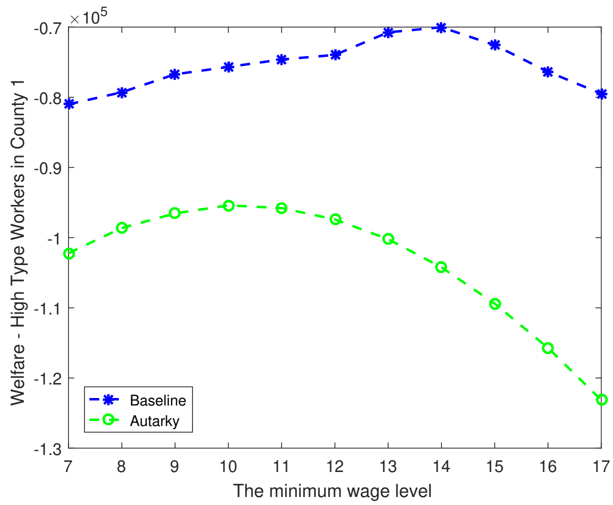
(a) Welfare - Low type in county 1



(b) Welfare - Low type in county 2



(c) Welfare - High type in county 1



(d) Welfare - High type in county 2

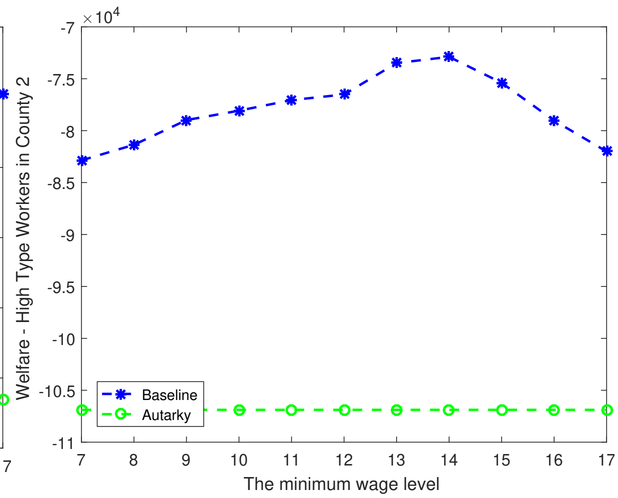
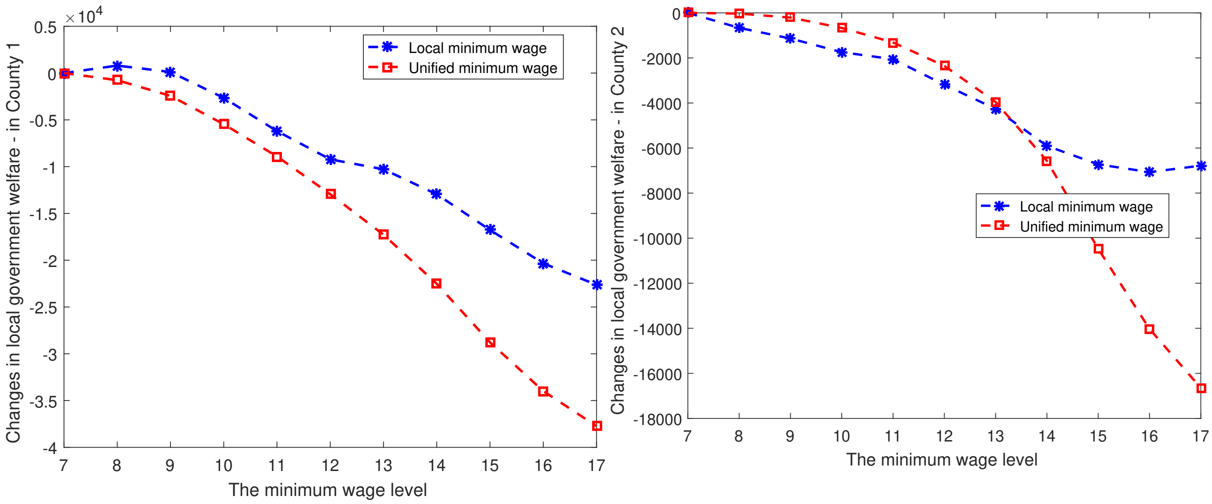


Figure 10: Changes in total welfare under local and under universal (federal) minimum wages

(a) Changes in local government welfare - county 1 (b) Changes in local government welfare - county 2



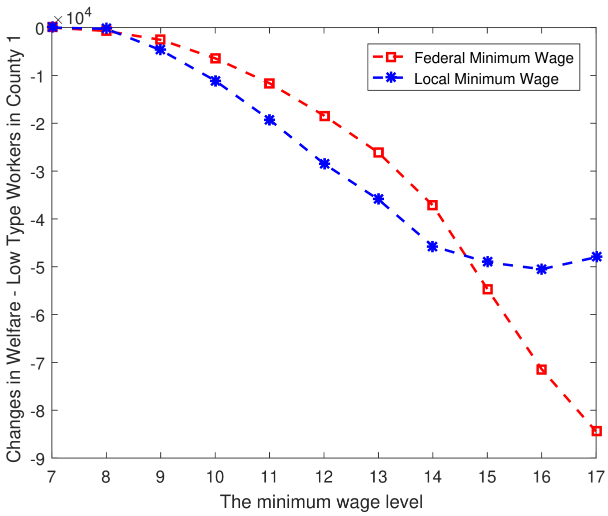
policy intervention adopted by either federal or state legislatures to avoid a "patchwork" of minimum wage levels within their justification. Sometimes, progressive legislatures offer a statewide/nationwide raise to avoid more aggressive local level minimum wage changes. To understand the trade-off between universal level minimum wage hikes and local level minimum wage hikes, I consider the case in which both counties have an identical increase of their same minimum wages. Thus welfare changes of heterogeneous workers when setting a universal federal minimum wage at m is defined as

