The Minimum Wage and Occupational Mobility: Appendix

September 30, 2021

A Data Construction and Robustness Check

A.1 Micro-level Data Analysis

This section presents regression analysis at the individual level. Similar to the baseline regression, I utilize the dependent coding system in CPS. Occupational switching and stay are defined as in section 2. State and federal level minimum wage data comes from Vaghul and Zipperer (2016). I use the regional price index to calculate real minimum wages.

At the individual level, the measurement error is likely greater than the aggregate ones. As Kambourov and Manovskii (2013) point out, if the interviewee switched employers or experienced usual activity changes, the occupational code is assigned independently by the coder. The larger measurement error will likely attenuate the estimates.

The empirical specification for the micro-level data is:

$$\text{Occupational Switch}_{ist} = \alpha + \beta lnMW_{st} + \delta_t + \lambda_s + \tau_s * t + \Gamma X_i + \Omega Z_{st} + \epsilon_{ist} \tag{A.1}$$

Since CPS allows at most 6 observations for one individual (the 1st and 5th month-in-sample do not allow defining occupational switching), it is not suitable to include the individual fixed effect. The outcome variable is an indicator for occupational switching and staying. Equation (A.1) includes the state and time fixed effect. The individual controls X_i include age, education, race, and gender. The aggregate controls Z_{st} include the state-level manufacturing and retail employment shares.

Table A.1 shows the results for the younger workers and younger, less-educated workers. I use the state-clustered standard error. As expected, the estimates are attenuated compared to the baseline results in table 1. However, the elasticity is not much different. For example, in the baseline regression, when adding the state-specific time trend, the elasticity of the younger workers' occupational mobility is -0.4. In the micro-level analysis, the elasticity is -0.3.

The estimate for the younger, less-educated workers is insignificant. The point estimate is imprecise. Compared to the aggregate regression equation (1), this suggests that the individual level measure error is more severe. In particular, the aggregate level standard error is much smaller.

As shown in the minimum wage literature, the low-wage workers are more affected by the minimum wage. To explore whether the low-wage workers' occupational mobility are more affected, I use the following specification:

(A.2)

$$\text{Occupational Switch}_{ist} = \alpha + \beta lnMW_{st} * Low\_Wage_{ist} + \delta_t + \lambda_s + \tau_s * t + \Gamma X_i + \Omega Z_{st} + \epsilon_{ist}$$

On top of equation (A.1), equation (A.2) interacts the minimum wage with an indicator of being in the five lowest-wage occupations ($Low\_Wage_{ist} = 1$) or the five highest-wage occupations ($Low\_Wage_{ist} = 0$). The list of the occupations is in table A.4. The demographic controls include age, race, gender, education, and an indicator for being in the low-wage occupations. The sample includes workers between 16 and 64. The results in column (3) and (4) in table A.1 suggests that the minimum wage has no significant effect. The result is consistent with the appendix section A.2.4, where at the aggregate level the result is also insignificant.

This is likely because the older workers, regardless of their occupations, are not affected by the minimum wage. I further restrict the sample to include only the younger workers aged 16 to 30. The results in the last two columns in table A.1 show that the minimum wage is negatively correlated with the younger, low-wage workers' occupational mobility while there is no effect on the younger, high-wage workers' occupational mobility. The point estimate is larger in magnitude than that of the younger workers' in equation (A.1). Since the younger, low-wage workers' monthly occupational mobility is 3.4%, the result suggests an elasticity of -0.4, slightly larger than the -0.3 elasticity for the younger workers.

Together with the aggregate level analysis in section 2, the results point to an elasticity of the occupational mobility with respect to the minimum wage of -0.4 to -0.3. The effect concentrates on the workers who are more likely to be affected by the minimum wage, namely the younger, less-educated, and low-wage workers.

Table A.1: The Effect of the Minimum Wage on Occupational Mobility

	(1) Age <30	(2) Age $<30\times$ High-School	(3) Age <64	(4) Age <30
$lnMW_{st}$	-0.008*	-0.008		
	(0.005)	(0.007)		
$lnMW_{st} * Low\_Wage_{ist}$			-0.003	-0.015**
			(0.002)	(0.007)
$lnMW_{st} * High\_Wage_{ist}$			0.001	-0.006
			(0.002)	(0.005)
Education	0.000	0.000***	0.000***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Age	-0.002***	-0.002***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.002***	-0.003***	-0.000	-0.002***
	(0.000)	(0.001)	(0.000)	(0.000)
Manufacturing	-0.022	0.049	-0.005	-0.006
Share	(0.076)	(0.096)	(0.041)	(0.084)
Retail	-0.45***	-0.394	-0.349***	-0.496**
Share	(0.151)	(0.237)	(0.115)	(0.185)
State FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
State-Trend	Y	Y	Y	Y
Low-Wage			Y	Y
Observations	N = 1287425	N = 516764	N = 2963718	N =661862

Notes. Table A.1 presents the results in equation (A.1) and equation (A.2). The sample period is from 2005 to 2016. I weight the regression by the CPS final weight. I use state-clustered standard errors. *** means significant at 1% level, ** means significant at 5% level, * means significant at 10% level.

A.2 Robustness

A.2.1 Placebo Test Details

I illustrate the placebo test in section 2 in detail. I first calculate the number of minimum wage increases during the sample period for each state. It turns out that Iowa has the lowest number of 2, while New York has the highest number of 11 times. 27 states change minimum wage less than five times, while the other 24 states have five or more minimum wage increases. I refer to the first group as the infrequent changers and the second group as the frequent changers.

The infrequent changers is similar to the placebo sample in Dube et al. (2010) and the frequent changers is similar to the actual sample in Dube et al. (2010). The idea is that, the infrequent changers have less variation in their minimum wage policy. They hence act like the controls whereas the frequent changers act like the treatments. By assigning the treatment states' minimum wage policy to the controls and run the regression, we should not see any significant results since the controls do not actually receive the treatment.

In the current context, every state increases its minimum wage at some point, so technically the infrequent changers are not controls. But since they have fewer variations, in the two-way fixed effect regression they act like the controls. I first describe the placebo test by separating the states into the frequent and infrequent changers. Then I repeat the exercise by separating the states into federal minimum wage and state-level minimum wage states, as well as states with large average percentage increases and small average percentage changes.

I randomly assign the minimum wage policies of the frequent changers to the infrequent changers in a one-to-one fashion. I first permute the minimum wage policies among the frequent changers, then assign them to 24 randomly chosen states from the 27 infrequent changers. The total number of mappings is $24! \times C(27, 24)$ in which C denotes the combination function. I take 500 of them and run regression equation (A.1) using only the infrequent changers, with the fictitious minimum wage data. For the younger workers, out of 500 repetitions, only 6.4% give significant estimates at the 5% level. For the younger, less-educated workers, out of the 500 estimates, only 5.3% of them are significant at the 5% level. This suggests there is no evidence of spatial confound.

I repeat the exercise for the above-mentioned groups. All permutations have less than 7% of significant results. The placebo tests hence do not find evidence of spatial confounds.

A.2.2 Endogenous Control Variables

Another potential issue not discussed in section 2 is that the control variables might be endogenous. If the minimum wage changes manufacturing and retail employment share, the effect would bias the estimates in the two-way fixed effect model. To show that this is not the case, I regress the controls on the minimum wage using the two-way fixed effect model. The result in table A.2 shows that the minimum wage has no significant effect on either of the controls.

Table A.2: The Effect of the Minimum Wage on Controls

	(1) Manufacturing Employment	(2) Retail Employment
$lmMW_{st}$	-0.008	0.004
	(0.007)	(0.004)
Observations	7344	7344
R-squared	0.9820	0.9816
State FE	Y	Y
Year Month FE	Y	Y

Notes. The first column regresses state monthly manufacturing employment share on log real minimum wages and state and year-month fixed effects. The second column uses state monthly retail trade employment share as the dependent variable. The sample period is from 2005 to 2016. Table A.2 uses state-clustered standard errors.

A.2.3 Different Sample Periods

I present regression results using equation (1) but with different sample periods. The sample periods in the baseline regression include the Great Recession. It is possible that concurrent policy changes during the Great Recession drives the results. Often, the concurrent policy changes would lead to pre-trend in the panel diff-in-diff regression. While the panel diff-in-diff result in section 2.4.2 suggests that there is no pre-trend, I skip the Great Recession and run the baseline regression using data after 2008.

Table A.3 shows that the conclusion does not dependent on the specific choice of sample periods: negative response remains mostly consistent for younger, less-educated workers.

A.2.4 Occupational Mobility by Wages

In section 2, I separate the workers into sub-groups based on age and education. In the subsection, I separate the workers by low-wage and high-wage occupations. Since occupational choices are endogenous, the estimate includes the selection bias of being in the low- or high-wage occupations. Nonetheless, it is useful to see whether the minimum wage has differential effect on the occupational mobility of workers in low-wage and high-wage occupations.

The regression specification is identical to the baseline two-way fixed effect regression:

$$\left(\frac{\text{Switcher}}{\text{Stayer+Switcher}}\right)_{st} = \alpha + \beta lnMW_{st} + \delta_t + \lambda_s + \Gamma X_{st} + \epsilon_{st}$$

The occupational mobility is constructed for the workers in the 5-lowest-wage occupations and the 5-highest-wage occupations. Table A.4 show the occupational codes.

Table A.3: The Effect of Minimum Wages on Occupational Mobility

	(1) Age 16-30	(2) Age 30-45	(3) High School	(4) College	(5) Age 16-30 × High School
			2008 to 2016 (N = 5508)		
$lnMW_{st}$	-0.018***	0.001	-0.007	-0.003	-0.022***
	(0.0050)	(0.0038)	(0.0045)	(0.0035)	(0.0078)
			2010 to 2016 (N = 4284)		
$lnMW_{st}$	-0.020***	-0.002	-0.011**	-0.004	-0.028***
	(0.006)	(0.0043)	(0.0051)	(0.0038)	(0.0088)
			2012 to 2016 (N = 3060)		
$lnMW_{st}$	-0.016**	0.000	-0.006	-0.001	-0.021*
	(0.008)	(0.0054)	(0.0058)	(0.0044)	(0.0113)
State FE	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y

Notes. Table A.3 presents the results using the baseline regression equation (1) with different sample periods. Table A.3 uses state-clustered standard errors. *** means significant at 1% level, ** means significant at 5% level, * means significant at 10% level.

Table A.4: Low-Wage and High-Wage Occupations 2010 Census Code

Low-Wage Occupations	**Census Code**	**High-Wage Occupations**	**Census Code**
Building & Grounds Clean	4200-4250	Management	0010-0430
Personal Care & Service	4300-4650	Computer & Mathematical	1000-1240
Sales & Related	4700-4965	Architect & Engineer	1300-1560
Office & Admin Support	5000-5940	Life & Social Science	1600-1965
Transportation	9000-9420	Legal	2100-2160

Table A.5 shows the results. I use different sample periods. There is some evidence that the minimum wage is associated with lower occupational mobility for the low-wage occupation workers. For the high-wage occupation workers, there is no significant effect, and the point estimates are positive. The results are consistent with the minimum wage having more bites for the low-wage occupation workers.

Compared to the baseline results and the results in the appendix section A.1, the low-wage workers' occupational mobility do not respond to the minimum wage by much. As discussed in the appendix section A.1, this could be that the older workers' occupational mobility is not sensitive to the minimum wage, regardless of their wages. I hence interact the low-wage workers and the younger workers. The minimum wage significantly decreases their occupational mobility, shown in the last column of table A.5. The estimate for the period 2005 to 2016 is very similar to the one using the micro-level data in table A.1. The result suggests that the elasticity of the younger, low-wage workers' occupational mobility with respect to the minimum wage is -0.4.

A.3 Occupational Switching Via Unemployment

The occupational mobility construction in section 2 does not take into account occupational switch via unemployment. Workers can change occupations by going through unemployment. In this section, I measure occupational mobility by employment-unemployment-employment transitions and study its response to the minimum wage. Note that the CPS allows at most 3 consecutive observations (I need to drop the first and the fifth month in sample). This means that I can only characterize occupational mobility via short-term unemployment.

