Optimal Minimum Wage Setting in a Federal System*

Andrew Simon
University of Michigan

Matthew Wilson
University of Michigan

May 6, 2020

Abstract

Minimum wages in the United States are jointly set by federal, state, and local governments, while many other countries have a single national policy. This paper provides a theoretical model and counterfactual simulations to understand the relative merits of centralized and decentralized policy setting. A binding policy is optimal if the benefits from redistribution outweigh the costs from migration, which are relatively steeper for local governments. Centralized policy, though uniform in practice, reduces horizontal migration externalities, which improves decentralized minimum wage setting. Our results therefore indicate that decentralized and centralized policy setting exhibit strategic complementarity; the extent of which depends importantly on mobility and regional heterogeneity. We then calibrate a model of the continental US and find that joint policy setting leads to a small welfare gain over centralization, and closely resembles the social planner's optimal policies.

1 Introduction

After Kansas City, and St. Louis, Missouri set their own minimum wages above the state level in 2017, the Missouri state legislature prohibited any city from setting its own policy. While there is likely a political motive for a red state to overturn a progressive policy in its blue cities, there are also potential economic benefits from centralized redistribution. The central government can more efficiently implement such policies when workers are mobile by internalizing spillovers, and a state may not want one of its cities to set a higher minimum wage because of the externality imposed on the rest of the state. However, if the state is constrained to setting a uniform policy, as in Oates' (1972) decentralization theorem, then local policy setting may be preferred. In this paper, we examine optimal minimum wage setting in a federation with mobile workers to understand the relative benefits of centralized, decentralized, and joint policy setting with interregional spillovers. Although our substantive focus is on the minimum wage, our framework has implications for a broad set of policies where local governments supplement federal decisions, like the EITC, income taxes,

*emails: Andrew Simon- arsimon@umich.edu, Matthew Wilson- mgwilson@umich.edu. We would like to thank seminar participants at the University of Michigan and the 14th Meeting of the Urban Economics Association, John Bound, Jim Hines, and Clémence Tricaud for their helpful comments.

and Medicaid. Understanding the appropriate level of government decision making is increasingly important as the US federal government has continued to shift more responsibility for these policies to states (Baicker et al. 2012).

Our work most directly builds on Lee and Saez (2012), who examine the role of the minimum wage in the competitive labor market of a single jurisdiction, where workers are implicitly immobile. They find that a binding policy is desirable, even when non-linear taxes are available, if the newly unemployed have zero surplus from working and the government values redistribution to low-wage workers. In their model, workers can migrate between low and high wage jobs, though they face different costs from working in each. When a worker's sector is fixed so that labor supply can only respond on the extensive margin, the minimum wage is second-best Pareto inefficient.[1] With only one jurisdiction, their framework is unable to consider the relative merits of having different levels of government set policy as well as the externalities that result from horizontal and vertical government competition. Vertical competition in a federation is particularly important for the minimum wage because it is a price floor policy; only the higher one matters. If the central government that is restricted to a uniform policy sets a binding minimum wage in the lower wage region, then it lowers the costs associated with horizontal competition.

We adapt the Lee and Saez (2012) model to a two-jurisdiction framework with mobile agents, regional governments, and a federal government to study the trade-off presented in the decentralization theorem. Workers are not mobile across sectors, but the high-skilled are mobile across jurisdictions.[2] Our goal is to understand the conditions under which setting a binding minimum wage is optimal for each type of government and analyze the welfare implications of different policy setting authorities. Local governments compete for high-skilled workers, which hinders their ability to redistribute through the minimum wage, but the central government is restricted to a uniform policy. Our stylized framework only includes two inputs for tractability, but the high-skilled can be thought of more broadly as mobile factors of production that may respond to increases in the minimum wage, like capital and firm location. In addition to the "tiered" US model of a federal uniform price floor in which states may "top off," or raise the floor, we consider federal uniform, decentralized, and federal non-uniform policy setting.

In the first part of the paper, we provide theoretical results to understand the extensive margin decision of implementing a binding minimum wage when governments maximize social welfare. The local governments will set a binding policy if the welfare loss from unemployment is zero on the margin, and the government values redistribution to low-skilled workers more than emigration. Our conditions are the same as in Lee and Saez, except governments are concerned with policy induced migration. The federal government only cares about migration insofar as it affects total output and moving costs, leading it to have stronger preferences for a minimum wage than local jurisdictions.[3]

[1] Gerritsen and Jacobs (2016) allow for educational investment in the Lee and Saez framework and find that a minimum wage is only optimal in the presence of non-linear taxes if the gains from more education outweigh higher unemployment.

[2] Blundell and MaCurdy (1999), and Lee and Saez (2012) both note that the extensive margin working decision where workers are not mobile across sectors is the most important and relevant in practice.

[3] We also discuss joint policy setting by the local and federal government as well as federal non-uniform policies

Since the theoretical model only gives predictions on the extensive margin, we proceed to calibrate a two region model to the aggregate US economy to illustrate how mobility and regional heterogeneity in productivity impact the relative benefits of centralization, decentralization, a hybrid system as in the US, and centralized non-uniform policy in general equilibrium. Higher levels of mobility increase the costs associated with horizontal competition, while higher levels of regional heterogeneity decrease the effectiveness of uniform central policy.

The key insight from the model is that different levels of government are strategic complements in policymaking and the extent of this complementarity depends on input mobility and regional heterogeneity in productivity. Centralized and decentralized policy setting alone rarely lead to the social planner's outcome. When jurisdictions are identical and inputs are mobile, centralized authority leads to greater social welfare while additionally allowing local governments to enact policy has no effect. However, when jurisdictions differ but inputs are immobile and therefore there are no interregional spillovers, decentralized authority is preferred. Simultaneously allowing the central government to also set a minimum wage in this case does not change the equilibrium. For the more realistic intermediate cases with heterogeneous regions and imperfectly mobile inputs, a hybrid system improves welfare since central uniform policy reduces horizontal competition and decentralized policy allows for a different minimum wage in each jurisdiction. Therefore, centralization raises the value of local government policy setting and decentralization raises the value of federal policy setting.

Regional heterogeneity in redistributive preferences may also drive differences in the minimum wage, as reflected in the Missouri example. Urban areas with high minimum wages tend to be both more productive and more progressive. We find that progressivity has a nonlinear effect on optimal policy since the minimum wage redistributes to the low-skilled from the high-skilled and the newly unemployed. A government with a very progressive social welfare function would not implement a binding minimum wage because of the negative effects of additional unemployment, while a government at the other extreme would not implement a binding policy because it does not value redistribution. We also document how local heterogeneity in preferences can lead to higher minimum wages. If one local government enacts a minimum wage because of its redistributive preferences, this can lead the other local government to set a binding policy, since a higher minimum wage in one region decreases the migration externality in the other.

To understand the welfare implications of tiered and centralized policy setting, we calibrate an economic geography model of the continental United States to match regional heterogeneity in productivity, high-skilled location decisions, employment, and the federal government's optimal policy. Consistent with our earlier findings, this model predicts that tiered policy setting yields a small welfare gain over centralization alone, as it allows the federal government to more effectively redistribute from high-skilled to low-skilled workers. Even though states are heterogeneous in productivity, and high-skilled workers are fairly mobile, we also find that tiered minimum wage

and several alternative models, such as rent-seeking governments, economy-wide aggregate production functions, and additional mobile factors of production, to understand how the model assumptions impact the sufficient conditions for a minimum wage.

setting closely approximates the social planner's optimal policies. Similar models have been used to study place based policies in general (Kline and Moretti 2014), corporate tax cuts (Suárez Serrato and Zidar 2016), and income taxes (Colas and Hutchinson 2020).

While we focus on the trade-offs of different minimum wage setting authorities, the previous normative literature focused on different tax systems. Stiglitz (1982), Allen (1987), and Guesnerie and Roberts (1987) find that the minimum wage is not optimal when non-linear taxes are available, however, Guesnerie and Roberts find they may be desirable under linear taxes. The literature has also considered more sophisticated minimum wage policies in a single jurisdiction where government competition is implicitly absent, like a graduated minimum wage tied to firm size (Danziger and Danziger, 2018), in-kind redistribution (Economedies and Moutos, 2017), monitoring job search (Broadway and Cuff, 2001), and bargaining power (Hungerbuhler and Lehmann, 2009).[4] Previous work on optimal tax that incorporates government competition in a federation (e.g. Wilson, 1982; Gordon, 1983; Hamilton and Pestieau, 2005; Gordon and Cullen, 2012; Divorkin, 2017) has not simultaneously allowed governments implement minimum wages. We connect these two previous literatures by understanding optimal minimum wage policies as a tool for redistribution in a federation.[5]

By putting the optimal minimum wage analysis into a federation, our work relates to the literature on tax competition. In our model, when a jurisdiction increases its minimum wage, high-skilled workers migrate to the other region, creating a horizontal externality that is increasing in mobility. Similar to models like Wilson (1986), and Zodrow and Mieszkowski (1986) on tax competition, agent mobility will lead to low local minimum wages. By adding a federal government to the Wilson-Zodrow-Mieszkowski set-up, Keen and Kotsogiannis (2002, 2004) find vertical externalities leave state taxes too high, which is in contrast with our result that federal policy decreases the cost of local minimum wage setting. Our results are therefore more similar to Janeba and Wilson (2011) who examine local public good provision with both horizontal and vertical externalities. In their model, local provision is too low due to competition, but central provision is inefficient because it is determined by a winning coalition in the legislature. In other political economy models, like Lockwood (2002) and Besley and Coate (2003), central provision is inefficient for similar reasons, whereas it is inefficient in our framework because the federal government is restricted to a uniform policy. Our assumption is based on the US system, but this restriction does not need to hold. In May 2019, European Commission Vice-President Frans Timmermans called for each member of the European Union to set a minimum wage of 60% of its median salary.

Although not centered on the minimum wage, there is a complementary literature on the interaction of mobility and redistribution, beginning with Stigler (1957) and Oates (1972). Our focus in this work is on the implications of different levels of government undertaking redistribution as

[4]We examine the implications in a competitive labor market, but other work has focused on the optimality when there are search frictions, as in search-and-matching model. See Flinn (2006), Dube et al. (2016)), and Lavecchia (2018), for example.

[5]There is also a large empirical literature on the effects of the minimum wage. See Neumark (2018) for a recent review.

in Oates (1977), Ladd and Doolittle (1982), Brown and Oates (1987), and Dixit and Londregan (1998). Our approach to understanding this mechanism is similar to Epple and Romer (1991), who also use a computational model to capture general equilibrium effects. The importance of mobility for policy making in a federation has also been studied in the context of public goods (Epple and Platt 1998; Caplan et al. 2000; Calabrese et al. 2012),[6] income redistribution (Wildasin 1997; Armenter and Ortega 2011; Gordon and Cullen 2012), and higher education (Wildasin 2000).[7]

In the following section, (Section 2), we present the theoretical model and some extensions. Based on theoretical work, Section 3 outlines our two region quantitative model and calibration, and Section 4 solves for the optimal minimum wages for different government structures, regional heterogeneity, and worker mobility. To asses the welfare implication of different minimum wage setting authorities, we calibrate a model of the United States in Section 5. Section 6 concludes.

2 Model

We consider a simple two-region, two-factor model to emphasize the role of regional heterogeneity and factor mobility in determining the relative effectiveness of centralized and decentralized minimum wage setting. The economy, I, consists of two regions, $i \in \{1, 2\}$, low-skilled and high-skilled workers, local firms, and local and federal governments. We consider each in turn.