$$\Delta W_0(a, j; m, m) = V_u(a, j; m, m) - V_u(a, j; 7, 7)$$

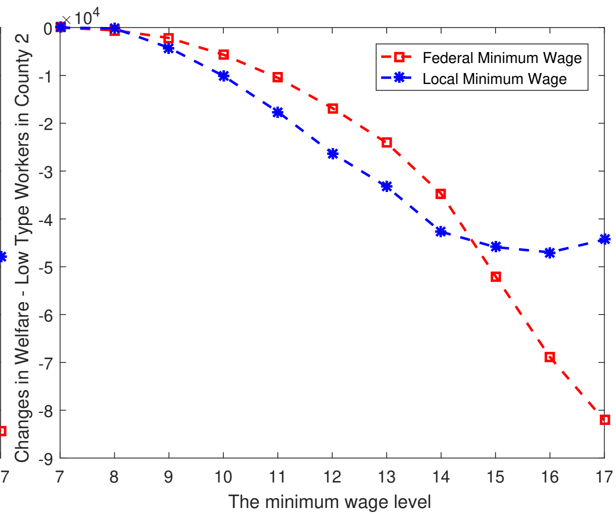
Figure 10 compares welfare changes under local minimum wage regulation and welfare changes under universal minimum wage regulation. Rather than keeping m_2 unchanged, a universal minimum wage policy equalizes the minimum wages in both counties, $m_1 = m_2$. Compared with the "Baseline" case, the increase of m_2 generates two offsetting effects. On one hand, the minimum wage hikes in county 2 dissolves previously acceptable matches. On the other hand, the increase of m_2 prevents the sorting of workers between two counties, encouraging firms to post more vacancies. As shown in the right panel of Figure 10, the benefit of preventing negative spillovers dominates the cost of losing acceptable matches when $m < 13.5$. When minimum wage is not dramatically high, the total welfare in county 2 is actually higher under universal minimum wage policy. When the minimum wage exceeds \$13.5, the total welfare in both counties is reduced, because the loss of sustainable matches

Figure 11: Change in worker welfare under local and universal (federal) minimum wage changes

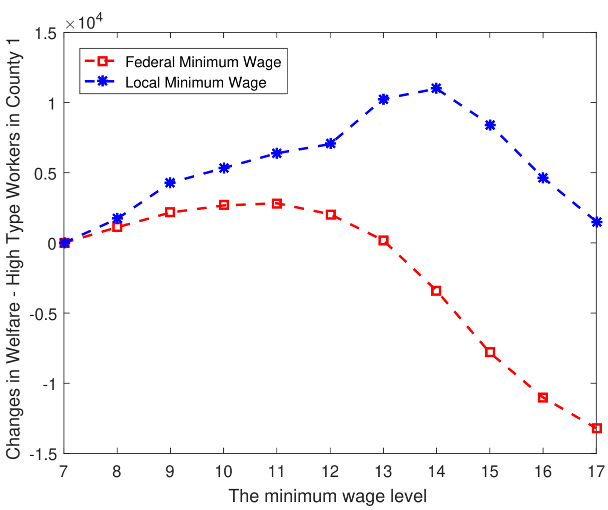
(a) Changes in welfare - low type in county 1



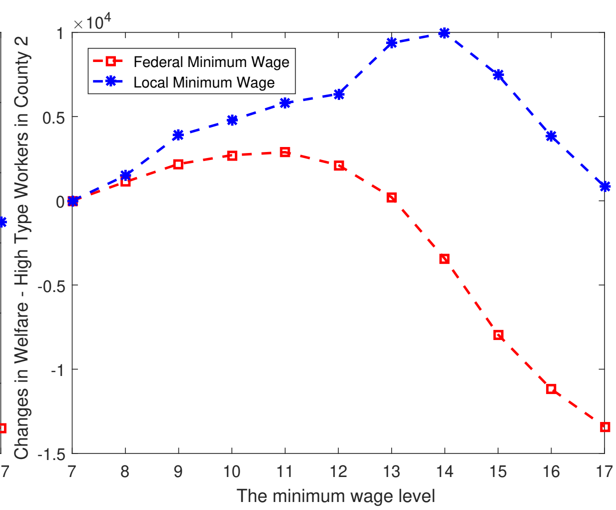
(b) Changes in welfare - low type in county 2



(c) Changes in welfare - high type in county 1



(d) Changes in welfare - high type in county 2



becomes the dominant effect.

When decomposing total local welfare by worker types, I find preferred minimum wage regulation (universal vs. local) in county 2 is driven by low type workers. Thus, a planner that cares for low type workers should opt for universal rather than local minimum wage intervention when the change is moderate ($m <$ \$14.5). However, this welfare gain is accompanied with a welfare loss for high type workers.

7 Conclusions

In this paper, I developed a spatial search model to study the effect of both local and universal (federal) minimum wage policies. In the model, firms endogenously choose where to post vacancies. Workers, differentiated by their type and location, engage in random search and can either accept a local job or migrate/commute to work in the neighboring county. My model captures three important effects associated with the minimum wage increases. First, conditional on being employed, a higher minimum wage shifts profits from firms to workers and increases workers' earnings. Second, a higher minimum wage also creates a disemployment effect by dissolving previously acceptable matches. This disemployment effect is more for low type worker. Third, firms reduce their vacancy postings in response to changing county-level worker composition and because they receive a smaller share of the matching surplus. Although the reduction in contact rates affects both counties, it has a larger effect on the neighboring county.

My analysis yields a number of interesting empirical findings when simulating the effects of minimum wage increases in county 1 with no change in county 2. First, minimum wage increases up to \$14/hour increase the welfare of high type workers but lower the welfare of low type workers, leading to an increase in inequality. Minimum wage increases in excess of \$14/hour lower the welfare of all workers, because the wage increases do not compensate for the disemployment effects. Second, the welfare of same type workers differs by locations ("lock-in effect") due to migration/commuting costs. Lastly, I find the disemployment effect of a minimum wage increase is underestimated if one ignores labor mobility. With the model, I obtain with the model a minimum wage elasticity of employment equal to -0.073; ignoring labor mobility cuts this value in half to -0.034. The bias is most severe for the counties with higher fractions of mobile workers.

I examine two counterfactual policies aiming for reducing spillover externalities: restricting labor mobility and preempting local minimum wage laws. In the experiment restricting labor mobility, the low type workers in neighboring county (the county without minimum wage change) prefer "Autarky" labor markets when the increase of local minimum wage is large ($m > \$10$). In the experiment of preempting local minimum wage laws, low type workers prefer a universal (federal) minimum wage rather than local minimum wages when the increa se of minimum wage is moderate ($m < \$14.5$). In contrast, the welfare of high type reduces unambiguously under both policies.

There are several ways to extend my analysis for future research. First, my model only compares the change between two steady states with minimum wage hikes. Adding transitional dynamics could capture the immediate effect of minimum wage hikes, which might differ from the long-term steady-state. Second, although I emphasize the worker selection and reallocation consequences of the local minimum wage policy, the local government is not a strategic player in my current model. Examining the competitive behavior of policy makers could be interesting. Third, local minimum wages also affect labor force participation. With higher minimum wages, individuals who were out of the labor force may also start to look for jobs in the labor market. This feature could be added into the model where government not only cares about the working population, but also the sub-population out of the labor force.

References

Altonji, Joseph G and Lewis M Segal, "Small-sample bias in GMM estimation of covariance structures," *Journal of Business & Economic Statistics*, 1996, *14* (3), 353–366.