I merge three monthly files together. An occupational switch is identified if a worker 1. is employed in the first month, unemployed in the second month, and employment in the third month; 2. has different occupational codes in the first and third months. An occupational stayer is identified if a worker 1. is employed in the first month and the

Table A.5: The Effect of Minimum Wages on Occupational Mobility

	(1) Low-Wage	(2) High-Wage	(3) Age $< 30\times$ Low-Wage
	2005 to 2016 (N=7344)		
$lnMW_{st}$	-0.003	0.004	-0.016***
	(0.004)	(0.004)	(0.005)
	2008 to 2016 (N=5508)		
$lnMW_{st}$	-0.009**	0.004	-0.028***
	(0.004)	(0.006)	(0.010)
	2012 to 2016 (N=3060)		
$lnMW_{st}$	-0.008	0.005	-0.028**
	(0.005)	(0.007)	(0.011)
Mobility > 0	97%	78%	83%
Average Mobility	2.1%	1.4%	3.4%
State FE	Y	Y	Y
Year Month FE	Y	Y	Y

Notes. Table A.5 shows the effect of the minimum wage on workers' occupational mobility by low-wage and high-wage occupations as in table A.4. Mobility > 0 is the fraction of positive occupational mobility at the state-month level from 2005 to 2016. Average mobility is calculated using sample period from 2005 to 2016. The standard error is clustered at the state-level. *** means statistically different from zero at the 1% level, ** at the 5% level, * at 10% the level.

third month; 2. has the same occupational code in these two months. 3. does not switch employer or does not have usual activity change in the third month.[53]

[53]The data allows for longer intervals for defining occupational mobility. For example, I could focus on the fourth and fifth month-in-sample and see if there is any change in the occupational code. However, the interval between these two periods is eight months, making it difficult to deduce

I study the effect of the minimum wage on the occupational mobility of the sub-groups as in section 2. I investigate the effect of the minimum wage in the first month denoted by $lnMW_{st}$ and the effect of three-month average minimum wages denoted by $\overline{lnMW}$. The empirical specification is

$$\left(\frac{\text{Switcher}}{\text{Stayer+Switcher}}\right)_{st} = \alpha + \beta lnMW_{st} + \delta_t + \lambda_s + \Gamma X_{st} + \epsilon_{st}$$

The specification is identical to the baseline regression. Table A.6 presents the results. The results are insignificant. This is likely because of poor data quality. For almost all the sub-groups, over 50% of the observations are empty. The monthly average occupational mobility for younger workers is 0.7%, 0.3% for older workers, and 0.9% for younger, less-educated workers. The mobility is much smaller than the one in the baseline regression.

A.4 The Effect of the Minimum Wage on Occupational Transition Rates

Section 2.5.2 shows for the younger workers, the minimum wage is associated with a lower transition rate from non-routine manual occupations to routine cognitive occupations. Table A.7 shows all the estimates.

Table A.8 shows the occupational transition matrix for the younger workers from 2005 to 2016. The transition rates are average monthly transition rates. The occupational mobility is defined at the 4-digit occupational code level, but aggregated to the four occupational categories. For example, the average monthly transition rate from the non-routine cognitive occupations to the non-routine cognitive occupations is 1.0%. This means that 1% of the non-routine cognitive occupation workers switch occupations to another non-routine cognitive occupation at the 4-digit code level in a month. Occupational switchers

whether an occupational code change across these two periods is truly an occupational switch. It is also difficult to interpret which kind of occupational mobility is affected because I do not observe what happens during the eight months.

Table A.6: The Effect of Minimum Wages on Occupational Mobility

	(1) Age 16-30	(2) Age 30-45	(3) High School	(4) College	(5) Age 16-30 × High-School
$lnMW_{st}$	-0.000	0.000	0.001	0.001	0.003
	(0.0030)	(0.0012)	(0.0022)	(0.0009)	(0.0058)
$\overline{lnMW}$	0.000	0.000	0.002	0.001	0.002
	(0.0030)	(0.0012)	(0.0022)	(0.001)	(0.0058)
Mobility > 0	46%	35%	52%	48%	30%
Average Mobility	0.7%	0.3%	0.5%	0.3%	0.9%
Observations	7344	7344	7344	7344	7344
State FE	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y

Notes. Table A.6 identifies an occupational switch if and only if a worker is employed in the first period, unemployed in the second period, employed in the third period, and has different occupational codes in the two employment. An occupational stayer is a worker who has the same occupational code in the first and the third month and who does not change employer or usual activity in the third month. I aggregate the switchers and stayers at the state level and monthly frequency using final weight. $lnMW_t$ is the real minimum wage in the first month. $\overline{lnMW}$ is the three-month average minimum wage. Controls include state manufacturing and retail trade employment shares. Table A.6 uses state-clustered standard errors. *** means significant at 1% level, ** means significant at 5% level, * means significant at 10% level.

and stayers are defined as in section 2.

The routine cognitive occupation workers tend to switch occupations more often than the other workers. The two primary destinations are the routine cognitive occupations and the non-routine cognitive occupations. For the non-routine manual occupation workers, they tend to switcher to other non-routine manual occupations. Although they also have a non-trivial tendency to switch to the routine-cognitive occupations.

For the non-routine cognitive and the routine manual occupations, the workers mainly stay in the occupation category even if they switch occupations at the 4-digit level. This

Table A.7: The Effect of Minimum Wages on Detailed Occupational Transition Rates

From \ To	Non-Routine Cognitive	Non-Routine Manual	Routine Cognitive	Routine Manual
Non-Routine Cognitive	0.005	0.001	0.002	-0.002
	(0.005)	(0.003)	(0.003)	(0.003)
Non-Routine Manual	0.004	0.001	-0.009*	-0.005
	(0.003)	(0.006)	(0.005)	(0.003)
Routine Cognitive	0.001	0.001	-0.004	-0.004
	(0.003)	(0.003)	(0.005)	(0.003)
Routine Manual	-0.001	0.005	-0.003	-0.011
	(0.003)	(0.003)	(0.004)	(0.009)
Observations	612	612	612	612
State FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y

Notes. Table A.7 shows the estimate results from equation (9). The numbers are coefficients of $lnMW_{st}$. For example, the number in the cell (Non-Routine Cognitive, Non-Routine Manual) is the coefficient of $lnMW_{st}$ in equation (9) using annual transition rate from non-routine cognitive occupations to non-routine manual occupations as the dependent variable. Equation (9) identifies an occupation switcher and a stayer the same as in section 2. Table A.7 uses state-clustered standard errors. *** means significant at 1% level, ** means significant at 5% level, * means significant at 10% level.

Table A.8: Occupational Transition Matrix

From \ To	Non-Routine Cognitive	Non-Routine Manual	Routine Cognitive	Routine Manual
Non-Routine Cognitive	1.0%	0.4%	0.5%	0.2%
Non-Routine Manual	0.4%	1.4%	0.8%	0.6%
Routine Cognitive	1.0%	0.7%	1.4%	0.5%
Routine Manual	0.4%	0.5%	0.5%	1.7%

can be seen by comparing the probably on and off the diagonal.

A.5 Federal Minimum Wage States

Table A.9: Federal Minimum Wage States

Alabama	North Dakota
Georgia	Oklahoma
Idaho	South Carolina
Indiana	Tennessee
Kansas	Texas
Kentucky	Utah
Louisiana	Virginia
Mississippi	Wyoming

Notes. Table A.9 includes states that a have binding minimum wage equal to the federal minimum wage from 2005 to 2016.

A.6 More Results from the GSC

In section 2, I show the GSC result for the workers aged 16 to 30. In this subsection, I show the GSC results for the other subgroups of workers.

Table A.10: The Effect of Minimum Wages on Occupational Mobility: GSC Estimation

	(1) Age 16-30	(2) Age 30-45	(3) High School	(4) College	(5) Age 16-30 × High School	(6) Age 16-30 × College
$lnMW_{st}$	-0.011**	0.001	0.000	-0.002	-0.011*	-0.010
	p=0.04	p=0.85	p=0.84	p=0.58	p=0.09	p=0.31
Mobility > 0	97%	94%	96%	99%	81%	67%
Average Mobility	2.9%	1.5%	2.0%	1.8%	3.1%	2.2%
N	7344	7344	7344	7344	7344	7344

A.7 Dynamic Effects of the Minimum Wage

In section 2.4.2, I show the dynamic effects of the minimum wage using a panel DiD method. A similar approach is to include leads and lags of the minimum wage in the two-way fixed effect regression. The so-estimated coefficients are not DiD estimates, but they could show the quasi-elasticity of the minimum wage before and after the increase.

I estimate the following equation:

$$\left(\frac{\text{Switcher}}{\text{Stayer+Switcher}}\right)_{st} = \alpha + \sum_{\tau=-5}^{5} \beta_\tau lnMW_{s,t-\tau} + \delta_t + \lambda_s + \Gamma X_{st} + \epsilon_{st} \tag{A.3}$$

Equation (A.3) differs from the two-way fixed effect regression by including 5 lags and 5 leads of the log real minimum wage. Both equation (A.3) and adding the state-specific time trends aim at capturing the dynamic effects of the minimum wage and testing for robustness. The former focuses on estimating the average effects of the minimum wage for both the lags and the leads across states, while the latter focuses on controlling for underlying linear trends specific to each state.

Figure E.6 plots the coefficients for the subgroups of workers, with the month prior to the minimum wage increase as the baseline. No estimates is significant at the 5% level. For the younger workers and the younger, less-educated workers, the estimate for the month of the minimum wage increase is significant at the 10% level. The other subgroups of workers have small point estimates for all months.

Figure E.6 shows the lack of pre-trend. In that regard, it serves a similar purpose as the panel DiD regression. The point estimates are not comparable to those in table 1. Instead, I calculate the employment effect following Cengiz et al. (2019) as $\Delta\beta \equiv \sum_{\tau=0}^{5} \beta_\tau/6$. The results are in table A.11. Similar to the results from the two-way fixed effect regression, the estimates are negative for the younger and the younger, high-school workers. The point estimates are smaller and significant at the 10% level.

Table A.11: The Effects of the Minimum Wage on Occupational Mobility: Dynamic Effects

	(1) Age 16-30	(2) Age 30-45	(3) High School	(4) College	(5) Age 16-30 $\times$ High School	(6) Age 16-30 $\times$ College
$\Delta\beta$	-0.003*	0.000	-0.002	0.000	-0.005*	-0.002
	(0.0018)	(0.0012)	(0.0012)	(0.0010)	(0.0028)	(0.0026)
Mobility > 0	97%	94%	96%	99%	81%	67%
Average Mobility	2.9%	1.5%	2.0%	1.8%	3.1%	2.2%
N	7344	7344	7344	7344	7344	7344
State FE	Y	Y	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y	Y	Y

Notes. Table A.11 shows the estimates from equation (A.3) as $\Delta\beta \equiv \sum_{\tau=0}^{5} \beta_\tau/6$. The standard error is state clustered. * means significant at the 10% level.

B Model Details and Proofs

B.1 Analytic Solution of the Value Function

In this subsection, I obtain the analytic solution to the value functions. The analytic solution helps study the relation between the minimum wage and the endogenous separation decision.

Proposition B.1. *The value function $J(x)$ is log concave and has the form:*

$$J(x) = \begin{cases} C_0^0 x^{\gamma_0^0} + C_1^0 x^{\gamma_1^0} - A(x,m), & \text{if } \underline{x} \leqslant x \leqslant x_s \\ C_0^1 x^{\gamma_0^1} + C_1^1 x^{\gamma_1^1} - B(x,m), & \text{if } x_s < x < \overline{x} \end{cases} \tag{B.1}$$

in which A and B are functions of productivity x and minimum wage m. The power coefficients γ_0^1, γ_1^0, γ_0^1, γ_1^1 are determined by model parameters and satisfy $\gamma_0^0, \gamma_0^1 < 0$, $\gamma_1^0, \gamma_1^1 > 0$.