Each region i has a unit mass of immobile low-skilled labor. Low skilled labor may either be employed, l, or unemployed, u, so that $u_i + l_i = 1$. If employed, they pay a cost θ of working, which is distributed according to a known distribution $\mu(\theta)$. There are $\bar{h}$ high-skilled workers who are (im)perfectly mobile across regions, so that $h_1 + h_2 = \bar{h}$. High-skilled workers face moving costs of $\xi \sim \zeta(\xi)$, but do not face a working cost.[8] We interpret this as a normalization on working costs so that the low-skilled costs are relative to those of the high-skilled. It also ensures that all high-skilled workers are in the labor force for any reasonable minimum wage policy. Workers can only be employed in their region of residence. Consumption, c_j, for each worker j is given by her wage minus her cost of working and moving costs paid, if applicable.

Our mobility assumptions are based on the high correlation between education level and mobility, and the fact that low-skilled workers sluggishly respond to employment shocks.[9] To better understand the implications of these restrictions on workers, we consider how low-skilled mobility affects our theoretical results in the next section. Then, in section 4, our calibration exercise varies $\zeta(\xi)$ to describe the relationship between the high-skilled migration elasticity and optimal minimum wage setting authority. Although the model and calibration are based solely on low and high-skilled workers for tractability, the high-skilled represent input factor mobility more broadly. For example,

[6] There is also a relatively large literature on the value of centralized environmental policies with spillovers. See Banzhaf and Chupp (2012) for a recent example.

[7] There is also a growing empirical literature on income taxation and mobility. See Kleven et al. 2019 for a review.

[8] In our simplified framework, we abstract away from congestion costs, which are likely important for high-skilled migration (Moretti 2013).

[9] Amior (2019) documents this relationship using the CPS. Our assumptions are similar to those in Wildasin (2000) who examines the public provision of human capital in a federation, but we allow for variable high-skilled mobility to understand the comparative statics. We discuss this assumption more explicitly later in this section.

firms may respond to a minimum wage in their capital investment or location decisions and local governments compete for these resources. Our choice of using the high-skilled as the mobile factor allows us to naturally incorporate the mobile input into a social welfare calculation. We discuss how additional mobile inputs affect our results later in this section.

Each region has a single perfectly competitive firm with production function $\gamma_i f(l,h)$. Without loss of generality, we assume $\gamma_1 = 1$ and $\gamma_2 = \gamma \leq 1$. Workers are paid their marginal product, and if the high-skilled are perfectly mobile, they migrate so that $w_1^h = w_2^h$. Relative low-skilled competitive wages across regions depends on γ and the relative stock of high-skilled workers. We assume that the optimal minimum wage is never binding for high-skilled workers, so that they are also always employed. Let Θ_{ie} be the set of θ's of the low-skilled employed workers in i and Θ_{iu} the unemployed. For a given wage, workers with relatively large values of θ are unemployed because they choose to exit the labor force. If there is a binding minimum wage, others become unemployed due to the change in low-skilled labor demand by the firm. Similarly, define Θ_{ih} as the set of high-skilled workers in i.

There is one central government, and each region has its own local government. Both the local and central governments have perfect information about workers' working and moving costs and regional productivity, and we explore the consequences of this assumption in Section 4.3. Governments care about the consumption of their constituents, according to the concave function $G(c)$, where $G' > 0, G'' < 0$, and $G'(0) < \infty$. Total utility in r is given by $\int_{j \in r} G(c_j) dj$, $r = \{1, 2, I\}$, i.e. each region maximizes the total utility of its own residents, while the central government maximizes the utility of both regions. We assume that $G(c_h) > 0$ for any feasible policy so that high-skilled workers are valued and local governments compete for them. Federal and local governments are either endowed or not with the ability to set a minimum wage. There are four potential authority structures: (1) only regional governments; (2) only central government uniform; (3) both central uniform and regional governments; and (4) central non-uniform. Governments do not need to raise revenue.

Equilibrium is determined in two stages. In the first stage, no government has minimum wage setting authority; this is the competitive equilibrium. This gives each worker a residence. In the second stage, some governments are endowed with minimum wage setting power and optimally set their policy, given the policies of other governments, if applicable, and the responses by firms and workers.

Definition 1. *The economy is in equilibrium after the second stage for a given authority structure if:*

1. *Each government with minimum wage setting authority optimally sets its minimum wage to maximize the welfare of its residents, given the policies of the other governments with authority.*
2. *Firms maximize profits, taking the residencies of workers and government policies as given.*
3. *Each worker optimally chooses whether to enter the labor force to maximize utility.*

4. No high skilled workers want to move.

2.1 Decentralized Minimum Wage

We consider the conditions under which it is optimal for local governments to set a binding minimum wage. Each jurisdiction i's social welfare function is defined as

$$SW_i = (1 - l_i)G(c_u) + \int_{\Theta_e} \mu(\theta)G(c_l)d\theta + \int_{\xi(\Theta_{ih})} \zeta(\xi)G(c_{h_i})d\xi \tag{1}$$

where $c_j = w_j - \theta_j$ for low-skilled employed individual of type θ_j, $c_h = w_h$ for high-skilled individuals, and c_u is the consumption of the unemployed. Without taxes and transfers, $c_u = 0$. Since moving costs are only paid if residents leave, they do not directly enter the regional government's social welfare function. As in Lee and Saez (2012), let $g_u = G'(c_0)/\lambda; g_l = \int_{\Theta_e} \mu(\theta)G'(c_l)d\theta/\lambda$ and $g_h = G'(c_h)/\lambda$, where λ is the Lagrange multiplier on the Social Planner's budget constraint.[10]

For a given level of the minimum wage in the other jurisdiction, $\bar{w}_{-i}$, i implements a binding minimum wage $\bar{w}_i$ if and only if:

$$0 < \frac{\partial SW_i}{\partial \bar{w}_i}\Big|_{\bar{w}_{-i}} \iff \tag{2}$$

$$\begin{aligned} 0 < {} & \frac{-\partial l_i}{\partial \bar{w}_i}G(0) + \frac{\partial l_i}{\partial \bar{w}_i}\int_{\Delta_{\bar{w}_i}\Theta_{iu}} \mu(\theta)G(c_l)d\theta \\ & + \int_{\Theta_{ie}(\bar{w}_i)} \mu(\theta)G'(c_l)d\theta\frac{\partial w_i^l}{\partial \bar{w}_i} + h_iG'(c_h)\frac{\partial w_i^h}{\partial \bar{w}_i} \\ & + \frac{\partial h_i}{\partial \bar{w}_i}G(c_h) \end{aligned}$$

The first line represents the change in welfare induced by newly unemployed low-skilled workers. $\Delta_{\bar{w}_i}\Theta_{iu}$ denotes the set of workers who become unemployed due to the minimum wage $\bar{w}_i$, which depends on how workers are separated from the firm after the policy change. Following Lee and Saez (2012), we assume efficient rationing, i.e. workers with the lowest surplus from working involuntarily lose their jobs first. Under the assumption, $\Delta_{\bar{w}_i}\Theta_{iu} = [\gamma_i f_l(l,h), \theta(\bar{w})]$, where $\theta(\bar{w})$ is the largest θ worker who becomes unemployed. $\theta(\bar{w})$ depends on the new policy level, the relative production functions in each jurisdiction, $-i$'s minimum wage, and the amount of migration. For a very small binding policy, $\theta(\bar{w}) = \gamma_i f_l(l,h)$. Therefore, the marginal low-skilled worker who loses her job has $c_j = 0$ before and after the policy change. The second line represents the change in welfare from the still-employed in region i and the third line is the change in welfare due to emigration.

First note that $\frac{\partial w_i^l}{\partial \bar{w}_i} > 0$ by construction and $\frac{\partial l_i}{\partial \bar{w}_i} < 0$ since it is less profitable to employ

[10]The Social Planner's budget constraint is $(1 - l_i)c_u + l_ic_l + h_ic_h \leq l_iw_l + h_iw_h$. Since the governments in our model do not redistribute and only set a minimum wage, they do not have a budget constraint. Also note that the construction of the g terms are analogous for the regional and central governments, although their respective populations are different.

low-skilled workers at a higher wage. Increasing the minimum wage also impacts the high-skilled workers since $\frac{\partial w_i^h}{\partial \bar{w}_i} < 0$ from the zero-profit condition.[11] As the marginal product of labor for high-skilled workers falls, some high-skilled workers leave, giving $\frac{\partial h_i}{\partial \bar{w}_i} < 0$ as well. For this reason, the third line of the first order condition only involves emigration and not immigration. Since the local government only cares about its residents, i.e. natives less those who move out, it does not internalize individual migration costs.

Proposition 1. *If (1) there is efficient rationing; (2) the demand elasticity is finite; (3) the supply elasticity of low-skilled workers is positive; (4) the government values additional wages to low-skilled workers more than the loss of wages to high-skilled (redistribution) and outmigration, then it is total welfare improving for a jurisdiction to impose a (small) binding minimum wage.*

Proof. Our assumptions imply that the first two terms sum to 0 and the magnitude of the third term is larger than the sum of the magnitude of the fourth and fifth terms. More explicitly, using the above expressions and the zero profit condition, we can rewrite the last three terms as $l_i\lambda[g_l - g_h] + \frac{\partial h_i}{\partial \bar{w}_i} G(c_h)$. This is identical to proposition 1 in Lee and Saez (2012) except jurisdictions are now also affected by potential outmigration, weakening the incentive to enact a binding minimum wage. □

Under these assumptions, this extensive margin decision to adopt a minimum wage depends on γ, $\bar{w}_{-i}$, the elasticity of substitution between low and high-skilled workers, and the mobility of the high-skilled. In section 3.2, we conduct a calibration exercise to better understand how these parameters impact equilibrium, on both the extensive and intensive margins.

We now consider how changes to the assumptions of Proposition 1 affect the desirability of a binding minimum wage. Without efficient rationing, the first two terms do not sum to zero for a small increase in the minimum wage because workers with a large surplus from working lose their jobs. If workers were fired randomly after the increase in the minimum wage, then social welfare would be lower in expectation as workers with positive value of employment are separated rather than the indifferent marginal employee. However, even if the lowest cost workers were separated first, a government may still implement a binding minimum wage if the value of redistribution is sufficiently large.[12] This reasoning leads to a corollary to Proposition 1.

Corollary 1. *Efficient rationing leads local governments to weakly prefer the highest minimum wages, all else equal, relative to any other separation assumption.*

Proof. Efficient rationing maximizes the sum of the first two terms, while not affecting any other terms. Otherwise, there is a first order welfare loss from unemployment. The condition is weak

[11] With perfectly competitive markets and constant returns to scale, the zero-profit condition implies that $ldw^l + hdw^h = 0$, giving the relationship between the minimum wage and high skilled wages. See Lee and Saez (2012) Appendix A.1. for more information.

[12] Lee and Saez (2012) note that while this is the most favorable assumption for optimal policy, it may not be realistic or may be costly to reach through queuing or search costs. However, there is some empirical literature supporting this assumption. Neumark and Wascher (2007) find that the minimum wage has larger impacts on teenagers and secondary earners who are more likely to have more elastic labor supply and Luttmer (2007) found that reservations wages do not increase with the minimum wage.

because a government may not desire a binding minimum wage for any separation policy if the other conditions of Proposition 1 do not hold. □

If conditions (2) or (3) do not hold, then the employment of low-skilled workers will change too dramatically for the first two terms to sum to zero. Condition (4) may fail if governments care too much about outmigration relative to redistribution, if they do not care about redistribution, or if they only care about redistribution to the unemployed. As a special case, consider when high and low-skilled labor are perfect substitutes. Then $\frac{\partial w_i^h}{\partial \bar{w}_i} = 0$ and there is no outmigration. In this case, the minimum wage is optimal if there is efficient rationing as the government trades off the unemployment of marginal workers with the additional earnings of all others. As the two types of labor become less substitutable, increasing the minimum wage will lead to lower high-skilled wages and more outmigration. Decreasing the elasticity of substitution will weaken the incentive to have a minimum wage, all else equal.