Baum-Snow, Nathaniel and Ronni Pavan, "Understanding the city size wage gap," *The Review of Economic Studies*, 2012, *79* (1), 88–127.

Boca, Daniela Del, Christopher Flinn, and Matthew Wiswall, "Household choices and child development," *The Review of Economic Studies*, 2014, *81* (1), 137–185.

Boscoe, Francis P, Kevin A Henry, and Michael S Zdeb, "A nationwide comparison of driving distance versus straight-line distance to hospitals," *The Professional Geographer*, 2012, *64* (2), 188–196.

Burkhauser, Richard V, Kenneth A Couch, and David C Wittenburg, "Who minimum wage increases bite: An analysis using monthly data from the SIPP and the CPS," *Southern Economic Journal*, 2000, pp. 16–40.

Busso, Matias, Jesse Gregory, and Patrick Kline, "Assessing the incidence and efficiency of a prominent place based policy," *The American Economic Review*, 2013, *103* (2), 897–947.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin, "Wage bargaining with on-the-job search: Theory and evidence," *Econometrica*, 2006, *74* (2), 323–364.

Card, David and Alan B Krueger, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *The American Economic Review*, 1994, *84* (4), 772–793.

_ **and** _ , "Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania: reply," *The American Economic Review*, 2000, *90* (5), 1397–1420.

Coen-Pirani, Daniele, "Understanding gross worker flows across US states," *Journal of Monetary Economics*, 2010, *57* (7), 769–784.

Cohen, Roger, Andrew Lai, and Charles Steindel, "The Effects of Marginal Tax Rates on Interstate Migration in the US," *New Jersey Department of the Treasury*, 2011, p. 10.

David, H, Alan Manning, and Christopher L Smith, "The contribution of the minimum wage to US wage inequality over three decades: a reassessment," *American Economic Journal: Applied Economics*, 2016, *8* (1), 58–99.

Deere, Donald, Kevin M Murphy, and Finis Welch, "Employment and the 1990-1991 minimum-wage hike," *The American Economic Review*, 1995, *85* (2), 232–237.

den Berg, Gerard J Van and Geert Ridder, "An empirical equilibrium search model of the labor market," *Econometrica*, 1998, pp. 1183–1221.

Dube, Arindrajit, Suresh Naidu, and Michael Reich, "The economic effects of a citywide minimum wage," *ILR Review*, 2007, *60* (4), 522–543.

_ , T William Lester, and Michael Reich, "Minimum wage effects across state borders: Estimates using contiguous counties," *The Review of Economics and Statistics*, 2010, *92* (4), 945–964.

_ , _ , and _ , "Minimum wage shocks, employment flows, and labor market frictions," *Journal of Labor Economics*, 2016, *34* (3), 663–704.

Eckstein, Zvi and Kenneth I Wolpin, "Estimating a market equilibrium search model from panel data on individuals," *Econometrica: Journal of the Econometric Society*, 1990, pp. 783–808.

_ , Suqin Ge, and Barbara Petrongolo, "Job and wage mobility with minimum wages and imperfect compliance," *Journal of Applied Econometrics*, 2011, *26* (4), 580–612.

Enrico, Moretti, "Local labor markets," *Handbook of Labor Economics*, 2011, *4*, 1237–1313.

Flinn, Christopher, Ahu Gemici, and Steven Laufer, "Search, matching and training," *Review of Economic Dynamics*, 2017, *25*, 260–297.

_ and James Heckman, "New methods for analyzing structural models of labor force dynamics," *Journal of Econometrics*, 1982, *18* (1), 115–168.

_ and Joseph Mullins, "Labor market search and schooling investment," *International Economic Review*, 2015, *56* (2), 359–398.

Flinn, Christopher J, "Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates," *Econometrica*, 2006, *74* (4), 1013–1062.

Glaeser, Edward L, Joshua D Gottlieb et al., "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, 2008, *39* (1 (Spring)), 155–253.

Greenwood, Michael J, "Research on internal migration in the United States: a survey," *Journal of Economic Literature*, 1975, pp. 397–433.

Horton, John J, "Price Floors and Employer Preferences: Evidence from a Minimum Wage Experiment," *Unpublished paper*, 2017.

Jardim, Ekaterina, Mark C Long, Robert Plotnick, Emma van Inwegen, Jacob Vigdor, and Hilary Wething, "Minimum Wage Increases, Wages, and Low-Wage Employment: Evidence from Seattle," Technical Report, National Bureau of Economic Research 2017.

Kennan, John and James R Walker, "The effect of expected income on individual migration decisions," *Econometrica*, 2011, *79* (1), 211–251.

Kline, Patrick, "Place based policies, heterogeneity, and agglomeration," *The American Economic Review*, 2010, *100* (2), 383–387.

_ and Enrico Moretti, "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 2013, *129* (1), 275–331.

Mabli, James and Christopher Flinn, "On-the-Job Search, Minimum Wages, and Labor Market Outcomes in an Equilibrium Bargaining Framework," in "2007 Meeting Papers" number 791 Society for Economic Dynamics 2007.

Manning, Alan and Barbara Petrongolo, "How Local Are Labor Markets? Evidence from a Spatial Job Search Model," *The American Economic Review*, 2017, *forcecoming.*

McKinnish, Terra, "Cross-state differences in the minimum wage and out-of-state commuting by low-wage workers," *Regional Science and Urban Economics*, 2017, *64*, 137–147.

Meghir, Costas, Renata Narita, and Jean-Marc Robin, "Wages and informality in developing countries," *The American Economic Review*, 2015, *105* (4), 1509–1546.

Molloy, Raven, Christopher L Smith, and Abigail Wozniak, "Internal migration in the United States," *The Journal of Economic Perspectives*, 2011, *25* (3), 173–196.

Monras, Joan, "Minimum Wages and Spatial Equilibrium: Theory and Evidence," *IZA Discussion Paper No. 9460.*, 2015.

Moretti, Enrico and Daniel Wilson, "The effect of state taxes on the geographical location of top earners: evidence from star scientists," *The American Economic Review*, 2017, *107* (7), 1858–1903.

Mortensen, Dale T and Christopher A Pissarides, "Job creation and job destruction in the theory of unemployment," *The Review of Economic Studies*, 1994, *61* (3), 397–415.

Neumark, David, "The employment effects of minimum wages: Evidence from a prespecified research design the employment effects of minimumwages," *Industrial Relations: A Journal of Economy and Society*, 2001, *40* (1), 121–144.

_ , JM Ian Salas, and William Wascher, "More on recent evidence on the effects of minimum wages in the United States," *IZA Journal of Labor Policy*, 2014, *3* (1), 24.