The firm's value function is decreasing in the minimum wage m. Fixing the other parameters, the minimum wage shifts the firm's value function downward. The downward shift means that the endogenous separation cutoff $\underline{x}$ moves towards the right, making the match more likely to dissolve. The rightward shift of the endogenous separation cutoff suggests that after the minimum wage increase, the firm is less tolerate for low productivity. Hence, while the minimum wage increase might not cause displacement immediately, it could increase the likelihood that the match ends.

The firm's value function has parameters $C_0^0, C_1^0, C_0^1, C_1^1, \underline{x}, x_s$. They are determined by the boundary conditions. The first two are $J(\underline{x}) = 0$ and $J'(\underline{x}+) = 0$. $J(\underline{x}) = 0$ indicates that at the endogenous separation cutoff, the value of the match is equal to 0, which is the firm's outside option. $J'(\underline{x}+) = 0$ is the smooth pasting condition. The next three boundary conditions are continuities at the on-the-job-search cutoff: $J(x_s-) = J(x_s+)$, $J'(x_s-) = J'(x_s+)$, $J''(x_s-) = J''(x_s+)$. The last boundary condition is an arbitrary value at the upper bound of the value function $J(\overline{x}) = \overline{J}$.

The workers' ability and the occupational skill requirement affects the power coefficients γ's. In particular, the positive power coefficients γ_1^0 and γ_1^1 are increasing in the workers' ability and decreasing in mismatch. This implies that higher-ability workers and better matched workers are less likely to quit to unemployment.

Proposition B.1 shows that for a fixed pair (a, j), the integral in equation (11) is a constant, allowing one to solve for the value function. However, when the value functions of all pairs of (a, j) are solved and integrated back as in equation (11), the resulting constant should be the one used to solve for the value function in the first place. Section B.5.3 proves the existence of the solution.

B.2 Wage Setting Details

In this section I briefly discuss the wage setting in the model. In particular, I first show that the use of generalized Nash bargaining is justified with the introduction of the minimum wage. Then I show that an adapted argument in Moscarini (2005) justifies the generalized Nash bargaining with on-the-job search and switching cost.

The introduction of the minimum wage restricts the valid wage interval. A potential problem with it is that the bargaining game might become non-convex. For example, if the set of wages over which workers and the firms bargain is non-convex, then Nash axiom would not be applicable and one cannot use Nash bargaining. Qin et al. (2015) show that Nash bargaining is justified as long as the game is log-convex. A sufficient condition for log-convexity is that the value functions over which the workers and the firms bargain are log-concave. Proposition B.1 shows that log-concavity is satisfied, justifying the generalized Nash bargaining with the minimum wage.

With the on-the-job search, the bargaining environment is more complex because a worker can bargain with two potential employers simultaneously (Flinn (2006)). In Dey and Flinn (2005) and Postel-Vinay and Robin (2002) where the productivity at the current firm and the poaching firm are observable, a Bertrand-type game in which the employers bid for the services of the employee until the one with the dominated match drops out solves the complexity. In Moscarini (2005) which has a similar environment as this paper, only the expected value of the match at the poaching firm is known. Two firms play an English first-price auction. In the auction, two firms take turns to bid for the worker, while the existing match remains active. The auction ends after a firm fails to raise the last bid. With this, the equilibrium wage remains an affine function of the productivity of the match, and the Nash bargaining is justified. My model introduces a switching cost on top of auction, which requires the workers who decide to switch occupations to pay a one-time fee. In the auction environment, this addition just means that the incumbent firm will be able to retain the worker if the expected productivity in the poaching firm is the same as the current productivity at the incumbent firm. For example, in Moscarini (2005), the worker would switch to the poaching firm if the expected productivity is higher than her current productivity. In my environment with the switching cost, the worker would switch only if the expected productivity is sufficiently above the current productivity, which would cover the switching cost.

The remaining of the section shows the detail derivation of the wage function equation (17). The generalized Nash bargaining holds for every productivity level x, and hence

$$\begin{aligned}
\beta J(x) &= (1-\beta)(V(x)-U) \\
\beta J'(x) &= (1-\beta)V'(x) \\
\beta J''(x) &= (1-\beta)V''(x)
\end{aligned} \tag{B.2}$$

Now let us look at the value functions $V(x)$ and $J(x)$. For simplicity, I suppress the dependence on x for functions $V(x)$ and $J(x)$ and write V and J. Using notation from above, I write the integral $\int_{\mathbb{T}^n} V(x_p, j)dH(j) - \phi$ as $\overline{V}$. Then the value functions are

$$\begin{aligned}(1-\beta)r(V-U) =& (1-\beta)w(x) + (1-\beta)V'\frac{a}{1+|a-j|}x + \frac{1}{2}(1-\beta)\sigma^2x^2V'' \\ &+ (1-\beta)\alpha\lambda\mathbb{I}_{sw}\left(\overline{V}-V\right) - \delta(1-\beta)(V-U) - r(1-\beta)U \\ \beta rJ =& \beta x - \beta w(x) + \beta J'\frac{a}{1+|a-j|}x + \frac{1}{2}\beta\sigma^2x^2J'' - \beta\alpha\lambda\mathbb{I}_{sw}J - \delta\beta J\end{aligned}$$

Using the relations in equation (B.2), the derivatives of J can be replaced by the derivatives of V. We can subtract the first equation by the second one and solve for $w(x)$ to get

$$w(x) = \beta x + (1-\beta)b + (1-\beta)\lambda(1-\alpha\mathbb{I}_{sw})\left(\overline{V}-U\right)$$

B.3 The Stationary Wage Distribution

The randomness in the productivity, together with separation and job finding, leads to the stationary productivity and hence wage distribution. It is important that the wage distribution of the model matches the empirical one in order to study the counterfactual. In particular, the wage distribution is an important factor in determining the effect of the minimum wage as it is related to the fraction of workers with binding minimum wages.

To that end, I derive the analytic solution of the stationary wage distribution equation (15) and summarize the result in proposition B.2.

Proposition B.2. *The stationary wage distribution is:*

$$f(x) = \begin{cases} B_0^0 x^{\eta_0^0} + B_1^0 x^{\eta_1^0}, & \text{if } \underline{x} \leqslant x \leqslant x_s \\ B_0^1 x^{\eta_0^1} + B_1^1 x^{\eta_0^1}, & \text{if } x_s < x < \overline{x} \end{cases} \tag{B.3}$$

The parameters η_0^0, η_1^0, η_0^1, η_1^1 satisfy $\eta_0^0, \eta_0^1 < 0$, $\eta_1^0, \eta_1^1 > 0$.

The stationary wage distribution is double-Pareto, consistent with the empirical wage distribution. Its left Pareto tail is crucial for matching the fraction of workers affected by the minimum. The right Pareto tail, determined by the parameters η's, governs the amount of randomness in the productivity process. The right Pareto tail is locally decreasing in the productivity volatility σ: a larger volatility implies a thinner right tail and hence less inequality at the upper end. The intuition is that when the probability of large negative shock is high, a high productivity is less likely.

B.4 The Wage Compression Channel

Here I explain in detail the wage compression channel in the model. Proposition B.3 is important to the wage compression channel:

Proposition B.3. *Conditional on employment, the worker's value function is increasing in the minimum wage. The amount of the increase is decreasing in worker's ability a but increasing in mismatch $|a - j|$.*

Proposition B.3 implies that the minimum wage has varying effects across the ability and mismatch distribution. High-ability workers will benefit less from the minimum wage increase because the minimum wage rarely binds for them. When the minimum wage increases, there is little increase in their value function. One can view the value function as the integral of future potential wage paths, weighted by the probability of each path. For the high-ability workers, it is rare for their wage paths to drop below the minimum wage. An increase in the minimum wage does not affect their value function by much.

Low-ability workers are more likely to face a binding minimum wage. An increase in the minimum wage would increase their value function by much more. The minimum wage increases the wage on each affected wage path and the measure of wage paths with a binding minimum wage.

Similarly, mismatched workers benefit more from a minimum wage increase. The ones that benefit the most from a minimum wage increase are the low-ability workers in mismatched occupations, conditional on staying employed. The distributional effect of the minimum wage on the workers' value function along two dimensions—ability and mismatch—is the key to the wage compression channel.[54]

To illustrate how the minimum wage can affect a worker's occupational mobility via the wage compression channel, consider a low-ability worker in a mismatched occupation. I fix a pair (a, j) such that a is small and $|a - j|$ is large.

The intercept of the value function and the outside option determines her on-the-job search cutoff point, as shown in figure B.1. Before the minimum wage, the cutoff point is the "old x_s" in the figure. By proposition B.3, when the minimum wage increases, conditional on staying employed, she experiences a large increase in her value function. The value function shifts upward. By comparison, her outside option increases by less. This is because she would have less mismatch in her outside option, which according to proposition B.3 would give her less gain when the minimum wage increases. The minimum wage increase narrows the wage gap between a mismatch and a good match. The new cutoff point "new x_s" moves to the left of the old cutoff point.

Figure B.2 illustrates the effect of the leftward movement of on-the-job-search cutoff point on occupation mobility. Before the minimum wage increase, the low-ability worker searches on the job for other occupations when her output is between $\underline{x}$ and x_s^{old}. Once her output exceeds x_s^{old}, she stops searching on the job, shown in figure B.2 (a). Figure B.2 (b) indicates what would happen when the minimum wage increases. The wage compression effect moves the on-the-job-search cutoff to the left. The on-the-job-search region shrinks. The probability of the worker's output being in the "search" region decreases. A decrease in occupational mobility follows.

[54]While proposition B.3 is intuitive, the proof reformulates the problem as an infinite horizon optimal switching problem. I leave the details in the appendix section B.

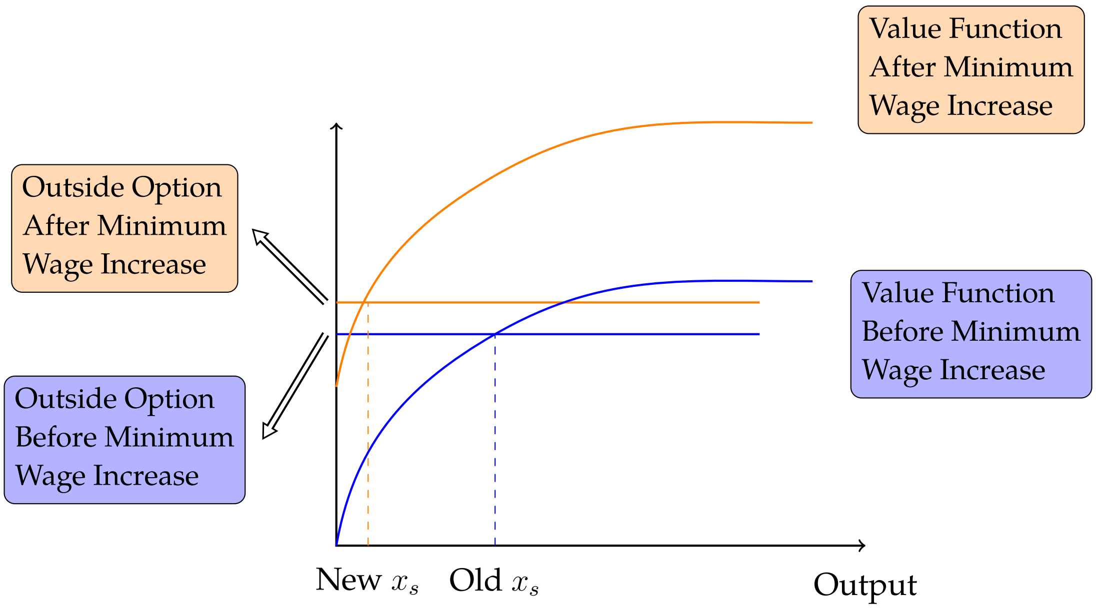


Figure B.1: The Wage Compression Channel

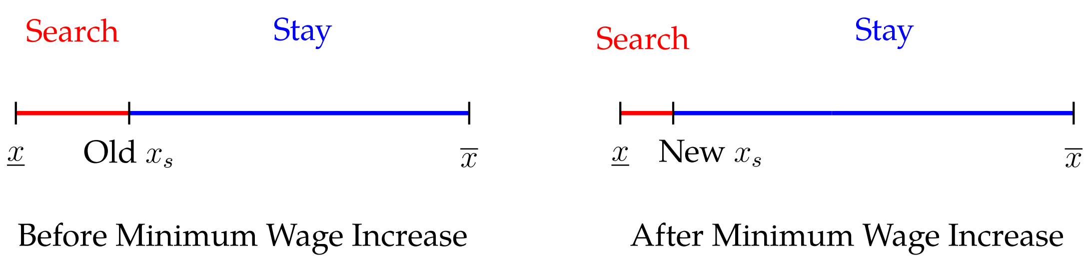


Figure B.2: On-the-Job Search and Occupational Mobility

B.5 Proofs

Throughout the proof

$$\tilde{a} \equiv \frac{a}{1 + |a - j|}$$

denotes the drift of the productivity process.