To further understand the migration externality, consider the impact of raising the minimum wage on the other jurisdiction. If high-skilled workers migrate from jurisdiction i because moving costs are sufficiently small on the margin, the marginal product of labor of high-skilled workers in jurisdiction $-i$ will fall. Government $-i$ is made relatively better off, all else equal, as there is a small change in high-skilled consumption, but a larger change in welfare due to the additional high-skilled workers. This effect lowers the cost for $-i$ to implement a binding minimum wage as fewer high-skilled workers will want to migrate compared to the case where i does not have a binding minimum wage. In our model, the marginal product of low-skilled workers increases when high-skilled workers immigrate, and wages and employment both increase, leading unambiguously to higher utility for those workers.

2.1.1 Mobility and the Minimum Wage

Throughout the model we assume that low-skilled workers are perfectly immobile for tractability. While these workers tend to have relatively low mobility rates overall, Monras (2019) finds that after a minimum wage increase, the relative share of low-skilled workers decreases because of changes in low-skilled in-migration. This effect is potentially driven by immigration from other countries. Cadena (2014) provides evidence that low-skilled immigrants prefer states with unchanged minimum wages relative to those experiencing increases because of the disemployment effects. Since our static framework considers a fixed point in time without population change, including from immigration, or shocks to workers that induce migration besides directly from the minimum wage, this concern is not particularly relevant in our highly stylized setting. Monras additionally finds no effect on out-migration. Taken together, these effects on migration are consistent with moving low-skilled not choosing to live in minimum wage areas, rather than migration due to changes in the minimum wage itself. In a dynamic setting where households may move at any given time for any reason, the immobility assumption becomes less plausible and more important.

We now consider the implications of relaxing the perfect immobility assumption for complete-

ness. Low-skilled workers initially sort in stage 1 so that w^l is equal across jurisdictions. After a minimum wage increase in i, the firm does not hire any additional low-skilled and some are separated. Therefore, no low-skilled in $-i$ have an incentive to immigrate to i after the policy change. The unemployed in i emigrate if the new wage is higher than the sum of working and moving costs. After a minimum wage increase, low-skilled employment weakly increases in $-i$ from high-skilled migration, but the model does not specify who is hired. If the firm in $-i$ only hires initially local low-skilled workers, then there is no incentive to move and the immobility assumption does not impact our results.[13] If the firm hires workers with the lowest working cost, then the incentive to move depends on γ. Absent any binding minimum wage policies, the marginal worker in jurisdiction A will have a weakly higher working cost than the marginal worker in B since $\gamma \leq 1$. Therefore, a minimum wage in A will not lead to low-skilled migration, but one in B will if moving costs are sufficiently small. In general, adding in low-skilled mobility has an ambiguous effect on the optimal minimum wage policy that depends on the value of $G(0)$ and whether the newly unemployed would be hired if they moved.[14]

The main migration mechanism in the model is the link between regions created from high-skilled. The extent to which high-skilled workers respond to local minimum wage policies, $\frac{\partial h_i}{\partial \bar{w}_i}$, determines the effectiveness of local government policy relative to centralization. If the derivative is zero, because moving costs are high or the marginal product of low and high-skilled workers are unrelated, then our framework reduces to the single region case as studied by Lee and Saez (2012). Monras (2019) estimates the relationship between wages and high-skilled out-migration, finding negative or zero effects; however, they are not statistically different from zero. Cadena (2014) also provides some suggestive, although mixed, evidence that this derivative is in fact negative. Taking the point estimates across all specifications at face value, he finds a negative relationship between the count of immigrants with at least some college and the minimum wage in a given state.[15] More generally, the high-skilled migration externality represents the impact on endogenous business locations of low-skilled employers which have been shown to be somewhat responsive to the minimum wage (Rohlin 2011 and Aaronson et al. 2018). Since our model treats the low and high-skilled as linked through production, we think of the high-skilled as the managers of those firms, with the very high-skilled not captured in our framework.

2.2 Extensions of the Decentralized Model

In our baseline model, the results and conditions for a binding optimal minimum wage are driven by the high-skilled migration externality: when a local government raises its minimum wage some high-skilled workers move out. The size of the externality depends on the choice of social welfare function

[13] Phillips (2018) finds that low wage employers discriminate against applicants with a long commute, which provides suggestive evidence for this assumption.

[14] If $G(0) < 0$, then local governments can set high minimum wages to induce the unemployed to migrate and raise total social welfare in the jurisdiction.

[15] When controlling for state-specific trends in the main results, the point estimate becomes small in magnitude and imprecise. The unweighted results are slightly larger in magnitude, but still imprecise. Although he interprets this as the minimum wage having no effect on high-skilled migration, the result is likely underpowered.

and the production technology. We extend the model to see how changes in these assumptions alter the conclusions. We consider a social welfare function where governments only care about natives or maximize housing values, when there is a single firm that operates in both regions, and the introduction of additional inputs to production.

2.2.1 Social Welfare Function

In our framework, we assume that the government only cares about the total utility of its final residents in the second stage equilibrium. In the first part of this section, we consider a government that only cares about natives, regardless of whether they move.[16] When governments maximize utility of their natives, the derivative of the social welfare function shown in equation 2 no longer includes the final term: $\frac{\partial h_i}{\partial \bar{w}_i} G(c_h)$. However, they internalize migration costs. The government still trades off lower consumption of high-skilled workers with higher consumption of the low-skilled through the minimum wage, but no longer faces a first order cost from emigration. The implied minimum wages are higher compared to our baseline model.

Proposition 2. *If conditions 1-3 from Proposition 1 hold, the government cares only about their initial residents, and it values additional wages to low-skilled workers more than the loss of wages to high-skilled (redistribution) and moving costs, then it is total welfare improving for a jurisdiction to impose a (small) binding minimum wage.*

Proof. This is immediate from the first order condition, where the last term now concerns moving costs and not outmigration. Under this new assumption on the social welfare function, h_i is fixed.

$$0 < \frac{-\partial l_i}{\partial \bar{w}_i} G(0) + \frac{\partial l_i}{\partial \bar{w}_i} \int_{\Delta_{\bar{w}_i} \Theta_{iu}} \mu(\theta) G(c_l) d\theta$$
$$+ \int_{\Theta_{ie}(\bar{w}_i)} \mu(\theta) G'(c_l) d\theta \frac{\partial w_i^l}{\partial \bar{w}_i} + \int_{\xi(\Theta_{ih})} G'(c_h) \left(\frac{\partial w_i^h}{\partial \bar{w}_i} - \frac{\partial \xi}{\partial \bar{w}_i} \right) d\xi$$

□

Corollary 2. *Local governments that care about their initial residents set weakly higher minimum wage policies than those that care about their final residents.*

Proof. Comparing the first order conditions for a small binding minimum wage, this requires $\frac{\partial h_i}{\partial \bar{w}_i} G(c_h) \leq -\int_{\xi(\Theta_{ih})} G'(c_h) \frac{\partial \xi}{\partial \bar{w}_i} d\xi$. By construction, $\frac{\partial \xi}{\partial \bar{w}_i} = \xi$ if the worker moves and 0 otherwise, allowing us to rewrite the inequality as $\frac{\partial h_i}{\partial \bar{w}_i} G(c_h) \leq -\frac{\partial h_i}{\partial \bar{w}_i} G'(c_h) \mathbb{E}_{\Delta\xi(\Theta_{ih})}[\xi]$. First consider when $-i$ does not have a binding minimum wage. Then $\frac{\partial h_i}{\partial \bar{w}_i} < 0$, and the inequality holds since $G(c_h), G'(\cdot), \xi > 0$ by assumption. If $-i$ has a binding minimum wage and its initial residents return after i puts on a binding minimum wage, then both sides of the inequality are 0, and it holds as well. If i residents migrate instead, then the argument from the first case applies. □

[16] See Wilson (2015) for an overview of different assumptions on government maximization with resident mobility and Cremer and Pestieau (2004) for a review of the literature on government maximization with factor mobility.

The preceding analysis assumed that governments are utility maximizers, but we now consider a model where rent-seeking governments maximize property values.[17] In the most general application of this approach to our framework, all households and firms demand land. Define $H^d(\bar{w})$ to be the housing or land demand under minimum wage policy $\bar{w}$ and $\Delta_j^{j'} H^d(\bar{w})$ be the change for workers who transition from type j to j' due to the policy change, or for firm f. We assume that housing demand is nondecreasing in income, and that the elasticity of housing supply is finite. Then, for jurisdiction i, a binding minimum wage is optimal if and only if

$$\frac{\partial H^d}{\partial \bar{w}} = \Delta_u^u H^d(\bar{w}) + \Delta_l^u H^d(\bar{w}) + \Delta_l^l H^d(\bar{w}) + \Delta_{h_s}^{h_s} H^d(\bar{w}) + \Delta_{h_s}^{h_{s'}} H^d(\bar{w}) + \Delta_f H^d(\bar{w}) > 0 \qquad (3)$$

It is immediate that $\Delta_l^l H^d(\bar{w}) \geq 0$. Workers who remain employed after the wage increase will have more income and demand more housing. Similarly, workers that lose their jobs will demand weakly less housing, giving $\Delta_l^u H^d(\bar{w}) \leq 0$. The magnitude of this term importantly depends on how we treat the working cost. If θ is a real cost, then for a small binding minimum wage, the workers who lose their jobs under efficient rationing will have zero net income both before and after the policy change. If it is a utility cost, they will have less income and demand strictly less housing. In the welfare maximization framework, this distinction did not affect optimal policy. In the case where θ is a utility cost, the minimum wage is less attractive because we no longer have an assumption akin to efficient rationing. The always unemployed also have no change in income, implying $\Delta_u^u H^d(\bar{w}) = 0$.

Similar to our analysis in the utility maximization case, there is weakly less demand from high-skilled workers when the minimum wage increases. $\Delta_{h_s}^{h_s} H^d(\bar{w}) \leq 0$ since their wages weakly decrease and $\Delta_{h_s}^{h_{s'}} H^d(\bar{w}) \leq 0$ due to outmigration from falling wages. The changes are zero in the extreme case where low and high-skilled labor are perfect substitutes.

Finally, consider firms' demand for land. We can decompose the effect on firms as: $\Delta_f H^d(\bar{w}) = \frac{\partial H_f^d}{\partial L(l,h)} \frac{\partial L(l,h)}{\partial \bar{w}}$, where $L(l,h)$ is the total amount of land used by low and high-skilled workers for production. We showed that $\frac{\partial L(l,h)}{\partial \bar{w}} < 0$ as low-skilled workers become unemployed and the high-skilled migrate. If we assume that more workers requires more land, then $\frac{\partial H_f^d}{\partial L(l,h)} > 0$, and firm demand for land decreases. However, if the firm engages in more land-intensive activity when it has fewer workers, then $\frac{\partial H_f^d}{\partial L(l,h)} < 0$, and firm demand for land increases.[18] Summarizing the above analysis gives us the following proposition.

Proposition 3. *A rent-seeking local government will implement a small binding minimum wage if conditions 1-3 from Proposition 1 hold and if the housing supply elasticity is finite, housing demand is nondecreasing in after working-cost net income, and the net change in housing demand from the low-skilled employed and firms is positive and larger than the decrease in demand from the high-skilled.*

[17] See Epple and Nechyba (2004) for a survey of these two different assumptions on the government's objective function.

[18] In Section 2.2.3 we derive the general effect of additional inputs on the optimal minimum wage.

Proof. Under the assumptions, the condition in equation 3 is satisfied as described above since the first two terms are 0, and the sum of the remaining are positive. Note that the final term can be either positive or negative. □

The optimal minimum wage trades off more housing demand from the still employed low-skilled with lower demand from the high-skilled due to wage decreases and outmigration. We discussed a comparable set of assumptions to the welfare maximizing case, (i.e. efficient rationing with no change in housing demand for the newly unemployed, and no land demand by firms), that leads to similar policies. We specify utility maximization in our main theoretical and quantitative analysis because the results are theoretically similar and it does not require us to also specify a specific housing market system. This allows us to focus our analysis on the high-skilled migration externality.