_ , _ , and _ , "Revisiting the Minimum Wage - Employment Debate: Throwing Out the Baby with the Bathwater?," *ILR Review*, 2014, *67* (3_suppl), 608–648.

Petrongolo, Barbara and Christopher A Pissarides, "Looking into the black box: A survey of the matching function," *Journal of Economic Literature*, 2001, *39* (2), 390–431.

Phibbs, Ciaran S and Harold S Luft, "Correlation of travel time on roads versus straight line distance," *Medical Care Research and Review*, 1995, *52* (4), 532–542.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L Violante, "Mismatch unemployment," *The American Economic Review*, 2014, *104* (11), 3529–3564.

Schmutz, Benoît and Modibo Sidibe, "Frictional Labor Mobility," 2016.

Schwartz, Aba, "Interpreting the effect of distance on migration," *Journal of Political Economy*, 1973, *81* (5), 1153–1169.

Serrato, Juan Carlos Suárez and Owen Zidar, "Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms," *The American Economic Review*, 2016, *106* (9), 2582–2624.

Sorkin, Isaac, "Are there long-run effects of the minimum wage?," *Review of Economic Dynamics*, 2015, *18* (2), 306–333.

Young, Cristobal and Charles Varner, "Millionaire migration and state taxation of top incomes: Evidence from a natural experiment," *National Tax Journal*, 2011, *64* (2), 255.

A Expression appendix

A.1 Deducing the expressions of $v_{uj}(a,j)$, $v_{uj'}(a,j)$ and $V_e(w,a,j)$

I now consider individual's search problem

$$
\begin{aligned}
v_{uj}(a,j) &= (1+\rho\epsilon)^{-1}[\quad ab_j\epsilon + \lambda_j\epsilon \underbrace{\int_{m_j}^{\infty} \max\{V_e(w,j), V_u(a,j)\} dF(w|a,\theta,j)}_{\text{A local offer arrives}} \\
&+ \quad \lambda_{j'}\epsilon \underbrace{\int_{m_{j'}}^{\infty} \max\{V_e(w,j') - c(a,j), V_u(a,j)\} dF(w|a,\theta,j'))}_{\text{A neighbouring offer arrives}} \\
&+ \quad +(1-\lambda_j\epsilon - \lambda_{j'}\epsilon)V_u(a,j) + o(\epsilon)]
\end{aligned}
$$

Multiplying $1+\rho\epsilon$ then subtracting $V_u(a,j)$ from both sides, I get

$$
\begin{aligned}
\rho\epsilon V_u(a,j) &= \quad ab_j\epsilon + \lambda_j\epsilon \underbrace{\int_{m_j}^{\infty} \max\{V_e(w,j), V_u(a,j)\} dF(w|a,\theta,j)}_{\text{A local offer arrives}} \\
&+ \quad \lambda_{j'}\epsilon \underbrace{\int_{m_{j'}}^{\infty} \max\{V_e(w,j') - c(a,j), V_u(a,j)\} dF(w|a,\theta,j'))}_{\text{A neighbouring offer arrives}} \\
&+ \quad -(\lambda_j\epsilon + \lambda_{j'}\epsilon)V_u(a,j) + o(\epsilon)
\end{aligned}
$$

Dividing both sides by ϵ and taking limits $\epsilon \to 0$, I arrive at

$$
\begin{aligned}
\rho V_u(a,j) &= \quad ab_j + \lambda_j \underbrace{\int_{m_j}^{\infty} \{V_e(w,j) - V_u(a,j)\}^+ dF(w|a,\theta,j)}_{\text{A local offer arrives}} \\
&+ \quad \lambda_{j'} \underbrace{\int_{m_{j'}}^{\infty} \{V_e(w,j') - c(a,j) - V_u(a,j)\}^+ dF(w|a,\theta,j'))}_{\text{A neighbouring offer arrives}}
\end{aligned}
$$

The value of employment with wage w is

$$V_e(w,a,j) = (1+\rho\epsilon)^{-1}\{w\epsilon + \eta_j\epsilon V_u(a,j) + (1-\eta_j\epsilon)V_e(w,a,j) + o(\epsilon)\}$$

Multiplying $1+\rho\epsilon$ then subtracting $V_e(a,j)$ from both sides, I get

$$\rho\epsilon V_e(w,a,j) = w\epsilon + \eta_j\epsilon V_u(a,j) - \eta_j\epsilon V_e(w,a,j) + o(\epsilon)$$

Dividing both sides by ϵ and taking limits $\epsilon \to 0$, I arrive at

$$V_e(w,a,j) = \frac{w + \eta_j V_u(a,j)}{\rho + \eta_j}$$

A.2 Solving for the bargained wage equation without the minimum wage constraint

Follow the same deduction procedure, the firm's value for a match with wage w, $V_t^f(w,a,\theta,j)$, is(I assume that the effective discount fact $\rho + \eta_j$ is the same as worker's):

$$V_f(w,a,\theta,j) = \frac{a\theta - w}{\rho + \eta_j}$$

Then the Nash bargaining $\hat{w}(\theta,a,j)$ without considering possible binding minimum wage is:

$$\begin{aligned}
\hat{w}(a,j,\theta) &= \arg\max\nolimits_w (V_e(w,a,j) - V_u(a,j))^{1-\alpha_j} V_f(w,a,\theta,j)^{1-\alpha_j} \\
&= \arg\max\nolimits_w (\tfrac{w+\eta_j V_u(a,j)}{\rho+\eta_j} - V_u(a,j))^{1-\alpha_j} (\tfrac{a\theta - w}{\rho+\eta_j})^{\alpha_j} \\
&= \arg\max\nolimits_w (\tfrac{w-\rho V_u(a,j)}{\rho+\eta_j})^{1-\alpha_j} (\tfrac{a\theta - w}{\rho+\eta_j})^{\alpha_j} \\
&= \alpha_j a\theta + (1-\alpha)\rho V_u(a,j)
\end{aligned} \tag{12}$$

A.3 The derivation of fixed point system of $\theta^*(a,j)$ and $\theta^{**}(a,j)$

I start from the expression of unemployed value $V_u(a,j)$, equation **??**:

$$\begin{aligned}
\rho V_u(a,j) &= ab_j + \lambda_j \underbrace{\int_{m_j}^{\infty} \{V_e(w,j) - V_u(a,j)\}^+ dF(w|a,\theta,j)}_{\text{A local offer arrives}} \\
&+ \lambda_{j'} \underbrace{\int_{m_{j'}}^{\infty} \{V_e(w,j') - c(a,j) - V_u(a,j)\}^+ dF(w|a,\theta,j'))}_{\text{A neighbouring offer arrives}}
\end{aligned}$$