B.5.1 Derivation of the value function

Proof. The workers' problem is

$$V(x) = \mathbb{E}\bigg[\int_0^{\tau} e^{-rt} w(X_t)dt + e^{-r\tau}G(\tau)\bigg] \quad \text{subject to}$$
$$dX_t = \tilde{a}X_t dt + \sigma X_t dZ_t, \quad dX_{0-} = x$$

in which τ is a stopping time that corresponds to either switching occupations or unemployment and $G(\tau)$ is the outside option. To simplify notations, I denote $\big(\int_{\mathbb{T}^n} V(x_p, \boldsymbol{j})dH(\boldsymbol{j}) - \phi\big)$ by $\overline{V}$. Let τ_1 be the stopping time that the worker is separated exogenously. Let τ_2 be the stopping time that the worker receives outside offer. τ is then the minimum of τ_1 and τ_2. The worker's value function can be written further as

$$\begin{aligned} V(x) = \mathbb{E}[\int_0^{h\wedge\tau_1\wedge\tau_2} & e^{-rt} w(X_t)dt + e^{-rh}V(X_h)\mathbb{I}_{\{h<\tau_1, h<\tau_2\}} \\ & + e^{-r\tau_1}U\mathbb{I}_{\{\tau_1<h,\tau_1<\tau_2\}} + e^{-r\tau_2}\overline{V}\mathbb{I}_{\{\tau_2<h,\tau_2<\tau_1\}}] \end{aligned} \tag{B.4}$$

The symbol $\wedge$ means minimum of the two. Denote the generator of $\{X_t\}_{(t>0)}$ by $\mathcal{L}$, which is given by

$$\mathcal{L}V = \tilde{a}xV' + \frac{1}{2}\sigma^2 x^2 V''$$

Using the Ito's lemma on $V(X_h)$, we have

$$V(X_h) = v(x) + \int_0^h (\mathcal{L}V)(X_t)dt + \text{local martingale}$$

Plug this back into equation (B.4), we have

$$\begin{aligned} \big(1 - e^{-rh}\big)V(x) = \mathbb{E}[\int_0^{h\wedge\tau_1\wedge\tau_2} & e^{-rt}(w + \mathcal{L}V)(X_t)dt \\ & + e^{-r\tau_1}(U - V(x))\mathbb{I}_{\{\tau_1<h,\tau_1<\tau_2\}} + e^{-r\tau_2}(\overline{V} - V)\mathbb{I}_{\{\tau_2<h,\tau_2<\tau_1\}}] \end{aligned} \tag{B.5}$$

Note that because τ_1 and τ_2 are independent,[55] we have

$$\mathbb{I}_{\{\tau_1<h,\tau_1<\tau_2\}} = (1-e^{\delta h})e^{\alpha\lambda h}$$
$$\mathbb{I}_{\{\tau_2<h,\tau_2<\tau_1\}} = (1-e^{\alpha\lambda h})e^{\delta h}$$

Divide equation (B.5) by $\frac{1}{h}$ and let $h \to 0$, I arrive at equation (11). □

B.5.2 Proof of proposition B.1

Proof. I solve the differential equation (14) in the interval $(\underline{x}, x_s)$. In this interval, the differential equation can be written as

$$(r+\delta)J - \tilde{a}xJ' - \frac{1}{2}\sigma^2x^2J'' - (1-\beta)x + (1-\beta)(b+\lambda A(m)) = 0 \tag{B.6}$$

where $A(m)$ is some constant depending on the minimum wage. Define $f(x,m) \equiv (1-\beta)x - (1-\beta)(b+\lambda A(m))$. The general solution to equation equation (B.6) is

$$J(x) = C_0^0 x^{\gamma_0^0} + C_1^0 x^{\gamma_1^0} \tag{B.7}$$

The parameters γ_0^0 and γ_1^0 can be calculated directly via

$$\begin{aligned} \gamma_0^0 &= -\frac{\tilde{a}}{\sigma^2} + \frac{1}{2} - \sqrt{(\frac{1}{2} - \frac{\tilde{a}}{\sigma^2})^2 + \frac{2(\delta+r)}{\sigma^2}} < 0 \\ \gamma_1^0 &= -\frac{\tilde{a}}{\sigma^2} + \frac{1}{2} + \sqrt{(\frac{1}{2} - \frac{\tilde{a}}{\sigma^2})^2 + \frac{2(\delta+r)}{\sigma^2}} > 0 \end{aligned} \tag{B.8}$$

Equation (B.6) also admits a special solution

$$A(m,x) = \frac{2}{\sigma^2(\gamma_1^0-\gamma_0^0)}\left[x^{\gamma_1^0}\int_x^{+\infty} s^{-\gamma_1^0-1}f(s,m)ds + x^{\gamma_0^0}\int_0^x s^{-\gamma_0^0-1}f(s,m)ds\right] \tag{B.9}$$

[55] To construct such independent stochastic processes, see e.g. Rogers and Williams (2000).

where $f(x,m)$ is a function depending on the minimum wage. Specifically, when the wage function is smaller than the minimum wage, $f(x,m) = m$. The solution to equation (B.6) hence has the form

$$J(x) = C_0^0 x^{\gamma_0^0} + C_1^0 x^{\gamma_1^0} - A(m,x) \tag{B.10}$$

Solving for $A(m,x)$, I arrive at equation (B.1) for $J(x)$ in $(\underline{x}, x_s)$. Similarly, I can solve for the value function in $(x_s, +\infty)$. The power coefficients of solution γ_0^1 and γ_1^1 are given by

$$\begin{aligned} \gamma_0^1 &= -\frac{\tilde{a}}{\sigma^2} + \frac{1}{2} - \sqrt{(\frac{1}{2} - \frac{\tilde{a}}{\sigma^2})^2 + \frac{2(\alpha\lambda + \delta + r)}{\sigma^2}} < 0 \\ \gamma_1^1 &= -\frac{\tilde{a}}{\sigma^2} + \frac{1}{2} + \sqrt{(\frac{1}{2} - \frac{\tilde{a}}{\sigma^2})^2 + \frac{2(\alpha\lambda + \delta + r)}{\sigma^2}} > 0 \end{aligned} \tag{B.11}$$

□

Two corollaries are immediate after proposition B.1.

Corollary B.1. *There are boundary conditions such that the value function equation* (B.1) *is strictly increasing.*

Proof. To see this, I abstract from the other boundary conditions and consider only $J(\underline{x}) = 0$ and $J(x_s) = B$. This define a mapping from $(\underline{x}, B)$ to (C_0^0, C_1^0). Fix $x_s > 1$ and let $B \longrightarrow +\infty$, it must be the case that $C_1^0 \longrightarrow +\infty$, otherwise the function remains bounded in the interval $(\underline{x}, x_s)$. C_1^0 can be made so large that $J(x)$ is increasing on $(\underline{x}, x_s)$ because we can choose x_s so that the first term remains bounded. This completes the proof. □

Corollary B.2. *The threshold x_s characterizes the workers' on-the-job search decision.*

Proof. If x_s is unique then this is clearly the case. By lemma B.1, the value function can be chosen to be monotonic. Under such a condition, x_s is unique and the workers search on the job if their output is between $\underline{x}$ and x_s. □

B.5.3 Existence of Solutions to the Family of ODEs Equation (11) and Equation (14)

This subsection proves that the family of ODEs equation (11) and equation (14) has a solution.

Proof. I utilize tools from functional analysis to prove the proposition. I make the ability and the occupation distribution to be more general so that it is given by a joint CDF $N(a,j)$ instead of independently by $G(a)$ and $H(j)$. To prove the existence of a solution, I need one more assumption:

Assumption: The joint CDF $N(a,j)$ has a continuous pdf $n(a,j)$.

The problem at hand can be restated in the following abstract form:

1. Denote the constant in the ODE equation (11) by $f(a)$. That is, I define $f(a) \equiv \int_{\mathbb{T}^n} V(x_p,a,j)n(a,j)dj$. I let the function to be in the Banach space $L^1(\mathbb{T}^n)$. That is, the function is integrable on $\mathbb{T}^n$. Given a value $f(a)$, there is a corresponding solution of the value function $V(x,a,j)$ that is twice continuously differentiable. I denote $\mathbb{V}(a)$ to be the $L^1([0,\overline{x}]\times\mathbb{T}^n)$ valued function $V(\cdot,a,\cdot)$. By the uniqueness of the solution, I can define a mapping T that maps from $L^1(\mathbb{T}^n)$ to $L^1([0,\overline{x}]\times\mathbb{T}^n)$:

   $$Tf = \mathbb{V} \tag{B.12}$$

   This mapping is bounded by standard result in the ODE literature. By the uniqueness of solution and linearity of the differential operator, it is also linear.

2. Having obtained a family of solutions $\{V\}_{(a,j)}$, I fix a point x so that the family of solutions is mapped into a family of functions on $\mathbb{T}^n\times\mathbb{T}^n$. I denote this mapping by P:

   $$PV = V(x,\cdot,\cdot) \tag{B.13}$$

This mapping is clearly linear and bounded:

$$P(V_1 + V_2) = V_1(x, \cdot, \cdot) + V_2(x, \cdot, \cdot) \tag{B.14}$$

3. I integrate the function obtained in step 2 by the CDF $N(a, j)$ with respect to j. Denote this mapping Q:

$$QV(x, a, j) = \int_{\mathbb{T}^n} V(x, a, j) n(a, j) dj \equiv g(a) \tag{B.15}$$

The end result is that I map the function $f(a)$ into the function $g(a)$ by the operators T, P, Q. That is,

$$QPT \circ f = g \tag{B.16}$$

Note that all three operators Q, P, T are linear and bounded. Let us denote the composite mapping by R. The problem reduces to finding a fixed point of the mapping R so that

$$Rf = f \tag{B.17}$$

This can be done by observing that the operator Q is a compact operator. For a proof, see Lax (2002). Since P and T are also linear bounded operators, the composite map R is compact. This mapping is non-trivial, so there is at least one non-zero eigenvalue ξ with eigen-function f such that $Rf = \xi f$. I could now choose the job finding rate to be ξ^{-1} so that $R'f = f$.

□

B.5.4 Proof of proposition B.2.

Proof. The Fokker-Planck equation that the stationary productivity distribution is equation (15). Equation (15) implies that the density flowing into and out of any interior point in $(\underline{x}, \overline{x})$ must be equal. Equation (15) holds regardless of the existence of the minimum wage. This is because the minimum wage does not affect the worker's productivity process equation (10). The existence of the cutoff points $\underline{x}$ and x_s is given by the optimal stopping problem which is also not affected by the minimum wage. The minimum wage only affects the locations of the cutoff points.