2.2.2 Aggregate Production Function

When each jurisdiction has its own production function, asymmetric minimum wage policies lead high-skilled workers to migrate as the policies affect their relative marginal products across locations. If there is a binding minimum wage, employed low-skilled workers are paid that wage, but the amount of unemployment depends on the features of the production function. To understand the implications of this assumption, we consider a model where there is only one aggregate production function in the economy. This is motivated by the importance of large firms that operate in multiple states and choose where to hire their labor. For a company that operates in many states, the marginal product of labor of any worker may not only depend on the number of workers in her place of work, but in the entire company.

Proposition 4. *If conditions 1-3 from Proposition 1 hold, there is a single aggregate production function, and the government values additional wages to low-skilled workers more than the loss of wages to high-skilled (redistribution), then it is total welfare improving for a jurisdiction to impose a (small) binding minimum wage.*

Proof. With one aggregate production function, the marginal product of high-skilled workers will be the same in both jurisdictions, regardless of the minimum wage policies. High-skilled workers have no desire to migrate and the final term in equation 2 equals zero[19]. We are then in the Lee and Saez (2012) model where workers' locations and skill sectors are fixed. □

Proposition 4 replaces the outmigration condition with the single production function assumption. All else equal, this leads to higher minimum wages than in the baseline model. High-skilled workers are perfectly inelastic to all governments and now bear more of the burden of minimum wage policies, allowing the government to engage in more redistribution to the low-skilled. Since low-skilled workers are also always paid their marginal product, and marginal products do not differ across locations, when a jurisdiction raises its minimum wage, some low-skilled workers will become unemployed in each jurisdiction due to efficient rationing. Under our assumptions, each

[19]Note that this also implies that there is no difference between a government that cares about natives or residents.

jurisdiction has low-skilled employed workers earning zero surplus from working in the competitive equilibrium. This negative externality imposed on the other jurisdiction will lead minimum wage policies to be too high from an efficiency perspective as governments engage in behavior akin to tax exporting.

2.2.3 Capital and Other Inputs

With two inputs to production and zero-profits, high-skilled wages decrease with the minimum wage. We introduce elastically supplied capital, k, with required return, $\bar{r}$, so that this is not necessarily the case. With three production inputs, the zero-profit condition implies $\frac{dw^h}{d\bar{w}} = \frac{-l}{h} - \frac{k}{h}\frac{d\bar{r}}{d\bar{w}}$. If $\frac{d\bar{r}}{d\bar{w}}$ is relatively large in magnitude and negative, the marginal product of high-skilled workers increases with the minimum wage. The increased marginal product will induce migration into the jurisdiction as high-skilled workers move for higher wages. In this case, it is Pareto efficient[20] to implement a very small binding minimum wage since the marginal low-skilled worker is indifferent between working and not due to efficient rationing, and the wages of all other still employed low-skilled and all high-skilled workers increase. Raising the minimum wage further will not be Pareto efficient, and the government will trade off additional consumption from the employed low-skilled and high-skilled, and migration into the region from high-skilled workers with reduced consumption from the newly unemployed.

From the zero-profit condition, we can generalize $\frac{dw^h}{d\bar{w}}$ when there are additional inputs, indexed by ι with price p^ι. Then,

$$\frac{dw^h}{d\bar{w}} = \frac{-l}{h} - \sum_{\iota \neq h,l} \frac{\iota}{h}\frac{dp^\iota}{d\bar{w}} \tag{4}$$

Proposition 5. *When there are additional inputs to production besides high and low-skilled labor, a small binding minimum wage is Pareto Efficient if conditions 1-3 from Proposition 1 hold and* $\frac{dw^h}{d\bar{w}} \geq 0$

Proof. When $\frac{dw^h}{d\bar{w}} \geq 0$, there is no outmigration, and by efficient rationing all workers are weakly better off. □

2.3 Alternative Configurations for Minimum Wage Setting Authority

Using our baseline model, we consider the implications of alternative minimum wage setting authorities. As in the decentralized case, we develop conditions under which each type of government finds it optimal to implement a binding minimum wage.

[20] We are implicitly assuming that capital is paid its marginal product and always earns zero profit so we do not need to worry about rents to capital owners.

2.3.1 Only Federal Government

In this section, we consider optimal minimum wage setting when only the federal government can set a minimum wage. The government's social welfare function is defined as

$$SW_f = \sum_i (1 - l_i)G(0) + \int_{\Theta_{ie}} \mu(\theta)G(c_l)d\theta + \int_{\xi(\Theta_{ih})} \zeta(\xi)G(c_h)d\xi \tag{5}$$

Note that in the federal government's optimization problem, it cares about the utility of all high skilled workers regardless of residence. Unlike in the regional government's problem, when high skilled workers move, their utility still matters for the federal government. The government also cares about the migration costs paid by individual workers.[21]

2.3.2 Federal Non-Uniform Policy

When the federal government can set a different policy in each region, the policies are given by:

$$(\bar{w}_1, \bar{w}_2) = \arg\max SW_f(\bar{w}_1, \bar{w}_2) \tag{6}$$

The FOC for $\bar{w}_1$ imply that it will be binding in jurisdiction 1 if:

$$\begin{aligned} 0 < & -\frac{\partial l_1}{\partial \bar{w}_1} G(0) + \frac{\partial l_1}{\partial \bar{w}_1} \int_{\Delta_{\bar{w}_1}\Theta_{1u}} \mu(\theta)G(c_l)d\theta \\ & - \frac{\partial l_2}{\partial \bar{w}_1} G(0) + \frac{\partial l_2}{\partial \bar{w}_1} \int_{\Delta_{\bar{w}_1}\Theta_{2u}} \mu(\theta)G(c_l)d\theta \\ & + \int_{\Theta_{1e}(\bar{w}_1)} \mu(\theta)G'(c_l^1)\frac{\partial w_l^1}{\partial \bar{w}_1} + \int_{\Theta_{2e}(\bar{w}_1)} \mu(\theta)G'(c_l^2)\frac{\partial w_l^2}{\partial \bar{w}_1} \\ & + G'(c_h)\left(\bar{h}\frac{\partial w^h}{\partial \bar{w}_1} - \int_{\xi(\Theta_{ih})} \frac{\partial \xi}{\partial \bar{w}_1}\zeta(\xi)d\xi\right) \end{aligned}$$

The expression above is simplified by noting that $\frac{\partial h_1}{\partial \bar{w}_1} = -\frac{\partial h_2}{\partial \bar{w}_1}; h_1 + h_2 = \bar{h};$ and $\frac{\partial w_h^1}{\partial \bar{w}_1} = \frac{\partial w_h^2}{\partial \bar{w}_1}$.[22] We also use the fact that moving costs are randomly assigned after households choose a residence in the competitive equilibrium and that under our assumptions, there will be net migration to jurisdiction 2. The FOC for $\bar{w}_2$ is defined analogously.

The first row of the FOC gives the welfare loss from newly unemployed low-skilled workers in jurisdiction 1. The second row gives the welfare gain from new low-skilled employment in jurisdiction 2. As high-skilled move to jurisdiction 2, the marginal product of low-skilled labor increases, leading workers to enter the labor market. The third row gives the welfare gain to the still-employed in both jurisdictions. The wage, conditional on employment, increases in jurisdiction

[21] The final term of the government's social welfare function therefore involves integrating over the types of households in each jurisdiction.

[22] The last equality only holds if the lower bound on the moving costs are zero, which we maintain throughout, and for a small minimum wage.

1 due to the new binding minimum wage, while the wage in jurisdiction 2 increases because of induced high-skilled immigration. The final row gives the welfare loss to high-skilled workers, due to both the change in wages and moving costs paid by those who move to jurisdiction 2.

Proposition 6. *If (1) there is efficient rationing; (2) the demand elasticity is finite; (3) the supply elasticity of low-skilled workers is positive; (4) the government values additional wages to low-skilled workers more than the loss of wages to high-skilled; and (5) the first worker to move to the other jurisdiction has zero moving costs, $\xi = 0$, then it is total welfare improving for the federal government to impose a (small) binding minimum wage.*

Proposition 6 is the same as Proposition 1, except the central government only cares about migration insofar as it affects efficiency and moving costs paid. Condition (5) is similar to efficient rationing, but for moving, so that moving costs are zeros for a small binding minimum wage. It implies that the lower bound on the moving cost distribution is zero.[23] Even without condition (5), the federal government may implement a binding minimum wage if it values redistribution enough. An increase in the federal minimum wage leads to lower high-skilled wages, regardless of moving costs, and is therefore not Pareto improving. Comparing Propositions 1 and 6, we expect the federal government to prefer higher minimum wages than regional governments, all else equal.

The case in which the federal government can set a different minimum wage in each region is synonymous with the social optimum. The social optimum is defined in our context by the sum of total welfare for all agents in all regions. This is exactly the function maximized by the federal government. When its policy is allowed to be asymmetric, it has the flexibility to reach the socially optimal minimum wage in each region.

2.3.3 Federal Uniform Policy

In this case, we restrict the central government to set a single uniform policy. The minimum wage is given by:

$$\bar{w} = \arg\max SW_f(\bar{w}) \tag{7}$$

The FOC is quite similar to the non-uniform case, where the responses are given with respect to $\bar{w}$ instead of $(\bar{w}_1, \bar{w}_2)$. When $\gamma = 1$, the solution will be the same in the non-uniform and uniform cases. If $\gamma < 1$, then the competitive equilibrium low-skilled wages in jurisdiction 1 will be greater than those in 2. Therefore, when considering a small binding minimum wage, it will only be binding in 2. Under the assumptions of Proposition 6, the Social Planner will always want a binding minimum wage in each jurisdiction, however it may be the case that even when the assumptions hold, the federal government does not set a binding minimum wage if constrained to only set a

[23] This assumption on the moving cost distribution is not realistic and is chosen to provide sufficient, but not necessary, conditions for a federal minimum wage. However, in a long-run version of our model where workers do not have a fixed residence, the moving costs can be interpreted as preferences for one region compared to the other. In that case, it is plausible that some workers are indifferent, all else equal.

uniform policy. When the Social Planner sets a binding policy in each jurisdiction, it can mitigate the welfare loss due to migration, both moving costs and the suboptimal allocation of individuals across jurisdictions. When the federal government can only set one policy and it is only binding in a jurisdiction, then there is an externality imposed on the other. This is seen by noting that all of the components of the government's first order conditions depend on the enforced minimum wage in both jurisdictions. For $\gamma < 1$, a small binding uniform policy is equivalent to the Social Planner only setting a minimum wage in the low productivity region, which the previous section showed is suboptimal.

2.3.4 Tiered System

The US has a tiered system in which the federal and local government both set minimum wages, and the larger of the two is enforced. The federal government is restricted to set a uniform policy. Based on our previous analysis, under the assumptions of Proposition 1, each local government sets a binding minimum wage, and under the assumptions of Proposition 6, the federal government also sets a binding policy for at least the less productive region. As a result, there will be binding policy enacted in each. In the symmetric case, the federal minimum wage will be weakly higher than both jurisdictions' preferred policies because of the differential effects of outmigration and moving costs. The effect of moving costs must be less than outmigration by construction.

3 Quantitative Model

3.1 Parameterization

In the interest of exploring the relative welfare consequences of different configurations and the relative advantages of decentralized and centralized policy setting, we consider a parameterized version of the model. The calibrated model illustrates the mechanisms highlighted in the theory section, particularly regional heterogeneity and competition. In the previous sections, we are unable to give theoretical predictions for each agent (policymaker) beyond the extensive margin conditional on the other agent's policies. The calibrated simulations in the next section on optimal policy provide results for the intensive margin decisions when governments compete while factoring in the general equilibrium effects of their actions.