Now, I replace the term $V_e(a,j,\theta)$ in the above equation using the following step-wise function:

$$V_e(a,j,\theta) = \begin{cases} \frac{m_j + \eta_j V_u(a,j)}{\rho+\eta_j} & \theta \in [m_j, \hat{\theta}(a,j)) \\ \frac{\alpha_j(a\theta - \rho V_u(a,j))}{\rho+\eta_j} + V_u(a,j) & \theta \in [\hat{\theta}(a,j), \infty) \end{cases}$$

Then I replace $\rho V_u(a,j)$ with its equivalent definition $a\theta^*(a,j)$ then get:

$$
\begin{aligned}
a\theta^*(a,j) \quad = \quad & ab_j + \tfrac{\lambda_j}{\rho+\eta_j}[\underbrace{\boldsymbol{I}\,(\theta^*(a,j) < \frac{m_j}{a})(m_j - a\theta^*(a,j))\left(\tilde{G}(\hat{\theta}(a,j)) - \tilde{G}(\frac{m_j}{a})\right)}_{\text{Local offer with wage } m_j} \\
+ \quad & \underbrace{\int_{\max\{\hat{\theta}(a,j),\theta^*(a,j)\}} a\alpha_j(\theta - \theta^*(a,j))dG(\theta)]}_{\text{Local offer with wage } w_j > m_j} \\
+ \quad & \tfrac{\lambda_{j'}}{\rho+\eta_{j'}}[\underbrace{\boldsymbol{I}(\theta^{**}(a,j) < \frac{m_{j'}}{a})(m_{j'} - a\theta^*(a,j'))\left(\tilde{G}(\theta^{**}(a,j)) - \tilde{G}(\frac{m_{j'}}{a})\right)}_{\text{Neighbouring offer with wage } m_{j'}} \\
+ \quad & \underbrace{\int_{\max\{\hat{\theta}(a,j'),\theta^{**}(a,j)\}} a\alpha_j(\theta - \theta^*(a,j'))dG(\theta)}_{\text{Neighbouring offer with wage } w_{j'} > m_{j'}} \\
+ \quad & \underbrace{(\rho+\eta_{j'})\,(\frac{a(\theta^*(a,j) - \theta^*(a,j'))}{\rho} + c(a,j))\tilde{G}(\theta^{**}(a,j))]}_{\text{The unemployed value difference between staying/moving}}
\end{aligned}
$$

B Preliminary regression results

B.1 Both migrants and commuters are responsive to minimum wage hikes

This section presents the responses of migrants and commuters to minimum wage hikes. I find that low educated workers tend to commute/migrate away from states with higher relative minimum wage (compared to its neighboring state) rather than towards them. More specifically, the fraction of workers commuting out of the state increases and the number of individuals migrating into the local county from other states decreases.

I use the following regression to measure the effect of the relative minimum wage ratio on worker's migration and commuting behaviors:[34]

$$
\log y_{c,t} = \beta_0 + \beta_1 \log \frac{MW_{s(c),t}}{MW_{s'(c),t}} + \epsilon_{c,t} \tag{13}
$$

[34]Ideally, I would distinguish the effect of the own state's minimum wages from the effect of the neighboring state's minimum wages by using the following regression:

$$
\log y_{c.t} = \beta_0 + \beta_1 \log MW_{s(c),t} + \beta_1^{'} \log MW_{s'(c),t} + \epsilon_{c,t}
$$

However, due to the high correlation between $MW_{s(c),t}$ and $MW_{s'(c),t}$, the estimates suffer multicollinearity and become too sensitive to model specification. Therefore, I put the restriction $\beta_1 = -\beta_1^{'}$ to deliver more stable estimates.

Here $y_{c,t}$ is the ratio of migrants or commuters in county c, at time t. I estimate separate regressions for each education group. The minimum wage ratio $\frac{MW_{s(c)t}}{MW_{s'(c)t}}$ compares the minimum wage of $s(c)$, the state containing county c, to the minimum wage of $s'(c)$, the neighboring state of county c. The coefficient β_1 is the primary parameter of interest, which is the elasticity of outcomes y_{it} with respect to the relative minimum wage ratio.

Regression estimates are reported in Table 10. Column (1) reports the elasticity of the flows of migrants and commuters with respect to the change of relative minimum wage ratio. I use the relative minimum wage ratio rather than the absolute minimum wage levels to allow the flexibility that migration and commuting could be driven by either the own state's minimum wage hikes or the neighboring state's minimum wage increases. I find that minimum wage changes have a statistically significant negative effect for low educated migrants. In response to a 1% hike in the relative minimum wage ratio, the flows of low-educated migrants decrease by 0.539%. For commuters, these flows increase by 0.458% in response to a 1% increase in the relative minimum wage ratio.

However, observed commuting and migration changes could respond to other factors happening simultaneously with minimum wage increases. For example, if the local economic conditions are declining for the states with minimum wage increases, I would misattribute these changes to minimum wage changes instead of local economic conditions. Column (2) estimates the same regression model for high educated workers. If local conditions were underlying the observed changes of labor mobility, then high educated workers should present similar patterns, but that is not the case. There is no statistically significant migration response and only moderate commuting response to the same minimum wage increase.[35] While the evidence above does not prove causality, it is consistent with the view that minimum wage policy should have asymmetric effects on workers with different educational levels. Compared with low educated workers, the high educated group receives a higher wage on average, yielding a lower probability to be bound by minimum wage increases. Another concern is that the state-level minimum wage policy may move in tandem with other redistribution policies, such as unemployment insurance benefits, which may also cause asymmetric effects on workers with different levels of education. To minimize this concern, I restrict my sample to the period covered by The Fair Minimum Wage Act of 2007.[36] It is worth noting

[35] I ran the same regression only for the high-school graduates, which are more closely related to high-school dropouts. The estimates are very close to the estimates for the whole high educated group. (This regression result is not reported in table 10)

[36] The Fair Minimum Wage Act of 2007 was implemented by three stages. Stage one increased the minimum wage from $5.15 to $5.85 in 2007. Stage two continued to increase it to $6.55 in 2009. Then the final stage finalized the minimum wage in the level of $7.25 in 2009. Thus I restrict my sample to year 2007-2009 to include the total effect of federal minimum wage change.