Equation equation (15) is a second-order ODE. At the lower-bound $\underline{x}$, the distribution $f(x)$ satisfies an absorbing boundary condition. The corresponding equation is $f(\underline{x}+) = 0$. The worker will quit to unemployment once her productivity is below $\underline{x}$. At the upper-bound, the solution has an reflecting boundary condition by assumption:

$$(\tilde{a} - \sigma^2) f(\overline{x}) = \frac{1}{2}\sigma^2 \overline{x} f'(\overline{x})$$

I solve the equation equation (15) in the interval $(\underline{x}, x_s)$. The corresponding ODE is

$$\frac{\sigma^2}{2} x^2 f''(x) + (2\sigma^2 - \tilde{a}^2) x f'(x) + (\sigma^2 - \tilde{a} - \delta) f(x) = 0 \tag{B.18}$$

Similar to the proof of proposition 1, the general solution is

$$f(x) = B_0^0 x^{\eta_0^0} + B_1^0 x^{\eta_1^0} \tag{B.19}$$

in which the parameters η_0^0 and η_1^0 are given by

$$\begin{aligned} \eta_0^0 &= -\frac{2\sigma^2 - \tilde{a}}{2} + \frac{1}{2} - \sqrt{(\frac{1}{2} - \frac{2\sigma^2 - \tilde{a}}{2})^2 + \frac{2(\tilde{a} + \delta - \sigma^2)}{\sigma^2}} < 0 \\ \eta_1^0 &= -\frac{2\sigma^2 - \tilde{a}}{2} + \frac{1}{2} + \sqrt{(\frac{1}{2} - \frac{2\sigma^2 - \tilde{a}}{2})^2 + \frac{2(\tilde{a} + \delta - \sigma^2)}{\sigma^2}} > 0 \end{aligned} \tag{B.20}$$

Similarly I can solve for $f(x)$ in the interval $(x_s, \overline{x})$ and calculate that

$$(B.21)\qquad \begin{aligned} \eta_0^1 &= -\frac{2\sigma^2 - \tilde{a}}{2} + \frac{1}{2} - \sqrt{(\frac{1}{2} - \frac{2\sigma^2 - \tilde{a}}{2})^2 + \frac{2(\tilde{a} + \delta + \alpha\lambda - \sigma^2)}{\sigma^2}} < 0 \\ \eta_1^1 &= -\frac{2\sigma^2 - \tilde{a}}{2} + \frac{1}{2} + \sqrt{(\frac{1}{2} - \frac{2\sigma^2 - \tilde{a}}{2})^2 + \frac{2(\tilde{a} + \delta + \alpha\lambda - \sigma^2)}{\sigma^2}} > 0 \end{aligned}$$

Two more conditions pin down the stationary distribution. The first is that total flow in and out of unemployment is constant. The second condition is that the total flow in and out of employment must balance. This completes the proof. □

B.5.5 Proof of proposition B.3.

Proof. Let us first formulate the worker's problem in a different way. The worker chooses an increasing sequence of stopping times, τ_n, $\tau_n \to +\infty$, representing the decision of when to search on the job and when to quit to unemployment, and j_n is the occupational the worker wants to switch to or the state of unemployment. Let $p = \{\tau_n, j_n\}_{n\in\mathbb{N}}$, the worker's problem can be written as:

$$(B.22)\qquad V(x) = \sup_p \mathbb{E}\left[\int_0^{+\infty} e^{-rt} w(X_t^{x_s,j}, j)dt - \sum_{n=1}^{\infty} e^{-r\tau_n}\phi\right]$$

where $w(x, j)$ is the payoff function in occupation j, in this case given by equation (17). $X_t^{x_s,j}$ is the output process when the worker is in occupation j with the initial output x_s. $\sum_{n=1}^{\infty} e^{r\tau_n}\phi$ is the sum of the discounted switching cost.

It can be shown that the solution is given by the variational inequality

$$(B.23)\qquad \min\left[rV_i - \tilde{a}xV_i' - \frac{1}{2}\sigma^2 x^2 V_i'' - w(x), V_i - \int_{\mathbb{T}^n} V_j dH(j) + \phi\right] = 0$$

V_i is the value function on occupation indexed i, similarly for V_j. There is an output level x_m such that if $x < x_m$, $w(x)$ is equal to the minimum wage m. Clearly, x_m is increasing m. It is also straight forward to see that when $x < x_m$, $\partial V_i/\partial m > 0$: solving the ODE when $x < x_m$ we get that the minimum wage enters the value function linearly. By the smooth pasting condition and the uniqueness of the solution, this implies that $\partial V_i / \partial m > 0$ everywhere. The interpretation of the result is that conditional on employment, workers' value function is increasing in the minimum wage.

The increase in the value function because of the minimum wage is decreasing in the drift $\tilde{a}$ because $\partial^2 V_i / \partial m \partial \tilde{a} < 0$. The result comes from solving for the value function when $w(x) = m$, in which the solution is linear in m with multiply coefficient equal to $2/\sigma^2(\gamma_1^0 - \gamma_0^0)$. Taking the derivative of this multiply coefficient with respect to $\tilde{a}$ shows that it is decreasing in $\tilde{a}$, or the difference $(\gamma_1^0 - \gamma_0^0)$ is increasing in $\tilde{a}$. Mismatch decreases $\tilde{a}$, leading to a larger positive effect of the minimum wage on workers' value function. The rest of the proof is the same as the one shown in section 4. $\square$

C A Discussion of the Effect of the Minimum Wage on Search Effort and Labor Force Participation

There is literature that emphasizes the effect of minimum wages on increasing the search effort of unemployed workers (e.g. Acemoglu (2001), Flinn (2006), Ahn et al. (2011)). The model can extend to allow for such a response: let $e(m)$ denote the increase in search effort as a function of the minimum wage. The effective measure of job seekers, i.e. equation (18) becomes

$$\begin{aligned} s =& e(m)\left[1 - \int\int\int_{\underline{x}}^{\overline{x}} f(x,a,j,m)dxdG(a)dH(j)\right] \\ &+ \alpha \int\int\int_{\underline{x}}^{x_s} f(x,a,j,m)dxdG(a)dH(j) \end{aligned} \tag{C.1}$$

The increase in search effort would counteract the decrease in vacancy posting. If it dominates, the job arrival rate could increase.

I note first that there is little empirical evidence of higher search effort after a minimum wage increase. Adams et al. (2018) find that recent minimum wage increases lead to a transitory increase in the search effort of unemployed workers. In the stationary equilibrium, such a transitory increase becomes irrelevant. More importantly, the wage compression channel is independent of the search effort response. It unambiguously reduces occupational mobility.

Another related issue is that I do not model labor force participation. If the minimum wage induces labor market entry of non-participants, the effective measure of job seekers could increase, leading to an increase in job arrival rate. It could also mean that the minimum wage can increase efficiency by ex post changing the bargaining power of the workers.[56] The model can accommodate the argument in a reduced-form way: there exist

[56]This is the point made by Flinn (2006). The minimum wage has a similar effect in my model, but the welfare implication is less relevant because there is no labor force participation decision.

an equilibrium after the minimum wage increase with a higher job arrival rate λ. To see this, note that while firm's value function is decreasing in the minimum wage, it is also increasing in λ by equation (B.11). If the increase in the job arrival rate λ dominates, e.g. because of increased labor force participation, there exists an equilibrium in which both λ and the firm's value function increases after a minimum wage increase so that the free entry condition equation (19) still holds.

The implication of the preceding argument is that while the model does not include labor force participation, it is still present when estimating the model. Note that the wage compression channel only concerns the employed workers and hence is independent of the potential increase in labor force participation. It exists as long as the workers' wage is negatively correlated with mismatch.

It is worth pointing out that an increase in the job arrival rate will also increase the occupational mobility. If the labor force participation effect is strong such that the job arrival rate increases, a decrease in the occupational mobility would imply that the wage compression channel dominates.

D Numerical Details

I present the details of the model estimation in this section. I discretize the worker ability distribution so that there are ten levels of ability: $(0.1, 0.2, ..., 1)$. I do the same for the occupation distribution. I pool 2005 to 2016 CPS Merged Outgoing Rotation Groups (MORG) data to be the sample. In this sample, 28.3% of the workers complete college degree and above, 28.5% of the workers have associate degree or vocational training, and 43.2% of the workers have high school degree or less. I set grid 1 and 4 to be the low ability workers, 5 to 7 to be the medium ability workers, and 8 to 10 to be the high ability workers. I calibrate the parameters of the Beta distribution to match the empirical composition of the workers by education exactly. The resulting parameterization is $Beta(0.8877, 0.9415)$.

As mentioned in section 5, I introduce the search accuracy parameter ρ which determines the probability that a worker is sorted into her optimal occupation. The joint distribution of ability and occupation is hence implied by the distribution of workers $Beta(0.8877, 0.9415)$ and ρ. ρ impacts the occupational mobility rate, the effect of minimum wages on occupational mobility, and wage gain from switching occupations. The main target for ρ is the 1% wage gain from switching calculated by Guvenen et al. (2020).

An important parameter in the model is the initial output x_p. It determines the occupational mobility and the measure of unemployment. I set it as a function of ability a:

$$x_p(a) = c_0 + c_1 a \tag{D.1}$$

I set c_0 to be 5.35 and c_1 to be 15 to match the empirical wage distribution.[57] Both the on-the-job-search cutoff equation (23) and the endogenous separation cutoff equation (24)

[57] I need to modify the proof in the appendix section B.5.3 to adapt for the case in which initial output is an affine function of ability rather than fixed. When I have discrete distribution of abilities and occupations this can be done easily.

are functions of the initial output.

Having calibrated the distribution of (a, j) and the initial output x_p, the economy for a fixed pair of (a, j) is

$$
\begin{aligned}
rV(x) &= w(x) + \frac{a}{1+|a-j|}xV'(x) + \frac{1}{2}\sigma^2 x^2 V''(x) - \delta[V(x) - U] \\
&\quad + \alpha\lambda\mathbb{I}_{sw}\left[\int_{\mathbb{T}^n} V(x_p, j)dH(j) - V(x) - \phi\right] \\
rU &= b + \lambda\left[\int_{\mathbb{T}^n} V(x_p, j)dH(j) - U\right] \\
rJ(x) &= x - w(x) + \frac{a}{1+|a-j|}xJ'(x) + \frac{1}{2}\sigma^2 x^2 J''(x) - \delta J(x) - \alpha\lambda\mathbb{I}_{sw}J(x) \\
w(x) &= \max\left\{\beta x + (1-\beta)b + \lambda(1-\beta)(1-\alpha\mathbb{I}_{sw})\left[\int_{\mathbb{T}^n} V(x_p, j)dH(j) - \phi - U\right], m\right\} \\
\kappa &= \int\int \lambda^{\frac{\zeta}{\zeta-1}} J(x_s, a, j, m)dG(a)dH(j) \\
\frac{dX_t}{X_t} &= \frac{a}{1+|a-j|}dt + \sigma dZ_t \\
\frac{\sigma^2}{2}x^2 f''(x) &+ \left[2\sigma^2 - \left(\frac{a}{1+|a-j|}\right)^2\right]xf'(x) + (\sigma^2 - \frac{a}{1+|a-j|})f(x) \\
&\quad - (\delta + \alpha\lambda\mathbb{I}_{\{x<x_s\}})f(x) = 0
\end{aligned}
$$

The first three equations are the value functions. The fourth equation is the wage function. The fifth equation is the free-entry condition, which relates the value function with the job-finding rate. The last two equations are the evolution of the state variables, namely the productivity and the productivity distribution. Note that the productivity distribution does not affect the value functions. It determines the measure of unemployed workers (equation (18), equation (19), equation (D.1)) and the occupational mobility (equation (20), equation (D.1)).

Proposition B.1 shows that the firm's value function has an analytic solution with six undetermined coefficients. The value function of the workers has a similar analytic solution. One way to solve for the value functions is to use the boundary conditions to pin down the undetermined coefficients. Alternatively, I could specify the on-the-job-search cutoff x_s and the endogenous separation cutoff $\underline{x}$ as a function of the state variables, solve for the value function, and verify that the solution is consistent.