We specify the following functional forms. We assume f is a constant elasticity of substitution (CES) production function with elasticity of substitution between high and low-skilled labor $\sigma = \frac{1}{1-\rho}$ and weight α on the low-skilled.

$$f(l,h) = (\alpha l^{\rho} + (1-\alpha)h^{\rho})^{\frac{1}{\rho}} \tag{8}$$

With this production function, $\rho \leq 1$. If $\rho = 1$ the two types of labor are perfect substitutes and if $\rho = -\infty$ then the two types are perfect complements. Worker heterogeneity is determined by

the distributions on θ and ξ, which we assume to be $\theta \sim \mathcal{U}(0, \bar{\theta})$ and $\xi \sim \mathcal{U}(0, \bar{\xi})$. The sufficient conditions from propositions 1 and 6 require that the lower bounds are zero.

We also assume that the social welfare aggregator is defined as:

$$G(c_j) = \int_{j \in r} \log(c_j + \nu) dj \tag{9}$$

where ν is a utility function shifter. Since some low-skilled workers are always unemployed with $c_u = 0$, ν must be positive. $\log(\nu)$ is therefore the value of unemployment to the government. Smaller values of ν do not affect labor supply, but increase the social cost of unemployment. We choose individual utility to be net, after migration and disutility from working, consumption so that there are no income effects.

3.2 Calibration

We calibrate our model to match aggregate moments of the US economy. We have seven parameters: α, γ, and ρ for the production function; $\bar{h}, \bar{\theta}$, and $\bar{\xi}$ to describe the worker population; and ν for the social welfare function. We vary γ to see how different levels of regional heterogeneity impacts our results; for the calibration, we set $\gamma = 1$, corresponding to the symmetric case. Additionally, we let $\bar{h} = 1$ to reflect the fact that about one-third of the population has a Bachelor's degree (CPS), and $\rho = 0.286$, based on the estimate of the elasticity of substitution between high and low-skilled workers in Katz and Murphy (1992).[24]

Based on the timing assumptions of our model, we calibrate the remaining parameters in two stages. In the first stage of the calibration, based on the competitive equilibrium, we set α and $\bar{\theta}$. We calibrate α to match the relative wages across skill types. Using the Current Population Survey Outgoing Rotation Group microdata, the Economic Policy Initiative reports hourly wages in 2017 by highest degree received. Based on their report, we set $\frac{w^h}{w^l} = 1.82$. We calibrate $\bar{\theta}$ to match the 2017 labor force participation rate of high school graduates. Based on estimates from the Bureau of Labor Statistics in 2017, the value for high school graduates is 57.7%, while the value for those with only a Bachelor's degree is 73.3%. Since our model assumes that high skilled workers are always in the labor force, we scale up the high school rate to $\frac{57.7}{73.3} = 78.7$. ν is normalized to 1 so that the utility of the unemployed is 0. It also ensures that $G(c_h) > 0$

In the second stage, after all workers have established residencies from the competitive equilibrium, we calibrate the upper bound on the moving cost distribution, $\bar{\xi}$ to match the elasticity of migration with respect to wages of 0.5 from Kennan and Walker (2011). For a 10 percent decrease in γ, which would imply a 10 percent decrease in wages in jurisdiction 2 without migration, $\bar{\xi}$ is found so that there is an increase in the stock of high-skilled workers in jurisdiction 1 of 5 percent. With migration, the initial change in γ results in less than a 10 percent change in equilibrium wages.

[24] For the purposes of the calibration, we set $\bar{h}$ using college graduates to match the definition of high-skilled in Katz and Murphy.

The parameter estimates are presented in Table 1. The estimate column gives the values used in the welfare calculations and the residuals note the distance between the model implied value and the target, when relevant.

Table 1: Calibration Results

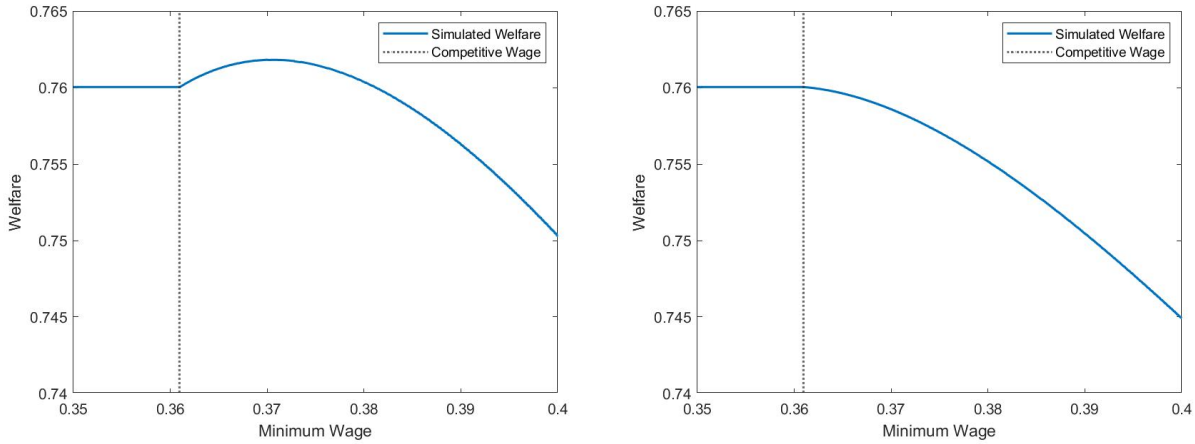
Parameters	Estimate	Residual
α	.43	10^{-9}
$\bar{h}$	1	-
$\bar{\theta}$	.46	10^{-9}
ρ	0.29	-
$\bar{\xi}$	1.31	10^{-13}
ν	1	-

From the baseline calibrated model, Figure 1 plots total social welfare in the economy as a function of the minimum wage. In the left panel, the minimum wage is effective in both regions and so this figure represents the problem of the central government restricted to a uniform policy. A small binding minimum wage increases welfare since the government values redistribution, the marginal separated low-skilled worker has 0 surplus from working, and no high-skilled workers move since the regions are identical and a uniform policy is enforced. As the minimum wage increase further to 102.5 percent of the competitive wage, welfare also increases. After this point, the higher price floor causes relatively too much additional unemployment and lost high-skilled wages. When the minimum wage is 5.6 percent above the competitive level, the policy is worse than no policy at all.

The right panel of the figure plots total welfare as a function of the minimum wage in region 1 with no minimum wage in the other. If region 2's government does not set a binding policy, this figure depicts the decentralized problem for region 1. For a small binding policy, total social welfare decreases in region 1 because the loss of high-skilled workers is more than the value from additional wages to the low-skilled. Welfare in region 2 increases from the inflow of new high-skilled workers, which increases the total population and low-skilled wages but at the expense of high-skilled wages. The increase in region 2 is not enough to offset the loss in region 1 and so a small binding minimum wage in only one region decreases total welfare in the economy. The migration externality is too large relatively to the benefits from redistribution.

4 Optimal Minimum Wage Setting

To supplement the theoretical work on the extensive margin, we use the calibrated model to trace out the relative benefits of centralization verse decentralization as a function of mobility and regional heterogeneity for optimal minimum wage setting. Our results illustrate the trade-off captured by the decentralization theorem, where the federal government can only set a single uniform policy and local governments can better adjust to their own needs. We find that centralization of minimum wage setting authority is always weakly preferred to decentralization in our baseline cal-



(a) Single Minimum Wage in Both Regions

(b) Minimum Wage in Region 1 Only

Figure 1: Welfare and the Minimum Wage

Note: Both panels of the figure plot the total social welfare in the economy for various levels of the minimum wage. The left panel plots welfare as a function of a single minimum wage applied in both regions while the minimum wage only applies in region 1 in the right panel. Since the regions are identical, the implications from a minimum wage in only 1 or 2 are the same. The vertical line denotes the competitive low-skilled wage in the absence of any policy.

ibrated model, even when regions are heterogeneous, because of interregional spillovers and the corresponding competition for high-skilled workers. Local governments never want to set binding minimum wages, while the central government loses its incentive to set a uniform minimum wage as heterogeneity across regions increases.

Panel A of Table 2 presents results for the symmetric case, i.e., when relative productivity $\gamma = 1$. The five rows correspond to the five different types of government structure that we allow in our context. The top row is the competitive equilibrium with no government, and the bottom row is the case of a benevolent central government setting non-uniform policy, which is synonymous with a social planner solution. The middle three rows correspond to cases with local governments only, a central government restricted to uniform policy only, and a US-style hybrid system with both of these types, respectively. Columns list the values for unemployment, wages for high-skilled and low-skilled workers, and the fraction of (mass 1) high skill workers that sort to region 1. If the wages of a particular case match the wages in the competitive case, then there is no binding minimum wage. The far right column lists the welfare loss relative to the federal non-uniform or social planner.

In the symmetric case, local governments do not want to set a binding minimum wage because the resulting outmigration is too large to be offset by the increase in wages for the low-skilled. Condition 4 of Proposition 1 does not hold. In contrast, the central government is not affected by this race-to-the-bottom behavior to keep high-skilled workers because migration does not change the set of agents in the central governments welfare function. In the symmetric case, the central government exactly matches the social planners policy, since the social planner sets identical

minimum wages in each region.

In our calibrated equilibrium, we find that the social planner and central government set the optimal minimum wage to 102.5% of the competitive wage. The binding minimum wage leads to a small total welfare increase over the competitive and decentralized cases of 0.23%. With higher low-skilled wages, the always employed low-skilled have a welfare gain at least 2.14% that depends on their working costs. The gain comes at the expense of the newly unemployed, whose welfare decreases by 100% as consumption drops to 0, and the high-skilled, whose welfare falls by 1.65%. As more low-skilled workers become unemployed, the marginal product of the high-skilled workers falls, leading to lower consumption. The always unemployed are unaffected by changed in the policy.

With heterogeneous regions, the central governments ability to match the social planner's solution disappears. Panel B of Table 2 presents the outcomes when γ is reduced to .995. With this very small productivity difference, the central government desires a binding minimum wage that is between the social planners two policies: slightly too high for the less productive region, but too low for the more productive area. The centralized and tiered environments both result in a positive welfare loss relative to the social planner, though they are still preferred to the decentralized case where local governments do not implement a binding policy.

Finally, we present results for $\gamma = 0.99$ in Panel C of Table 2. In this case, the local and central uniform governments do set a binding minimum wage, while the social planner's solution sets a different binding policy in each region. As in the previous cases, the migration incentive is too strong for the local governments to set a binding policy, and now that is also the case for the central uniform as well. A small binding central uniform policy only affects the less productive jurisdiction, which causes high-skilled workers to migrate. The migration externality decreases welfare as total output in the economy shrinks, even more low-skilled workers in region 2 become unemployed, and some high-skilled pay migration costs. In the $\gamma = 0.995$ case, the jurisdictions were similar enough that the same policy implemented in both did not cause too large of an externality. With large amounts of heterogeneity, like in the US, central uniform policies are inefficient. The optimality of a central uniform minimum wage is decreasing in regional heterogeneity. With two policies to set, the social planner is able to set the binding minimum wage in each jurisdiction so that the populations remain the same as in the competitive equilibrium.