Table 10: Migrant and Commuter Flows in Response to Minimum Wage Ratio Changes

	Baseline sample			Restricted sample		Alternative	Extended sample
y_{it}	*(1) Low education*	*(2) High education*	*(3) Whole sample*	*(4) Low education*	*(5) High education*	*(6) Whole sample*	*(7) Low education*
Migrants	**-0.589*****	-0.101	-0.093	**-0.682****	0.082	**-0.148*****	**-0.417*****
	(0.160)	(0.112)	(0.107)	(0.315)	(0.156)	(0.026)	(0.140)
	5,828	8,266	8,330	1,711	2,664	10,459	7,123
Commuters	**0.458****	**0.263****	**0.278****	**0.678***	**0.378****	**0.212*****	**0.442****
	(0.215)	(0.133)	(0.134)	(0.379)	(0.139)	(0.079)	(0.205)
	4,501	7,270	7,129	934	1,794	6,491	5,117
Controls							
Pair FEs	Y	Y	Y	Y	Y	Y	Y
Year FEs	Y	Y	Y	Y	Y	Y	Y
Centriods <75mi	Y	Y	Y	Y	Y	Y	

*The table reports coefficients associated with the log of relative minimum wage ratio ($\log \frac{MW_{st}}{MW_{st}}$) on the log of the dependent variables noted in the first column. All regressions include both county fixed effects and year fixed effects. Columns (1)-(3) provide estimates for all individual between 16-30 based on pseudo county-level variation constructed by ACS PUMS between year 2005-2015. Column (6) uses IRS data. The variable "Migrants" is collected by the IRS Statistics of Income Division (SOI), year 05-15. The variable "Commuters" is extracted from the county-level ACS (09-15) through the interface called American FactFinder (web:https://factfinder.census.gov/) In Column (5)-(6), the sample is restricted to year 2007-2009 when the the Fair Minimum Wage Act of 2007 is enforced. For Column (7), the sample is extended to all county-pairs. Robust standard errors, in parentheses, are clustered at the the paired-county levels. * for 10%. ** for 5%, and *** for 1%. Sample sizes are reported below the standard error for each regression.*

that the federal minimum wage compresses the minimum wage difference between contiguous counties. Therefore, the federal minimum wage should generate the opposite effect for states bound by the federal minimum wage: the commuting flows increase while the migration flows decrease. Columns (4) and (5) report values that are slightly higher (-0.682 and 0.678 compared to -0.589 and 0.458) than my baseline estimates for the low-education group, but not significantly different. The estimates for the high educated group are also similar to my baseline estimates. The elasticity of migration is not statistically significant and the elasticity of commuting is significantly positive but moderate in its magnitude. To sum up, my results are robust to the restricted sample only using the federal-level minimum wage variation, which supports the hypothesis that the potential endogenity of state-level minimum wage change does not bias the estimates.

Another concern is that pseudo county-based statistics may be imprecise. To mitigate this concern, I re-run the same regressions using different data sources in Column (6). The alternative migration data comes from the Internal Revenue Service (IRS) which collects the year-to-year address changes reported on individual income tax returns between 2005-2015.[37] The alternative commuting data comes from the 2009-2015 aggregated county-level ACS.[38] Unfortunately, these two alternative data sets lacks workers' demographic characteristics. Therefore, I can only compare estimates based on the full sample rather than estimates of subgroups classified by their education levels. The estimates using alternative data have similar values but different level of statistical significance. The elasticities of migration and commuting when using alternative data sets are -0.148 and 0.212 compared to my baseline estimates of -0.093 and 0.278.[39]

Lastly, I do another robustness check on the selection of contiguous county pairs. Following Dube et al. (2016), the baseline regression includes county pairs whose centriods are within 75 kilometers because the counties with closer centriods have more similar labor markets. In column (7), I run the same specification using all county pairs. Compared with

[37] IRS data is more robust than other data for a few reasons. First, IRS data covers 95 to 98 percent of the individual income tax filing population. Furthermore, the IRS and ACS display similar declines in migration after 2005, which is not true for other data such as the Panel Study of Income Dynamics (PSID), the Survey of Income and Program Participation (SIPP), and the Current Population Survey (CPS). A detailed discussion comparing different migration data sets can be found in Molloy et al. (2011).

[38] This is collected from the American FactFinder which only provides aggregate moments. Thus it is impossible to further disaggregate moments to get conditional ones on workers' characteristic. See https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtmlfor details.

[39] The larger variance of my baseline estimates is due to the imputation process. One PUMA usually contains several counties, which washes away the inter-county variation when converting the PUMA-based statistics into the county-based statistics. Consequently, the "pseudo" county-level variation should be smaller than the "true" county-level variation, which results in less significant estimates.

Table 11: Minimum wage elasticity for employment stocks and flows

y_{it}	(1)	(2)
Hires	**-0.156*****	0.012
	(0.017)	(0.045)
	84,140	83,280
Separations	**-0.190*****	-0.024
	(0.017)	(0.022)
	84,120	83,246
Employment	**-0.068*****	**-0.039****
	(0.017)	(0.017)
	84,140	83,280
Earnings	**0.056*****	-0.016
	(0.015)	(0.015)
	84,140	83,280
Controls		
County fixed effect	Y	Y
Common time effects	Y	
Pair-specific time effects		Y
Centriods <75mi	Y	Y

*Data source: 2005-2015 Quarterly Workforce Indicator (QWI). This table reports the elasticity of the labor market outcomes listed in the first column. The regression sample is restricted to the counties from 964 county-pairs whose centriods are within 75 miles and includes all workers whose age is between 14-34. Robust standard errors, in parentheses, are clustered at the the paired-county level. * for 10%. ** for 5%, and *** for 1%.*

the estimates using the baseline sample, the elasticities for low educated group are smaller but not different from my baseline estimates. This makes sense because as physical distance increases, workers have less incentive to take opportunities in the neighboring market since moving costs are higher. The regression results in this section suggest that low educated workers tend to move away from counties with minimum wage increases, either by commuting or migration.

B.2 The disemployment effect of local minimum wage hikes

In this section, I show additional evidence that the increase of outflows in response to a minimum wage increase is caused by the decline of local working opportunities. Following

Dube et al. (2007) and Dube et al. (2016), I run the following regression:

$$\log y_{c,t} = \beta_0 + \beta_1 \log MW_{s(c),t} + \beta_2 X_{c,t} + \phi_c + \eta_{p(c),t} + \epsilon_{c,t} \tag{14}$$

where $y_{c,t}$ refers to the local labor market variables, including earnings, employment, separations and hires, in county c and period t. $X_{c,t}$ is the log of the total local population. The coefficient β_1 is the primary variable of interest representing the elasticity of y_{it} with respect to the local minimum wages. Table 11 reports two regressions which only differ in their specification of the time-fixed effect. In Column (1), I restrict the time fixed effect to be common across all county pairs ($\eta_{p(c),t} = \eta_t$) and I find statistically significant disemployment effects in response to local minimum wage changes. The estimated elasticity of employment stock is -0.156. Meanwhile, the elasticities of employment flows are also substantial with minimum wage increases. The hire elasticity and separation elasticity are -0.190 and -0.156, both of which are statistically significant. The fact that the separation elasticity is larger than the hire elasticity is consistent with the negative effect of minimum wage on employment stock. However, when I account for the pair-specific time fixed effect (to control for time-varying, pair-specific spatial confounders), the estimates for the hire elasticity and separation elasticity are not distinguishable from zero. I attribute this change to the existence of spatial spillover effect. After the local county increases its own minimum wage, unemployed workers may seek their jobs in the neighboring county (either by migration or by commuting), which causes disemployment in the neighboring county. As a result, this spillover effect generates a common trend between the counties in one pair. When this pair-specific co-movement is teased out by pair-specific time effect, the estimates of local disemployment effect become less substantial.