I follow Lise and Robin (2017) and use GMM to estimate the parameters. In constructing the variance-covariance matrix of the vector of moments, the moments consist of only means. I use the Newey-West estimator to estimate the variance-covariance matrix as in Lise and Robin (2017) and choose the lag order to be 6.

The numerical algorithm is the following:

1. Fix the vector of parameters in table 4.

2. Fix the minimum wage to be \$7.25. Fixed the job-finding rate to 0.36.

3. Given the cutoffs x_s and $\underline{x}$ (equation (23) and equation (24)), solve for the value functions by invoking proposition B.1.

4. Update the job finding rate via the free-entry condition equation (19).

5. Calculate the wage distribution, the measure of unemployed workers, and the occupational mobility using equation (21). Calculate the fraction of workers with wages less than \$15.

6. Increase the minimum wage to \$8 (10% increase). Set a new job-finding rate λ'.

7. Repeat 3 and 4.

8. Calculate the elasticity of employment and occupational mobility.

9. Move to the next vector of parameters. Repeat until the GMM squared error is minimized.

In the GMM estimation, I set weights to be 100 for moments 5 to 13. The weights for the other moments are set to be 1. The model matches the moment targets quite well given the heterogeneity. Specifically, the employment elasticity and the occupational mobility elasticity are matched quite well which are the main focus of the paper.

Figure D.1 plots the empirical wage distribution in blue and the simulated wage distribution in orange. I construct the empirical wage distribution using CPS merged outgoing rotation group data from 2005 to 2016. I calculate the real wages using the 2012 chain-type price index. The simulated wage distribution comes from averaging 500 periods of the realized wages in the estimated model when the minimum wage is $7.25. The Y-axis is cutoff at 0.06 for better visual. There is a pike at $7.25 because of the minimum wage workers.

The simulated wage distribution gives more weight on the medium wage range and falls short of accounting for the density in low wage and extreme high wage range. The overall shape and the decay of the right tail match the empirical distribution.

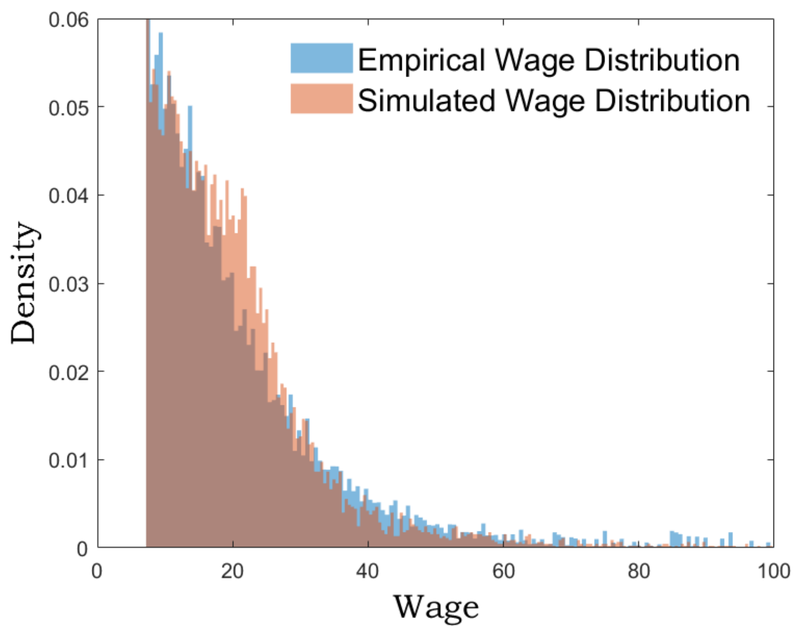


Figure D.1: Model Simulated and Empirical Wage Distribution

In section 5 I decompose the effect of minimum wages on occupational mobility and aggregate output by looking at the wage compression channel alone and by looking at the overall effect. I obtain the results without employment effect by setting p_2 to be 0 and λ' to be 0.36. p_2 governs the displacement effect of the minimum wage while λ' governs the effect on vacancy posting.

I plot the average wage by workers' ability in (a) of figure D.2. The average wage increases as the worker's ability goes up. The sharp rise shows that ability difference contributes to the fat right tail of the wage distribution. The increase in average wage by occupation skill intensity is less significant, as shown in (b) of figure D.2. The comparison implies that sorting is imperfect because workers cannot target their optimal occupations with certainty.

To shut down the wage compression channel and focus on the employment effect channel, I set the parameter s_2 to be 0, which determines the response of occupational mobility to the minimum wage. The occupational mobility for low ability workers decreases by 3%, and overall occupational mobility decreases by 2%. The effects are almost additive: when I restrict the model to have no employment effect, the corresponding estimates are 42% and 28% while the results with both channels are 44% and 30% respectively.

Turning to the effect on aggregate output, when I restrict the model to have only employment effect, the estimate shows that aggregate output decreases by 0.15% when the minimum wage increases to $15. This estimate is insignificant at 5% level.

D.1 Wage Inequality

While the aggregate output declines after the minimum wage increases to $15, the wage inequality, measured by the median-to-10th-percentile ratio, is 56% of that when the minimum wage is $7.25. This is because a larger fraction of the surplus goes to workers and firms bear the profit loss after the minimum wage increase.

The reduction in wage inequality is partly offset by declines in occupational mobility. In other words, some of the low-ability workers would have earned a wage above the $15 minimum wage by switching to the occupation that better match their skills. After shutting down the wage compression channel so that the reduction in occupational mobility becomes negligible, the median-to-10th-percentile ratio is 55% of that when the minimum wage is $7.25. In other words, if the minimum wage increase were not to decrease occu-

(a) Average Wage by Ability

(b) Average Wage by Occupation

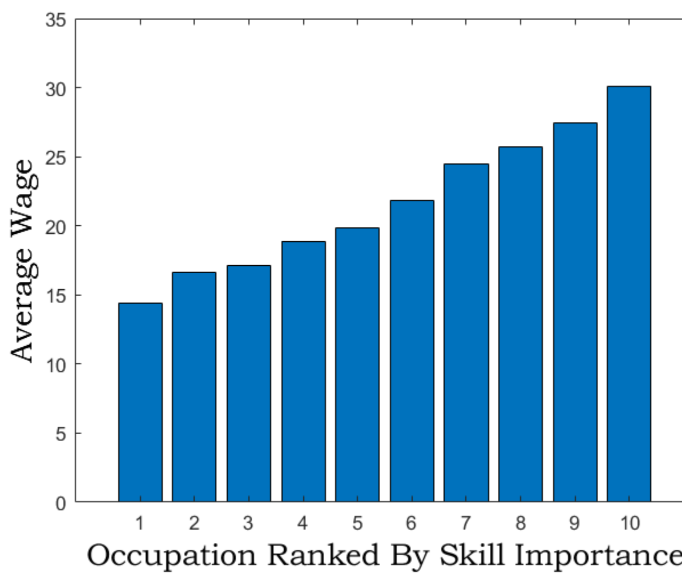


Figure D.2: Model Simulated Average Wage

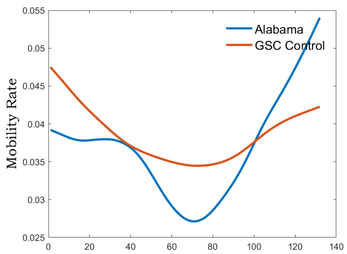
pational mobility, the reduction in wage inequality would be more significant.