4.1 Variable Mobility

The relative differences in government competition of mobile high-skilled agents drives our main results in the previous section. Local governments are averse to minimum wage induced outmigration, as it creates a first order welfare loss. The central government faces a much smaller cost from migration, especially since the first workers who move have no migration costs. However, both types of governments bear the burden of decreased total output as households sort across jurisdictions. We now explicitly vary the mobility of the high-skilled to better understand this mechanism. With high migration costs, the model captures the short-run effects as workers are

Table 2: Baseline Results

Authority	u	w_1^l	w_2^l	avg w^h	h_1	SW Loss
Panel A: $\gamma = 1$						
Competitive	.426	.361	.361	.657	.5	.23%
Decentralized	.426	.361	.361	.657	.5	.23%
Federal Uniform	.523	.370	.370	.643	.5	0
US	.523	.370	.370	.643	.5	0
Federal Non-Uniform	.523	.370	.370	.643	.5	-
Panel B: $\gamma = .995$						
Competitive	.429	.362	.358	.655	.506	.23%
Decentralized	.429	.362	.358	.655	.506	.23%
Federal Uniform	.533	.370	.370	.640	.509	.16%
US	.533	.370	.370	.640	.509	.16%
Federal Non-Uniform	.533	.372	.368	.640	.507	-
Panel C: $\gamma = .99$						
Competitive	.432	.364	.356	.653	.513	.23%
Decentralized	.432	.364	.356	.653	.513	.23%
Federal Uniform	.432	.364	.356	.653	.513	.23%
US	.432	.364	.356	.653	.513	.23%
Federal Non-Uniform	.522	.372	.364	.640	.513	-

Note: This table presents the simulated equilibria for our baseline calibrated model different minimum wage setting authorities and levels of regional heterogeneity. The competitive equilibrium is provided for comparison. Social Welfare (SW) Loss is defined relative to the Federal Non-Uniform, or Social Planner.

Table 3: $\gamma = .95$ and $\xi \sim \mathcal{U}[0.0025, \bar{\xi}]$

Authority	u	w_1^l	w_2^l	w^h avg	h_1	SW Loss
Competitive	.456	.374	.334	.638	.566	.21%
Decentralized	.497	.378	.338	.632	.566	.08%
Federal Uniform	.466	.374	.336	.367	.566	.19%
US	.542	.383	.343	.626	.566	0%
Federal Non-Uniform	.542	.383	.343	.626	.566	-

Note: This table presents the simulated equilibria for the case where $\gamma = .95$ and $\xi \sim \mathcal{U}[0.0025, \bar{\xi}]$ for different minimum wage setting authorities. The competitive equilibrium is provided for comparison. Social Welfare (SW) Loss is defined relative to the Federal Non-Uniform, or Social Planner.

unable to move in response to the policy. As migration costs decrease, the model better captures the long-run equilibrium effects.

When migration costs increase, minimum wages become more desirable for regional governments; when high-skilled workers are immobile, regional governments enact exactly the social planner policy. Without migration, there are no interregional spillovers and the local governments problems exactly match that of the social planner because the two economies are additively separable in the social welfare function. From the baseline symmetric model, if the minimum of the moving cost distribution is increased slightly to 0.0025 from 0, or 0.3% of the competitive high-skilled wage, then the decentralized outcome would match the social planner as well. The lowest moving cost is large enough that the high-skilled are effectively immobile in equilibrium. As long as the minimum moving cost is more than 0.03% of the high-skilled competitive wage, regional governments will enact binding policy in the baseline model. Small changes in the moving cost have large implications for government competition.

To understand the interaction between productivity heterogeneity and mobility, we set γ to .95 and the minimum of the migration cost distribution to 0.0025. Table 3 presents the results. As before, the social planner implements a different binding minimum wage in each jurisdiction. In the decentralized case, migration costs lessen horizontal competition enough so that both local governments set binding policies, but not enough to replicate the social planner's solution. The central uniform policy alone leads to greater welfare loss than the decentralized policy setting because there is considerable heterogeneity, but the government only has one instrument. However, the hybrid system captures the social optimum. When the central government sets its policy equal to that of the social planner in jurisdiction 2, it now becomes optimal for jurisdiction 1 to raise its policy. The local government faces a smaller threat from outmigration than the decentralized case where the regional government in jurisdiction 2 sets its policy too low. In this more realistic case, where jurisdictions are heterogeneous and all migration is at least a little costly, the hybrid minimum wage setting authority is welfare improving over either by itself. Centralized and decentralized policy setting are strategic complements. This justifies the policies we see in Seattle, for example, while not supporting the policies by the Missouri state government of prohibiting its cities from setting higher minimum wages or in Europe where many countries set a single national policy.

At the other extreme, we consider perfectly mobile high-skilled workers. In this case, the outmigration response to a minimum wage will be even stronger than in the baseline case, such that local governments will have even less of an incentive to implement binding minimum wages. For the central government, the effect is ambiguous. The lack of moving costs means that migration is less costly, since what matters to the central government is the costs paid by individuals rather than the loss of population. However, the central government also cares about the total output in the economy, which is decreasing with migration. When $\gamma = 1$, no households move as a result of federal policy, and the calibrated migration and perfect mobility equilibria are identical. When $\gamma = 0.95$, the federal non-uniform optimal policies are 0.379 in jurisdiction 1 and 0.340 in jurisdiction 2, compared to 0.383 and 0.343 in the baseline calibrated case. Higher levels of mobility push down optimal federal policy.

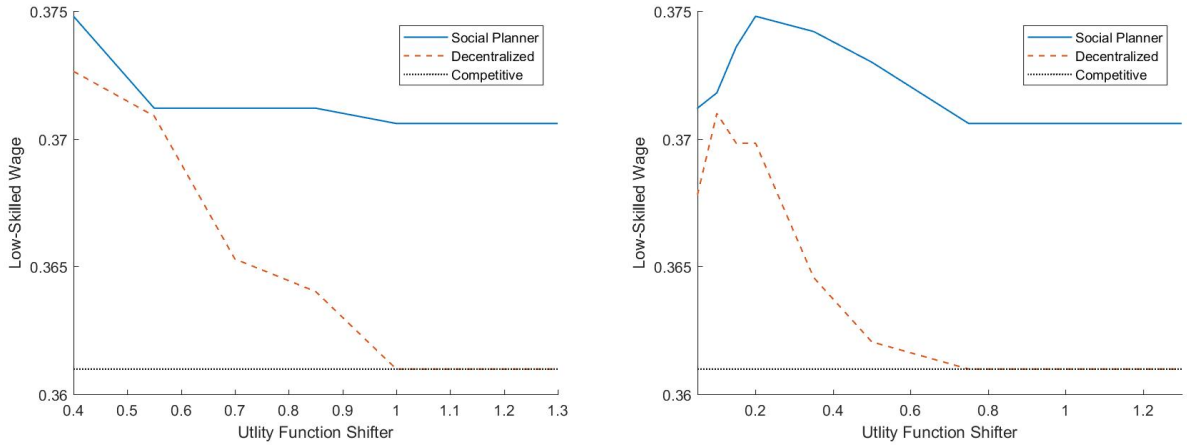
4.2 Alternative Social Welfare Functions

The previous quantitative results rely on a single specification of the social welfare function, namely, $U(c) = \log(c + \nu)$, where we set $\nu = 1$. ν is a shifter on the progressivity of the welfare function, with lower values corresponding to a more progressive desire for redistribution. The right panel of Figure 2 presents the equilibrium minimum wages desired by local governments and by the social planner as a function of the ν for the symmetric case holding all other structural parameters fixed. Since our theory requires that $G(c_h) > 0$, ν cannot be too small in our baseline model. To fully investigate the importance of progressivity, the left panel of the figure uses the social welfare function $\log(2(c + \nu))$ so that $G(c_h) > 0$ for all $\nu > 0$ since in the symmetric baseline equilibrium, $c_h > .6$ for any feasible minimum wage policy.

The optimal minimum wages are an inverted-U shape in progressivity. As ν approaches zero, the welfare of the unemployed quickly heads toward negative infinity, such that adding to unemployment becomes increasingly costly from a social welfare perspective. Minimum wages distribute from the high-skilled and the newly unemployed to the low-skilled employed. If ν were 0, then the social welfare would be negative infinity for all policies, including no binding minimum wage, since there will always be some unemployed workers in our model. In that case, the optimal minimum wage would not be defined for any government. The pattern is identical when $\gamma < 1$.

We now consider a utilitarian social welfare function. Since our agents maximize consumption, this social welfare function is equivalent to maximizing the sum of wages minus any working costs or moving costs paid. In the symmetric $\gamma = 1$ case, no binding minimum wages are optimal for any government. With a relatively large amount of regional heterogeneity, for example when $\gamma = 0.95$, the social planner desires a small binding minimum wage of about 0.3% over the competitive wage in each jurisdiction. Under our baseline social welfare function, when $\gamma = 0.95$, the social planner would set optimal minimum wages of about 2.5% above the competitive.

In addition to the productivity heterogeneity explored previously, regions may also differ in their progressivity. St. Louis and Kansas City, Missouri as blue cities in a red state may have set higher minimum wages because they value redistribution more. To test this implication, Figure



(a) Baseline model

(b) Utility given by: $\log(2(c+\nu))$

Figure 2: Optimal Minimum Wage as a Function of ν

Note: In the right panel, we vary ν from 0.4 to 1.3 using the baseline calibrated model. In the left panel, we vary ν from 0.05 to 1.3 changing the social welfare function to $\log(2(c+\nu))$ The low-skilled wage in a jurisdiction is the maximum of the competitive wage and that government's minimum wage policy. The Social Planner's optimal policy (solid line) equals the Central Uniform and Tiered optimal policies since jurisdictions are symmetric.

3 varies ν for only jurisdiction 1, holding the value for jurisdiction 2 fixed at 1, and plots the equilibrium decentralized policies and the competitive outcome when jurisdictions are otherwise identical. When ν is slightly less than 1, jurisdiction 1 values redistribution sufficiently to set a binding minimum wage. Unlike in the baseline, condition 4 of Proposition 1 is now satisfied. Although the welfare cost of setting a minimum wage for the government in jurisdiction 2 is now lower because the binding minimum wage in the other region reduces the value of migration for the high-skilled, it remains at the corner of no binding policy. It is not until ν in jurisdiction 1 falls below 0.7 that jurisdiction 2 sets a binding minimum wage. As ν in jurisdiction 1 falls further, jurisdiction 2 is induced to set a binding minimum wage. In this region of ν, governments engage in race to the top. When ν decreases even further, the benefit of setting a high minimum wage also decreases and the policies begin to level off. Even when regions are equally productive, differences in the value of redistribution can lead competing governments to set binding decentralized policies.

4.3 Information

In addition to mobility and regional competition, the relative information quality of different levels of government affects the optimal minimum wage setting authority. Decentralized policy setting may be preferred because local governments have better knowledge of the labor supply, labor demand, and migration elasticities or local productivity. These parameters jointly determine how employment changes after a minimum wage increase. In a 2019 report, the CBO documented a large amount of variation in estimates of the employment elasticities from the recent minimum wage

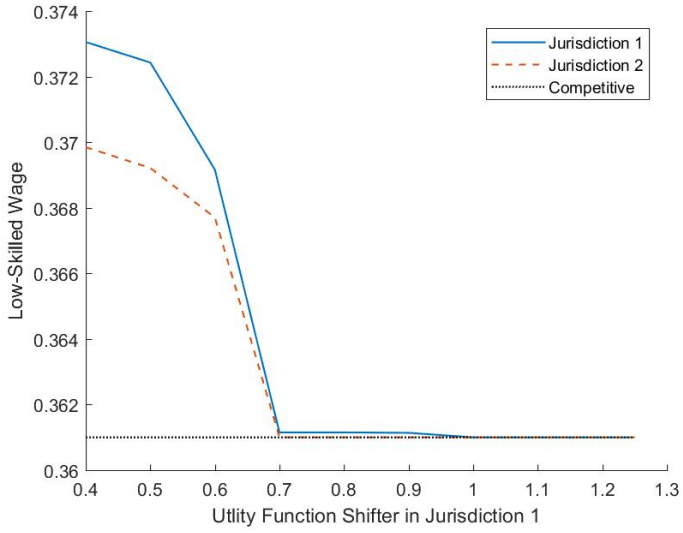


Figure 3: Optimal Minimum Wage as a Function of ν in Jurisdiction 1

Note: We vary ν in jurisdiction 1 from 0.4 to 1.3 keeping ν in jurisdiction 2 fixed at 1. The low-skilled wage in a jurisdiction is the maximum of the competitive wage and that government's minimum wage policy.