C Sample construction appendix

C.1 Minimum wage policies between 2005-2015

In this section, I consider changes of minimum wage policies both on the state and federal level (See Table 12).[40] Between 2005 and 2015, there was only one change to federal minimum wage law, the Fair Minimum Wage Act of 2007.[41] While 78 changes in minimum wage resulted from the Act, the other 164 events were due to state ordinances. Table 12 highlights two important patterns. First, at least 5 states change their effective minimum wage every year. Second, there is significant variation in how often states change their minimum wages. For example, Georgia only changed its minimum wage three times in line with federal minimum wage policy. On the contrary, its neighbor, Florida, makes the most minimum wage adjustments, changing 11 times.[42] Overall, the effective minimum wage increases $0.54 per change on average, but with substantial variation (Table 13). The largest change ($1.90) happened in Michigan in 2005, while the smallest increment ($0.04) happened in Florida in 2010.

One limitation is the scarcity of city-level minimum wage ordinances. Before 2012, only five localities had their own minimum wage laws. As of September 2017, 39 counties and cities have passed local minimum wage ordinances. Due to limited data, I evaluate the effect of county-level minimum wage indirectly. I estimate the baseline model using state-level minimum wage variation but focus on the resulting county-level labor market outcomes. Then, the effect of the county-level minimum wage will be inferred using contiguous border county pairs.

[40] David et al. (2016) document all minimum wage law changes between 1979-2012. My table differs slightly from David et al. (2016) because I extend the sample through 2015 and include DC. Additionally, I have corrected errors in the minimum wages of Pennsylvania and Colorado.

[41] The Act raised the federal minimum wage in three stages: to $5.85 60 days after enactment (2007-07-24), to $6.55 one year after that (2008-07-24), then finally to $7.25 one year after that (2009-07-24).

[42] Two changes happened in 2009.

Table 12: Variation in State Minimum Wages (2005-2015)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Changes
Federal MW	5.15	5.15	5.15	5.85	6.55	7.25	7.25	7.25	7.25	7.25	7.25	3
Alabama												3
Alaska	7.15	7.15	7.15	7.15	7.15	7.75	7.75	7.75	7.75	7.75	8.75	3
Arizona			6.75	6.90	7.25		7.35	7.65	7.80	7.90	8.05	8
Arkansas			6.25	6.25							7.50	4
California	6.75	6.75	7.50	8.00	8.00	8.00	8.00	8.00	8.00	9.00	9.00	3
Colorado			6.85	7.02	7.28	7.28	7.36	7.64	7.78	8.00	8.23	8
Connecticut	7.10	7.40	7.65	7.65	8.00	8.25	8.25	8.25	8.25	8.70	9.15	6
Delaware	6.15	6.15	6.65	7.15	7.15					7.75	8.25	5
D.C.	6.60	7.00	7.00	7.00	7.55	8.25	8.25	8.25	8.25	9.50	10.5	7
Florida	6.15	6.40	6.67	6.79	7.21		7.31	7.67	7.79	7.93	8.05	11
Georgia												3
Hawaii	6.25	6.75	7.25	7.25	7.25							3
Idaho												3
Illinois	6.50	6.50	7.00	7.63	7.88	8.13	8.25	8.25	8.25	8.25	8.25	5
Indiana												3
Iowa			6.20	7.25	7.25							2
Kansas												3
Kentucky												3
Louisiana												3
Maine	6.35	6.50	6.75	7.00	7.25	7.50	7.50	7.50	7.50	7.50	7.50	5
Maryland		6.15	6.15	6.15							8.25	4
Massachusetts	6.75	6.75	7.50	8.00	8.00	8.00	8.00	8.00	8.00	8.00	9.00	3
Michigan			7.05	7.28	7.40	7.40	7.40	7.40	7.40	8.15	8.15	4
Minnesota		6.15	6.15	6.15						8.00	9.00	5
Mississippi												3
Missouri			6.50	6.65	7.05				7.35	7.50	7.65	7
Montana			6.15	6.25	6.90		7.35	7.65	7.80	7.90	8.05	10
Nebraska											8.00	4
Nevada			6.24	6.59	7.20	7.55	8.25	8.25	8.25	8.25	8.25	5
New Hampshire												3
New Jersey		6.15	7.15	7.15	7.15				7.25	8.25	8.38	5
New Mexico				6.50	7.50	7.50	7.50	7.50	7.50	7.50	7.50	4
New York	6.00	6.75	7.15	7.15	7.15					8.00	8.75	6
North Carolina			6.15	6.15								3
North Dakota												3
Ohio			6.85	7.00	7.30	7.30	7.40	7.70	7.85	7.95	8.10	8
Oklahoma												3
Oregon	7.25	7.50	7.80	7.95	8.40	8.40	8.50	8.80	8.95	9.10	9.25	10
Pennsylvania			6.70	7.15	7.15							6
Rhode island	6.75	7.10	7.40	7.40	7.40	7.40	7.40	7.40	7.75	8.00	9.00	5
South Carolina												3
South Dakota											8.50	4
Tennessee												3
Texas												3
Utah												3
Vermont	7.00	7.25	7.53	7.68	8.06	8.06	8.15	8.46	8.60	8.73	9.15	10
Virginia												3
Washington	7.35	7.63	7.93	8.07	8.55	8.55	8.67	9.04	9.19	9.32	9.47	10
West Virginia			6.20	6.90	7.25						8.00	4
Wisconsin	5.70	6.50	6.50	6.50								4
Wyoming												3
Changes	12	17	47	45	47	5	9	8	10	18	24	242

Note: Two minimum wage changes happened in 2009 for Florida.