E Extra Figures

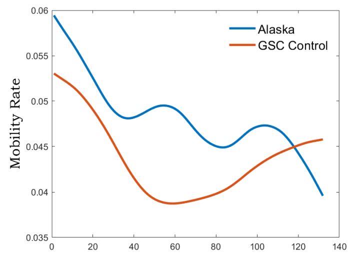
E.1 States and Their GSC Controls

Figure E.1: GSC Fit

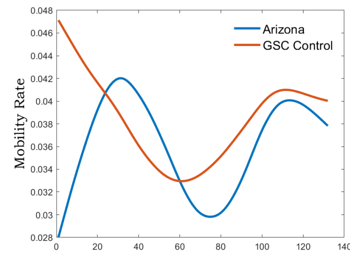
(a) Alabama



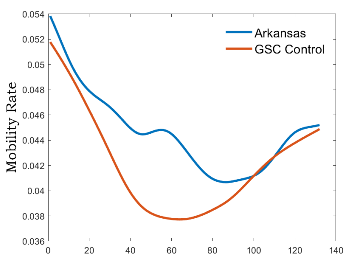
(b) Alaska



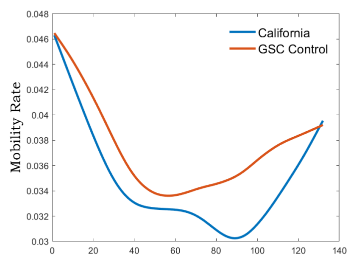
(c) Arizona



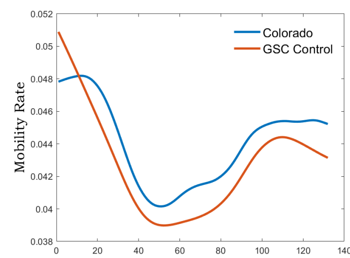
(d) Arkansas



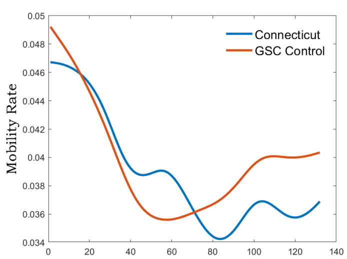
(e) California



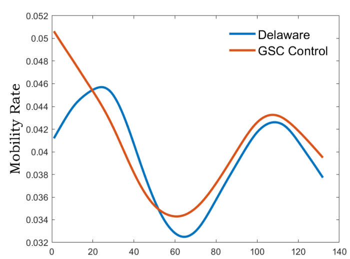
(f) Colorado



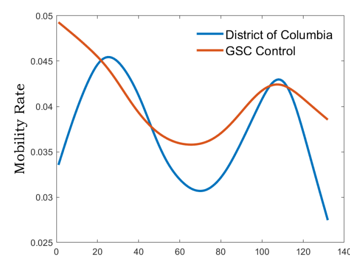
(g) Connecticut



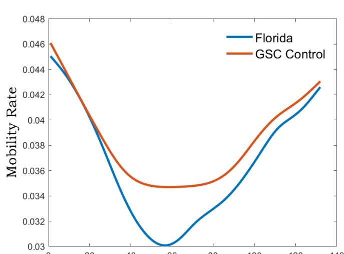
(h) Delaware



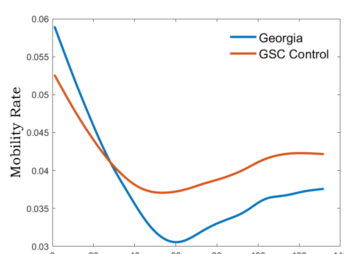
(i) District of Columbia



(j) Florida



(k) Georgia



(l) Hawaii

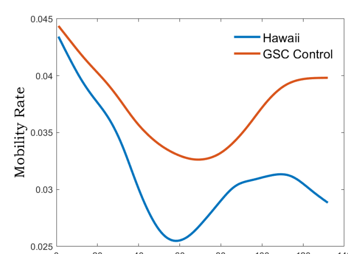
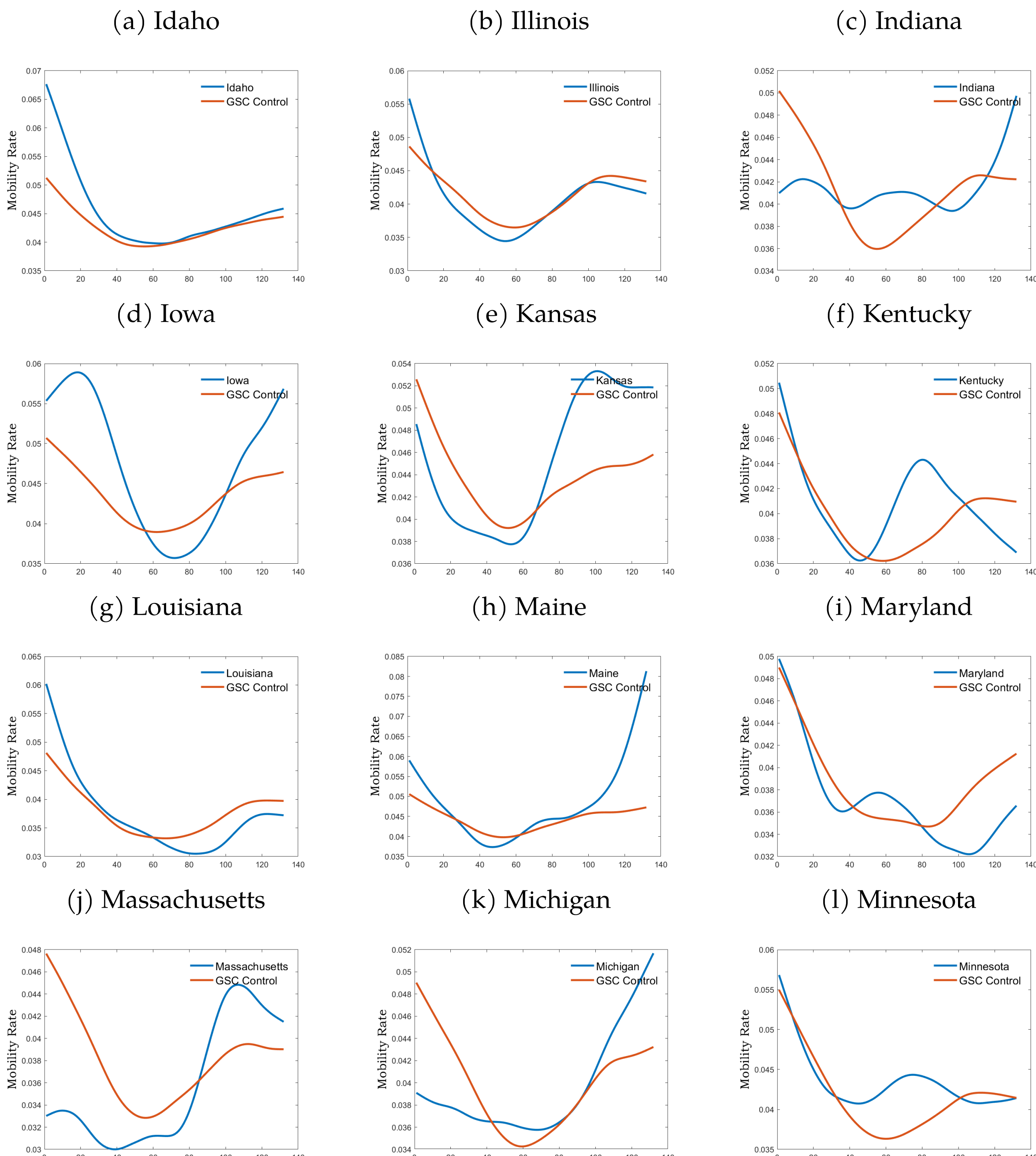


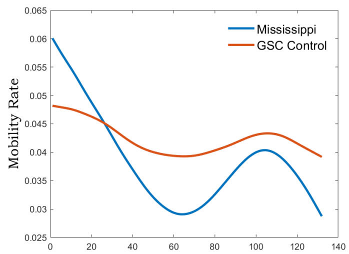
Figure E.2: GSC Fit



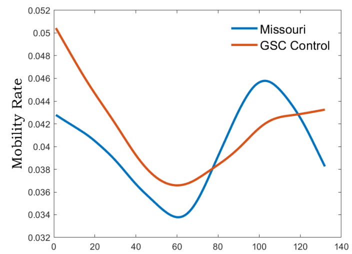
(a) Idaho (b) Illinois (c) Indiana
(d) Iowa (e) Kansas (f) Kentucky
(g) Louisiana (h) Maine (i) Maryland
(j) Massachusetts (k) Michigan (l) Minnesota

Figure E.3: GSC Fit

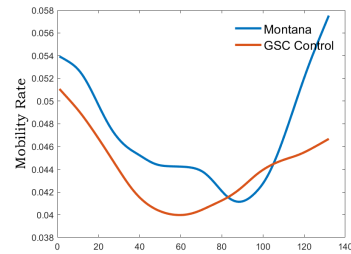
(a) Mississippi



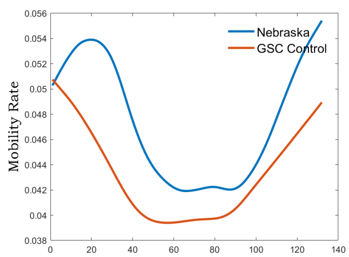
(b) Missouri



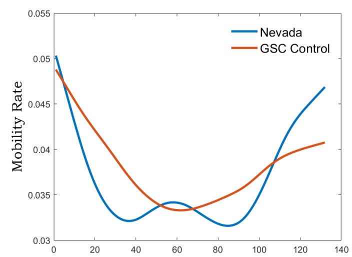
(c) Montana



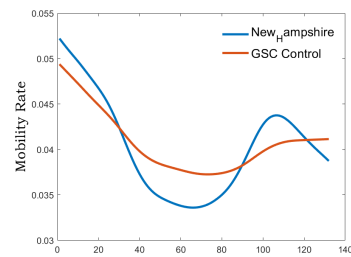
(d) Nebraska



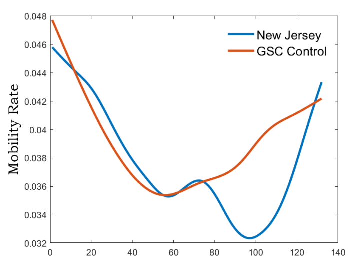
(e) Nevada



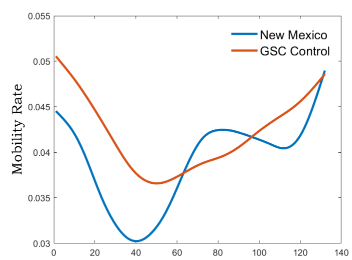
(f) New Hampshire



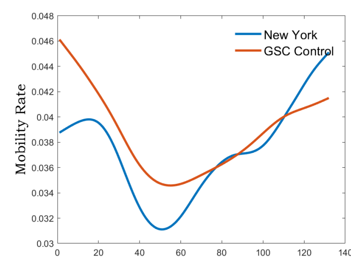
(g) New Jersey



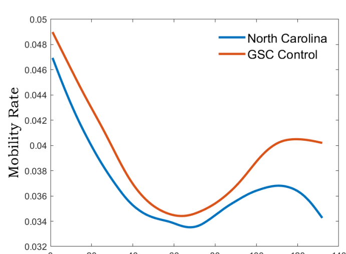
(h) New Mexico



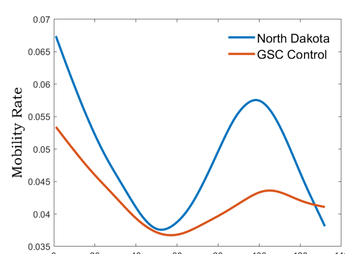
(i) New York



(j) North Carolina



(k) North Dakota



(l) Ohio

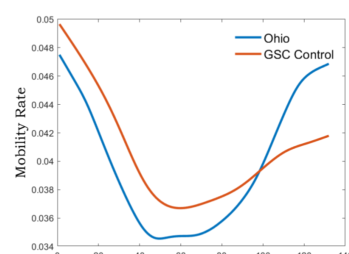
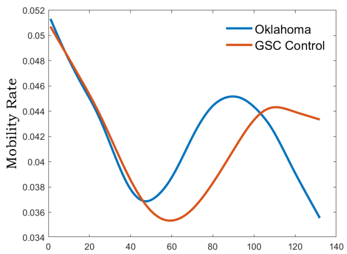
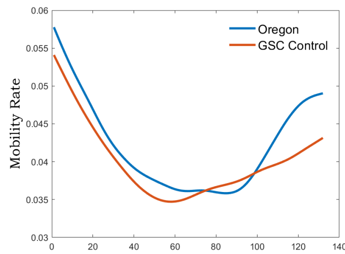


Figure E.4: GSC Fit

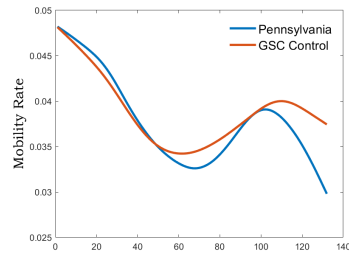
(a) Oklahoma



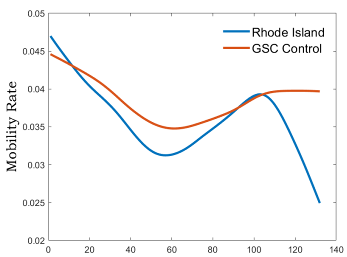
(b) Oregon



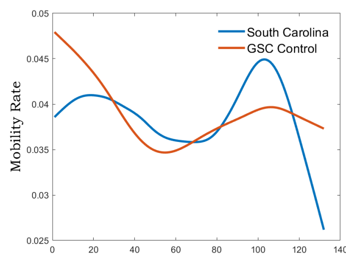
(c) Pennsylvania



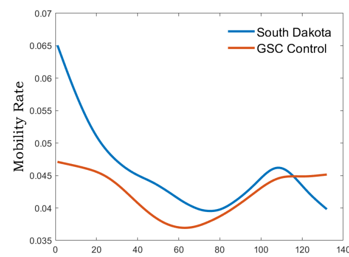
(d) Rhode Island



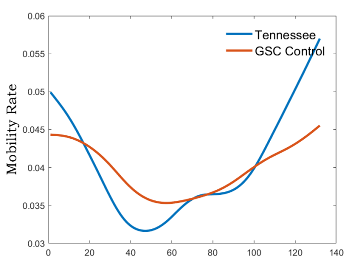
(e) South Carolina



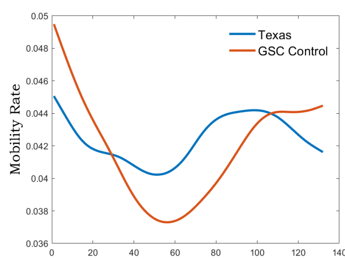
(f) South Dakota



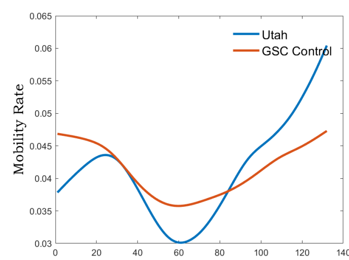
(g) Tennessee



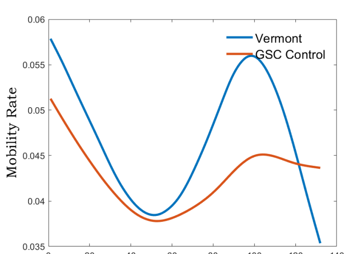
(h) Texas



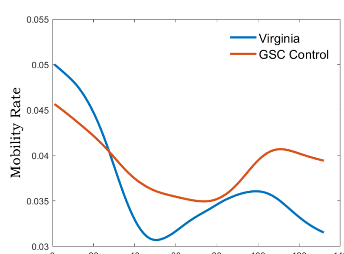
(i) Utah



(j) Vermont



(k) Virginia



(l) Washington

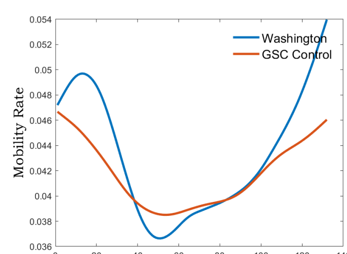
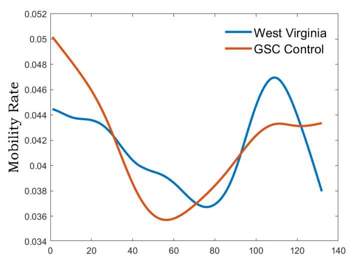
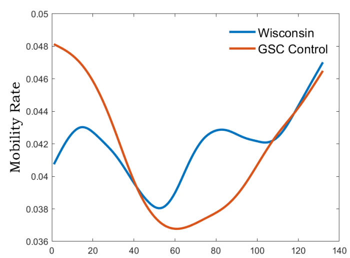