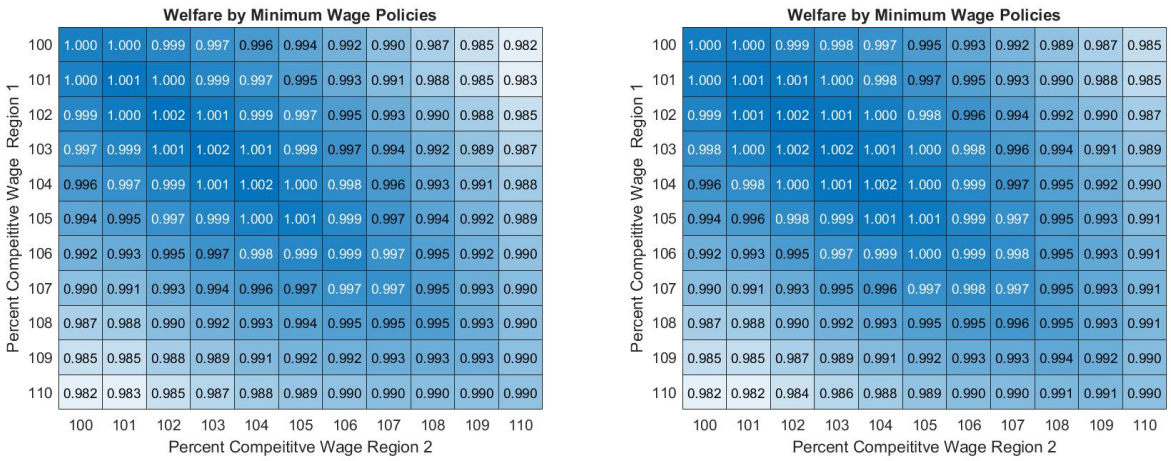
literature from 0.4 in Cengiz et al. (2019) to -1.7 in Clemens and Wither (2016). Which estimate should a government use when determining its policy? In the context of our model, if a government thought the employment elasticity were positive, then it would believe that a binding minimum wage is Pareto improving and only faces a welfare trade-off when the estimate is negative. Therefore different estimates lead to very different "optimal" policies. Local governments may additionally have better information about the market structure, which we do not explore in this paper, or political beliefs leading them to favor one estimate over another; for example, if it believes there is significant monopsony power in the region, it may place more weight on positive estimates. If any government has noisy or incorrect information about the regions or these parameters, then it will not be able to set policy optimally.

While the previous analysis assumed that both levels had perfect information about workers and regions, we now explore the consequences of misspecified policy.[25] Panel (a) of Figure 4 presents the welfare for different minimum wage policies in each region for the baseline model normalized by the welfare in the competitive case with no binding policy. In each row, the minimum wage varies from 0 to 10 percent above the competitive wage in region 1. The columns vary the price floor in region 2. The grid therefore constructs a discretized version of the social planner's problem and it confirms that the optimal minimum wage is about 2.5 percent above the competitive. Similarly, looking down the main diagonal gives the optimal central uniform policy. Based on these simulations, we see that very large uniform minimum wages can lead to lower total welfare than having no policy,

[25] We do not make explicit assumptions about why the minimum wage differs from the optimal policies found in the previous subsection but instead simulate how deviations affect welfare. These errors may be due to a lack of information.

which is also the decentralized case. Importantly for policy-making, these results suggest that when regions are symmetric, a minimum wage that is too high is better than one that is too low. Specifically, we find that a minimum wage 0.5 percentage points too high yields higher welfare than one that is 0.5 percentage points too low. The same is true when we compare policies that are 1.5 and 2.5 percentage points off in either direction.[26]

Asymmetric policies for which the difference is more than 2 percentage points also lead to lower welfare than having no policies at all. When the enforced minimum wage is too different in otherwise identical regions, the migration externality overpowers the benefits from redistribution. Since symmetric policies do not induce migration, relatively more of the burden falls on high-skilled workers that the progressive government values less. However, large differences in the enforced low-skilled wage across equally productive regions leads to a large amount of migration. This exacerbates the disemployment effects of the minimum wage in the region with a higher policy. In our simulations, we find that minimum wages of 101.5 and 103.5 percent of the competitive performs worse than a uniform policy of 106 percent.



(a) $\gamma = 1$, $\min(\xi) = 0$ (b) $\gamma = .95$, $\min(\xi) = 0.0025$

Figure 4: Welfare and the Minimum Wage

Note: Panel (a) presents the total social welfare in the economy for different minimum wage policies in both regions ranging from the competitive wage to 110 percent of the competitive. Welfare is normalized so that it is 1 in the competitive case where neither region has a binding minimum wage. Panel (b) conducts the same simulations but for the case where $\gamma = .95$ and $\min(\xi) = 0.0025$ as in Table 3. Minimum wage combinations with higher total welfare are shaded darker.

To understand how errors in optimal policy setting from a lack of information vary with productivity and mobility, Panel (b) displays the relative welfare based on the simulations shown in Table 3 where $\gamma = .95$ and $\min(\xi) = 0.0025$. With perfect information, the optimal uniform central policy is less than 101 percent of the competitive wage in region 2 and not binding in region 1

[26] We cannot compare more than 2.5 percentage points off of the optimal because then the low minimum wage would not be binding.

while the optimal non-uniform policy is about 102.5 percent above each region's respective competitive wage.[27] Since moving costs are positive on the margin, both regions set binding policies of about 1 percent above the competitive in the decentralized case as well. The patterns are very similar to the baseline example except the welfare matrix is no longer symmetric. Based on the asymmetry, we want to determine if the government should be relatively more cautious in the high or low productivity region. Our results indicate that setting too high a minimum wage in the low-productivity region while setting too low a minimum wage by the same percentage points in the high-productivity region is better than the reverse.

4.4 Taxes and Transfers

In the previous sections, we study the optimal minimum wage without taxes, which is an "optimal suboptimal policy. When the federal government has access to linear income taxes (a single income tax rate and a demogrant), it will always prefer linear taxes to a combination of a minimum wage and linear taxes in our calibrated model. The proof of this numerical result in our framework is presented in Appendix A. Lee and Saez (2012) prove in a model without migration, when workers are not mobile across skill levels and non-linear taxes and transfers are available, then it is not optimal to also have a binding minimum wage. This is analogous to the no mobility case of our model discussed above. In a search-and-matching framework, Lavecchia (2018) finds that both a minimum wage and optimal taxes are preferred only if the government has very strong redistributive preferences.[28] As shown in the previous sections, the migration externality from any policy that restricts low-skilled labor supply and induces migration is worse for the local governments than the federal, and income taxes are no exception. This is intuitive in light of the literature which tends to find that redistribution is better handled by the federal government (e.g. see Gordon and Cullen 2012).

If the government taxes income, then low-skilled workers will restrict their labor supply, impacting the marginal product of all workers and lowering total output in the economy. High-skilled labor supply is not perfectly inelastic to each jurisdiction, but it is perfectly inelastic to the economy and the federal government. The government trades off redistribution with efficiency. A binding minimum wage will increase low-skill labor supply, but decrease labor demand, while an income tax will decrease labor supply and have no effect on demand, potentially allowing the policies to be complements. While the optimality of jointly specifying a minimum wage and linear income tax is uncertain, we show that under the assumptions of our calibrated model, the federal government would never implement a binding minimum wage if linear taxes are available.

[27] The diagonal of this figure does not give the optimal central uniform policy because regions have different competitive wages. Since productivity in region 2 is much lower than 1, the optimal uniform minimum wage is found by looking across the top row.

[28] Redistributive policies may also have an insurance component that high-skilled workers care about (Hoynes and Luttmer 2012). If a high-skilled worker was separated from her job due to a shock, and then was able to find employment at the minimum wage rather than the low-skilled competitive one, the insurance value would be positive. However, in our static model with fixed worker skill levels, the high-skilled are made weakly worse off from a minimum wage increase.

Although it is not optimal to have both a binding minimum wage and a federal income tax, we incorporate this into our model and recompute the equilibrium to compare the policies. Our simulations confirm that the government does not want to set a binding minimum wage when it sets the income tax optimally. In our baseline symmetric calibrated model, we find that federal government would choose to set a tax rate of 24% and no binding minimum wage in both jurisdictions, as expected. The policy dramatically increases unemployment to 35.5% of the low-skilled workers from 26.2% under the optimal minimum wage. The restricted labor supply leads low-skilled wages to rise to 0.390, and high-skilled wages to fall to 0.616. The government trades off its redistribution incentive with decreased low-skilled labor supply that leads to increased unemployment and falling high-skilled wages. As the income tax increases, the marginal (zero consumption) low skilled worker exits the labor force. This raises the marginal product of all the still employed low-skilled workers, thus acting very much like a minimum wage, except now the government is also able to redistribute some of the additional income to the unemployed.

The optimal linear income tax increases total welfare over the optimal minimum wage by 2%. Compared with the optimal minimum wage model, where the utility of the unemployed is 0 by construction, the utility of unemployed is 0.126. The low-skilled worker with the lowest working cost sees her utility increase by 16.2% under the optimal linear income tax compared with a minimum wage, which is a lower bound on the welfare increase for low-skilled workers. High-skilled workers see their welfare fall by 5.1%, as the government redistributes total wealth. If the federal government were able to set a different tax rate on each skill type, it would tax low-skilled at 73% and the high-skilled at 100%, leading to a welfare gain of 8.5% over the optimal minimum wage. This result is driven by the fact that the high-skilled workers supply labor inelastically to the economy. Since the government can use that revenue to lessen the harm of unemployment, it then sets a much higher tax rate on the low-skilled as well.

Throughout this section, we assumed that the governments tax income. However, if the tax is on consumption, equal to income minus cost of working and migration, then there is no labor supply response, and a federal government with a concave social welfare function would optimally set a leveling tax so that all workers have the same utility. With a leveling tax, a minimum wage is suboptimal as the induced unemployment lowers total output in the economy that can be redistributed. A utilitarian government would have no desire to redistribute through a consumption tax since the total output consumed would not change.

5 Minimum Wage Setting in the United States

Our two region model demonstrated that decentralized and centralized minimum wage setting are strategic complements and that the extent to which they are depends on regional heterogeneity and mobility. While this approach highlights important mechanisms, it is unable to quantify the welfare gain from the tiered policy setting used in the US compared to other minimum wage setting authorities, like centralization. To do this, we now extend the model to the case of US states.

5.1 Model

We consider a spatial equilibrium model of local labor markets as in Rosen (1979) and Roback (1982) with two imperfectly substitutable types of labor, and state and federal governments. As in our previous model, low-skilled workers can be employed or unemployed, and are immobile. If a worker l is employed in state j, she pays working cost $\theta_l \sim \mathcal{U}(0, \bar{\theta}_j)$. High-skilled workers in state j have indirect utility given by

$$V_j = w_j + \phi_j + \kappa\epsilon_j \tag{10}$$

where w_j is the wage rate in j, ϕ_j is a state j amenity,[29] and ϵ_j is an i.i.d. type I extreme value idiosyncratic preference for j, with standard deviation κ. High-skilled workers pick the location that maximizes their utility, which implies that the probability a worker lives in j is

$$p_j = \frac{\exp\left(\frac{w_j+\phi_j}{\kappa}\right)}{\sum_h \exp\left(\frac{w_h+\phi_h}{\kappa}\right)} \tag{11}$$

Each state has one firm that hires low and high-skilled workers with production function

$$f_j(l,h) = \gamma_j(\alpha_j l^\rho + (1-\alpha_j)h^\rho)^{1/\rho} \tag{12}$$

This specification differs slightly from Equation 8 in that it allows both γ and α to vary across markets. The government's social welfare function is given in Equation 9.