Table 13: Summary Statistics of State-Level Effective Minimum Wage Changes (2005-2015)

Year	Counts	Mean	S.D.	Min	Max
2005	12	0.621	0.475	0.10	1.45
2006	17	0.605	0.463	0.15	1.85
2007	47	0.831	0.527	0.25	1.90
2008	45	0.541	0.285	0.10	1.35
2009	47	0.533	0.206	0.05	1.00
2010	5	0.548	0.234	0.04	0.70
2011	9	0.160	0.190	0.06	0.70
2012	8	0.315	0.032	0.28	0.37
2013	10	0.160	0.068	0.10	0.35
2014	18	0.362	0.321	0.10	1.00
2015	24	0.629	0.467	0.12	1.85
Total	212	0.538	0.370	0.04	1.90

Note: All units are in nominal dollars.

C.2 The raw ACS 2005-2015 PUMA database cleanup

First, I merge the three raw ACS 2005-2007, 2008-2010 and 2011-2015 data files into one that contains all the relevant variables between 2005-2015. The raw ACS files are downloaded directly from the US Census Bureau, following https://www.census.gov/programs-surveys/acs/data/pums.html. From year 2012, the ACS starts to use the 2010 version of Public Use Microdata Areas (PUMAs). Therefore, I further use the 2000-2010 PUMA crosswalk (https://usa.ipums.org/usa/volii/puma00_{}puma10_{}crosswalk_{}pop.shtml) to map the 2010 PUMA definitions to 2000 PUMA definitions for all the years after 2010. The variables obtained from the raw database are reported in Table 14. The wage measures are adjusted for inflation to be "2015 dollars" equivalent. I further put an age restriction $16 \leq age \leq 30$ on the population.

Next, I convert the individual-level observations into county-level moments, reported in Table 15. The biggest challenge in this process is that the basic geographic units for respondents in ACS is "Public Use Micro Areas"(PUMAs) rather than any jurisdiction geographic entity (i.e. county, city, etc.) in order to comply with census non-identifiable disclosure rule. Therefore, I instead construct the "pseudo" county-level statistics by the following two steps: (1) First, I construct the PUMA-level summary statistics from the corresponding individual-level variables. (2) Second, I impute the county-based measures from the corresponding PUMA-based measures following the crosswalk provided by Michigan

Table 14: Variables obtained from the raw ACS

Variables	Variable labels
serialno	Housing unit/GQ person serial number
puma	Public use microdata area code
st	State code
adjinc	Adjustment factor for income and earnings dollar amounts
agep	Age
pwgtp	Person's weight replicate
migpuma	Migration PUMA
migsp	Migration recode - state or foreign country code
powpuma	Place of work PUMA
powsp	Place of work - State or foreign country recode
schl	Educational attainment
esr	Employment status recode
wagp	Wages or salary income past 12 months
wkhp	Usual hours worked per week past 12 months
wkw	Weeks worked during past 12 months

Table 15: Converting individual-level observations to county-level moments

Individual-level variables	County-level variables	Definition	RAW ACS
High type dummy	High type fraction	Education attainment is high school graduate or above	schl
Low type dummy	Low type fraction	Education attainment is high school dropouts	schl
Employment dummy	Employment rate by types (high and low)	(1) Employed at work and (2) employed with a job but not at work	esr
Hourly wage	Average hourly wage by types (high and low)	"Wages or salary income past 12 months"(wagp) divided by the product of "usual hours worked per week past 12 months"(wkhp) and "weeks worked during past 12 months"(wkw)	wagp, wkhp, wkw
Migrants dummy	The fraction of migrants by types (high and low)	Individuals who report a migration states (not N/A)	migsp
Commuters dummy	The fraction of commuters by types (high and low)	Individuals who report the place of work different from the place of residence	powsp
Labor force dummy	Labor force participation rate by type (high and low)	(1) Employed at work, (2) employed with a job but not at work and (3) unemployed	esr

Table 16: County-level moments obtained from QWI

Variables	Definition	Raw QWI
Average monthly earnings	Average monthly earnings of employees who worked on the first day of the reference quarter.	EarnBeg
Employment	Estimate of the total number of jobs on the first day of the reference quarter.	Emp
Hire rate	The number of workers who started a new job at any point of the specific quarter as a share of employment	HirA/Emp
Separation rate	The number of workers whose job in the previous quarter continued and ended in the given quarter	SepBeg/Emp

Population Studies Center http://www.psc.isr.umich.edu/dis/census/Features/puma2cnty/. The new constructed county-level variables are reported in second column in Table 15, while the original individual-level variables are displayed in first column.

Finally, I label the adjacent counties on the state borderline, consistent with the classification showed in figure 1. Table 2 and the second panel in table 3 report conditional statistics both by educational types and by interior/borderline locations.

C.3 The raw QWI 2005Q1-2015Q4 database cleanup

The time series of county-level variables from QWI are directly obtained through LED extraction tool https://ledextract.ces.census.gov/static/data.html. The age group 19-21, 22-24, 25-34 are selected. The variables displayed in table 16 are calculated and used in this paper.

C.4 Creating the merged sample using multiple data sources

In this session, I will report the final step to merge multiple data sources together into the final completed sample. First, I will use QWI as the baseline data sample. Second, I will merge the ACS into QWI. Third, I will further merge other county-level moments from several different data sources.

- **Step 1: build the baseline data structure with QWI variables.** I create a balanced panel of all contiguous county-pairs with quarterly frequency between 2005Q1-2015Q4. (Obs. 43,596) Then I only keep the observations when one of the two counties changes its minimum wage at quarter t and the information for the following quarter

Table 17: Key moments from other data sources

Variables	Definition	Data source	Year	Completeness
Bargaining power α_j	The annual payroll expenditure in account for the employer value of sales, shipments, receipts, revenue, or business done in restaurant industry (NAICS:722)	Economy Wide Key Statistics (EWKS)	2012	2,243
Matching technology ω_j	The number of divided by the number of total ads and reflects the latest month for which unemployment data is available	The Conference board Help Wanted OnLine (HWOL)	2017.4	2,306
The centroid distance $d_{jj'}$	-	Dube et al. (2010).	-	2,314
The local amenity γ_j	Median gross rent	2011-2015 American Community Survey	2012	2,314

$t+1$ is still completed. (Obs. 3,278) I only keep one quarter observation if minimum wage changes multiple times in one year. (Obs. 2,886)

- **Step 2: merge with the pseudo county-level ACS 05-15 variables.** I merge the ACS into QWI using the indicator combining county-pair and year. I use the QWI data from step 1 as the master file for the merge. Then I only keep all the observation with positive shares of both migrants and commuters. (Obs. 2,314)

- **Step 3: merge additional other variables from several different databases.** I merge several key variables from other data sources which are displayed in the following table. The final sample covers 2,243 observations with all variables completed.