Figure E.5: GSC Fit

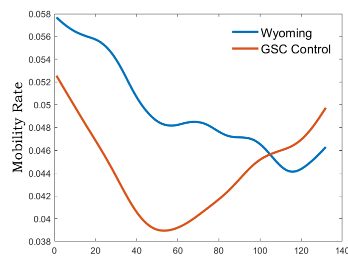
(a) West Virginia



(b) Wisconsin

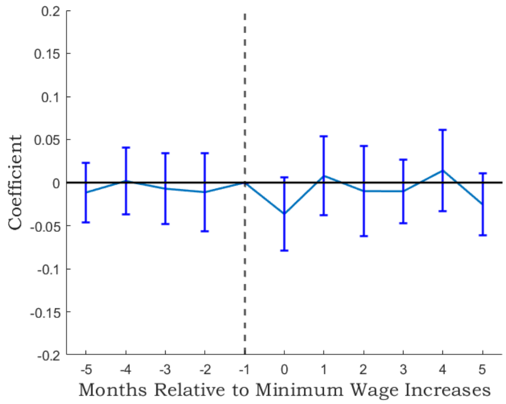


(c) Wyoming

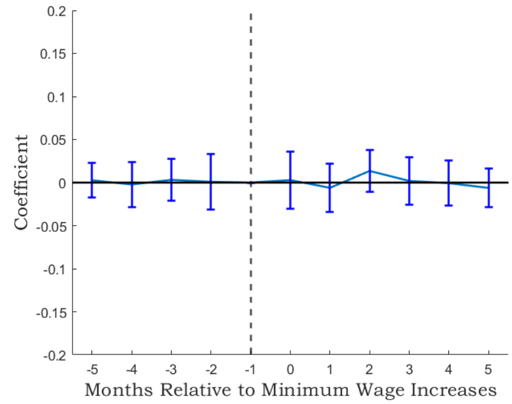


E.2 Dynamic Quasi-Elasticity of the Minimum Wage

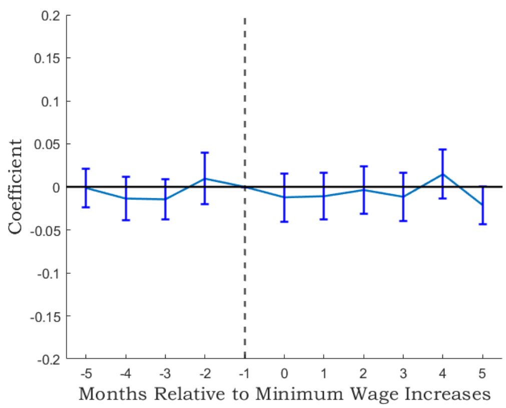
Figure E.6: Dynamic Quasi-Elasticity of the Minimum Wage for Subgroups of Workers



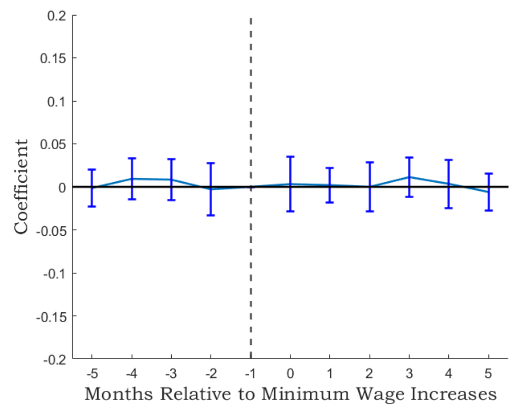
(a) Younger Workers



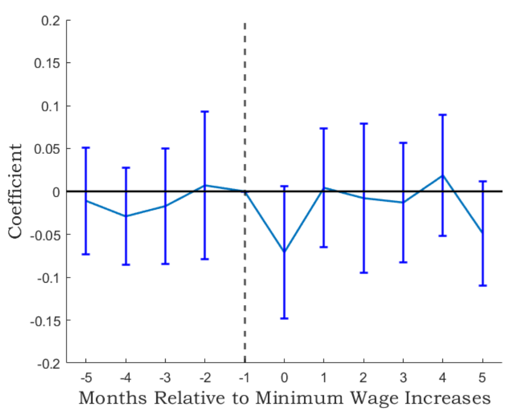
(b) Older Workers



(c) High-School Workers



(d) College Workers



(e) Young High-School Workers

(f) Young College Workers

F Micro-Found the Productivity Process

I micro-found the productivity process in section 3. Consider a Ben-Porath economy with human capital accumulation and labor supply decisions. Occupation-specific human capital determines productivity. Denote

$$\tilde{a} \equiv \frac{a}{1 + |a - j|}$$

To simplify, I assume that the match lasts for T periods and I normalized the outside option to be 0. One can think of T as a stopping time which does not affect the result in this section. The worker's objective is to minimize the disutility of labor supply and human capital accumulation, with initial human capital equal to h_1.

$$\min_{\{h_{t+1}\}} \sum_{t=1}^{T} \beta^t D(l_t, s_t) \tag{F.1}$$

The effective labor supply, l_t is decreasing in human capital h_t. A worker can accumulate human capital by exerting effort s_t. I assume that the effective labor supply is given by

$$l_t = g(h_t) \tag{F.2}$$

The function $g(h_t)$ is continuously differentiable and $g' < 0$. The interpretation is that human capital reduces labor disutility. Workers' effort leads to human capital growth subject to some shock ϵ_t.

$$h_{t+1} = h_t + f(\tilde{a}, \epsilon_t, s_t) \tag{F.3}$$

Equation (F.3) suggests that human capital accumulation depends on match-specific factor $\tilde{a}$, effort s_t, and some shock. I assume $\partial^2 f / \partial \epsilon_t \partial s_t > 0$, so that with good shocks, the same level of effort raises human capital by more. This is in line with the literature on the persistent effect of labor market entry shocks on wages, in which one explanation is that workers accumulate human capital faster with good entry shocks (see e.g. Oreopoulos et al. (2012)). In addition, I assume $\partial^2 f/\partial\tilde{a}\partial s_t > 0$, so the marginal gain in human capital by effort is increasing in match quality $\tilde{a}$. The shock ϵ has pdf $\phi(\epsilon)$.

The disutility function is separable in effective labor supply and effort.

$$D(l_t, s_t) = d_1(l_t) + d_2(s_t) \tag{F.4}$$

The recursive formulation of the problem is hence

$$V(h_t, \epsilon_t) = \min_{s_t} d_1(g(h_t)) + d_2(s_t) + \beta \int V(h_t + f(\tilde{a}, \epsilon_t, s_t), \epsilon_{t+1})\phi(\epsilon_{t+1})d\epsilon_{t+1} \tag{F.5}$$

The worker chooses the human capital accumulation effort s_t to trade off current disutility $d_2(s_t)$ with future labor disutility reduction. The first order condition reflects this trade-off:

$$d_2'(s_t) + \beta \int \frac{\partial V}{\partial h}\phi(\epsilon_{t+1})d\epsilon_{t+1} = 0 \tag{F.6}$$

This is an implicit function of s_t, which implies that the human capital increment is determined by $f(\tilde{a}, \epsilon_t, s_t(h_t, \epsilon_t))$. Note that the envelop condition gives

$$\frac{\partial V}{\partial h} = d_1'(g(h_t))g'(h_t) + \beta \int \frac{\partial V}{\partial h}\phi(\epsilon_{t+1})d\epsilon_{t+1} \tag{F.7}$$

With equation (F.7), it is easy to verify that $\partial^2 f/\partial h_t\partial\epsilon_t > 0$ and $\partial^2 f/\partial h_t\partial\tilde{a} > 0$. Using Taylor expansion, I have

$$f(\tilde{a}, \epsilon_t, h_t) \approx \beta_0 + \beta_1\tilde{a} + \beta_2\epsilon_t + \beta_3 h_t + \beta_4\tilde{a}h_t + \frac{\beta_5}{\sigma}\sigma\epsilon_t h_t \tag{F.8}$$

in which $\beta_4, \beta_5 > 0$. I also assume that shock per se does not affect human capital accumulation so that $\beta_2 = 0$.

Plug equation (F.8) into equation (F.3). With some normalization, I arrived at

$$h_{t+1} = h_t + (\tilde{a} + \sigma\epsilon_t)h_t \tag{F.9}$$

This is equivalent to

$$\frac{h_{t+1} - h_t}{h_t} = \tilde{a} + \sigma\epsilon_t \tag{F.10}$$

Equation (F.10) is the same as equation (10). The result suggests that in a Ben-Porath economy with endogenous human capital accumulation, the human capital process can evolve stochastically, depending on match-specific component $\tilde{a}$ and idiosyncratic shock ϵ.

G A Directed Search Model with Wage Posting

In this section I construct a simple directed search model with wage posting, and show that the minimum wage decreases occupational mobility and increases mismatch by the wage compression channel.

Consider the following one period model. There is a continuum of workers $i \in [0, 1]$ and two occupations. The workers are mismatched in occupation 1 and better matched in occupation 2. Mathematically, the output of workers in occupation 1, y_1 is less than the output in occupation 2, y_2.

The workers start in occupation 1 and can switch to occupation 2 subject to a switching cost ϕ. Their outside option is unemployment, with the value equal to U. The unemployment benefit is b. Let $\alpha_w(q)$ denote the probability that a worker finds a job. $q = u/v$ is the queue length. Let w_1 be the wage posted in occupation 1 and w_2 be the wage posted in occupation 2. Let q_1 be the queue length in occupation 1 and q_2 be the queue length in occupation 2. In equilibrium, we have

$$U = \alpha_w(q_1)w_1 + [1 - \alpha_w(q_1)]b = \alpha_w(q_2)w_2 + [1 - \alpha_w(q_2)]b - \phi \tag{G.1}$$

The firms choose to post vacancies in occupation 1 and 2 subject to the same flow cost of vacancy κ. Let $\alpha_e(q) = q\alpha_w(q)$ be the probability that a vacancy is filled. The firms' value functions are

$$J = \max_{w_j, q_j} -\kappa + \alpha_e(q_j)(y_j - w_j), \quad j = 1, 2 \tag{G.2}$$

Substituting equation (G.1) into equation (G.2), we have by the first order condition:

$$\begin{aligned} \alpha_e'(q_1) &= \frac{U - b}{y_1 - b} \\ \alpha_e'(q_2) &= \frac{U + \phi - b}{y_2 - b} \end{aligned} \tag{G.3}$$

Equation (G.3) implies that the difference between q_1 and q_2 depends on y_1, y_2, and U:

$$q_1 - q_2 = \alpha_e'^{-1}(\frac{U-b}{y_1-b}) - \alpha_e'^{-1}(\frac{U+\phi-b}{y_2-b}) \tag{G.4}$$

It is also easy to see that the wage is increasing in y because in equilibrium $w = b + \epsilon(q)(y-b)$. Now let us impose a minimum wage m which is between w_1 and w_2: $w_1 < m < w_2$. For the equilibrium to exist, the value of unemployment needs to increase. Now differentiating equation (G.4) with respect to U:

$$\frac{\partial q_1 - q_2}{\partial U} = \frac{1}{(y_1-b)\alpha_e''((U-b)/(y_1-b))} - \frac{1}{(y_2-b)\alpha_e''((U+\phi-b)/(y_2-b))} \tag{G.5}$$

Since $a_e(q)$ is concave, some algebra shows that $\partial q_1 - q_2/\partial U > 0$: more workers would stay in the mismatched occupation 1 and less workers would move to the better matched occupation 2. The increase in the minimum wage m and hence U leads to a decrease in switching from occupation 1 to 2 and an increase in mismatch. The example shows that a one period directed search model with wage posting is consistent with the implications of the model in section 3.