5.2 Data and Calibration

We calibrate the model to the continental US states in 2015. In 2013 and 2014, Congress considered the Minimum Wage Fairness Act, which would have raised the federal price floor in 2015 for the first time since 2009. After this failed, President Obama continued to push for an increase through 2015, although Congress did not act. Since the federal government considered changing the minimum wage, but chose not to, we take the 2015 policy to accurately reflect the government's preferences for redistribution. This allows us to back out ν for the federal government from the observed equilibrium.[30] While there has been a continued discussion of increasing the minimum wage, the

[29] Mapping our framework to standard economic geography model, ϕ captures amenities, rental value and taxes which we take as exogenous in this paper. As we note the in previous section, a richer model in future work should explore the interaction of these policies.

[30] More specifically, ν is calibrated so that \$7.25 is the welfare maximizing central minimum wage policy, holding the state policies fixed at their 2015 levels. This implies a Nash equilibrium where the federal government does not find it beneficial to deviate given the strategies of the states. However, the governments may be more sophisticated and know that changes in federal policy will also impact the state's optimal policies. We address calibrating ν for all states in Section 5.5.

federal government has not considered changing it since. We therefore use the 2015 American Community Survey (ACS) to estimate the model parameters.

From the ACS, we observe individual level wages by education and labor force attachment in each state. This gives the number of the low (high school equivalent) and high-skilled (college graduates) employed workers, as well as their average wages, the number of unemployed, and the number not in the labor force. Each state's relative number of employed, unemployed, and not in the labor force high school equivalent workers determines $\bar{\theta}_j$. The number of high-skilled residents in each state in 2015 pins downs ϕ_j.

The calibration of the state specific production function follows Section 3.2. First, we set $\rho = 0.286$ following Katz and Murphy (1992). We calculate α_j to match the wage gap between high and low-skilled workers and γ_j to match the state's minimum wage.[31]

Given the structural parameters, we then back out the federal government's redistribution preferences so that the observed policy of \$7.25 maximizes total social welfare in the economy. Table 4 summarizes the parameter estimates.

Table 4: US Model Calibration Results

National Parameters	Estimate	
κ	1.11	
ρ	0.29	
ν	3	
State Parameters	Mean	Std. Dev.
$\bar{\theta}_j$	14.99	1.31
α_j	0.43	0.06
γ_j	21.44	2.72

5.3 Welfare and US Minimum Wage Setting Authority

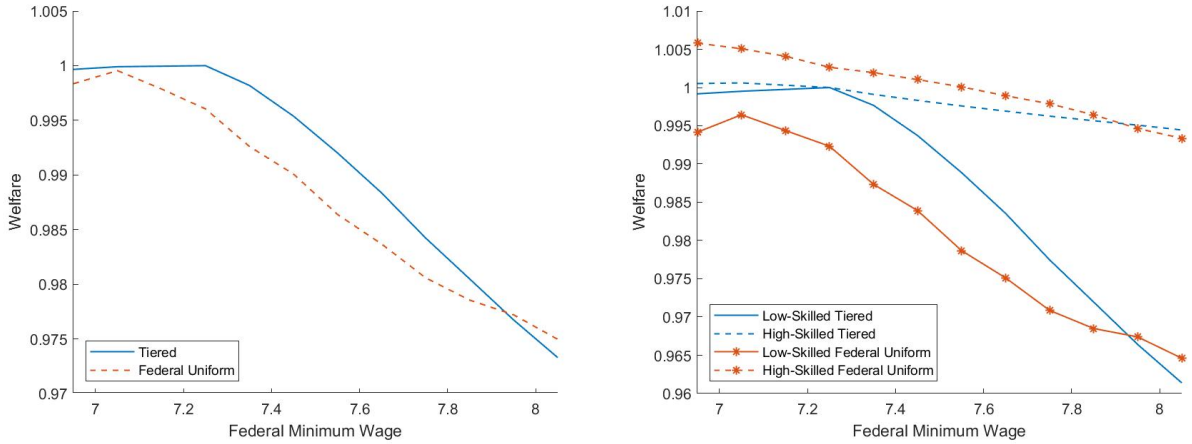
How large are the gains from state and the federal tiered minimum wage setting in the US over centralized policy alone? To measure the potential gain, Panel (a) of Figure 5 plots total welfare in the economy under both minimum wage setting authorities. The solid line shows how welfare in the tiered system changes as a function of the federal minimum wage. Based on the calibration, when the states' minimum wages are held at their 2015 levels, the federal government optimally sets a minimum wage of \$7.25, which confirms our calibration of ν is successful. The dashed line in the figure shows welfare as a function of centralized policy. Without the state policies, the federal government optimally chooses a lower minimum wage of \$7.05. Panel (a) of Figure 6 shows the relationship between tiered minimum wages and low-skilled wages under centralization.[32] This result is in line with the previous findings that state and federal polices are strategic complements. Higher

[31] For this calculation, we calculate the state's wage gap as the ratio of college graduates and high school equivalent or less in the ACS for employed workers between ages 16 and 70. Since the minimum wage is a floor, the ratio of high-skilled wages to the minimum wage would overstate their relative marginal product in the model.

[32] Based on our calibration, the tiered minimum wages are binding in all states, while our simulations predict that the optimal federal uniform minimum wage is only binding in 19 states.

state minimum wages reduce the high-skilled migration externality and allow the federal government to engage in more redistribution by setting a higher wage floor itself. Optimal centralized policy has a small welfare loss of about 0.3 percent.[33]

The figure also shows that suboptimal policy in the form of low wages, for example due to information frictions discussed earlier, has a lower social welfare cost under tiered policy setting. Small decreases in the federal minimum wage from the optimum in the tiered system have very little effect on aggregate welfare compared to centralized policy setting because many states set their own at or above the federal level; decreases only impact 7 of the 48 states. However, suboptimally high federal minimum wages under centralization and the tiered system lead to similar welfare losses.



(a) Aggregate Welfare

(b) Welfare by Worker Type

Figure 5: US Welfare and the Federal Minimum Wage

Note: Panel (a) plots total social welfare as a function of the federal minimum wage under tiered and federal uniform minimum wage setting authorities. Welfare is normalized to be 1 under optimal tiered policy setting with a federal minimum wage of $7.25. Panel (b) splits welfare into the contributions from low and high-skilled workers separately. Welfare is normalized to be 1 for each worker type under optimal tiered policy setting.

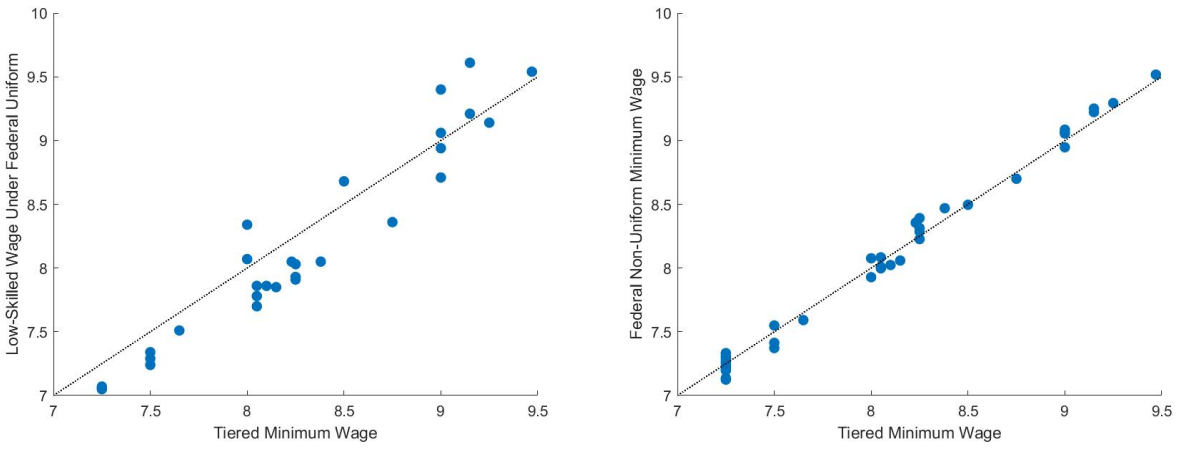
At very high federal minimum wages, the federal government prefers a uniform centralized policy over a tiered system. This is driven by the fact that states are heterogeneous in their preferences for redistribution and the most productive states do not always set the highest minimum wages. The correlation between the 2015 state policies and γ_j is 0.53.[34] If a relatively low productivity state sets too high a minimum wage relative to what is binding in other states, too many high-skill workers move. Since the state is of relatively low productivity, this disproportionately harms its low-skilled workers by increasing unemployment. Given this tension created from differences in preferences for redistribution, it is possible for optimal centralized policy to lead to higher total

[33]The competitive equilibrium has a welfare loss of about 0.8 percent.

[34]This calculation may overstate heterogeneity in preferences because states whose optimal policies are below $7.25 have no incentive to set their own. Using the 2015 effective minimum wages, the correlation is 0.85, but for similar reasons this likely overstates heterogeneity. Heterogeneity in other production parameters matters as well. The correlation between α_j and the state policies is -0.36.

welfare from the perspective of the federal government than the tiered system. However, in our model and data we do not find this to be the case near the optimal policies. It is important to note that we calculate welfare from the tiered system holding state policies fixed, while the model predicts that states will increase their minimum wages when the federal government does.

To understand who benefits and who bears the burden of the minimum wage and different wage setting authorities, Panel (b) of Figure 5 presents welfare by worker type normalized to 1 for both low and high-skilled workers under the optimal tiered policy. The results imply that the majority of the burden from the minimum wage falls on the newly unemployed. In both the tiered and federal uniform systems, high-skilled workers are minimally affected by federal policy since they are mobile. When the minimum wage increases in one or many states, high-skilled workers can move to other locations that offer similar welfare, which reduces the impact of the higher minimum wage on high-skilled workers who remain. However, outmigration in this model lowers the number of low-skilled workers employed at the minimum wage. Changes in the federal minimum wage have smaller impacts on the overall economy and high-skilled migration under the tiered system and so it allows the federal government to engage in more redistribution. Therefore, while low-skilled workers are mainly harmed by the minimum wage, the high-skilled bear most of the burden of the tiered system compared with centralized uniform policy.



(a) Centralization (b) Federal Non-Uniform

Figure 6: Comparing US Minimum Wage Setting Authorities

Note: Panel (a) plots the low-skilled wages in each state under tiered and federal uniform policy setting as well as a dotted 45 degree line for reference. Panel (b) similarly shows the relationship between tiered policy and the optimal federal non-uniform.

While the current US minimum wage setting authority yields higher welfare than centralized policy alone, it may be far from the social planner's solution. The 48 states are heterogeneous in their productivity and low-skilled working costs, but 21 of them have effective minimum wage policies of $7.25. This is in part because states face a different trade-off than the federal government as emphasized in equations 2 and 5, and because some states may be more or less progressive. Panel

(b) of Figure 6 shows the relationship between the observed tiered policies, and those under federal non-uniform. States below the 45 degree line have too high minimum wages under tiered policy setting. When not limited to a single policy, the government can better differentiate among states. Those subject to the \$7.25 minimum wage in the tiered system have federal non-uniform policies that range from \$7.12 to \$7.33. The federal government also sets slightly higher minimum wages for states with initially high policies in this case since it is able to better internalize the spillovers from high-skilled migration. Overall, current US policy closely matches the social optimum predicted in this model with only a 0.1 percent welfare loss.

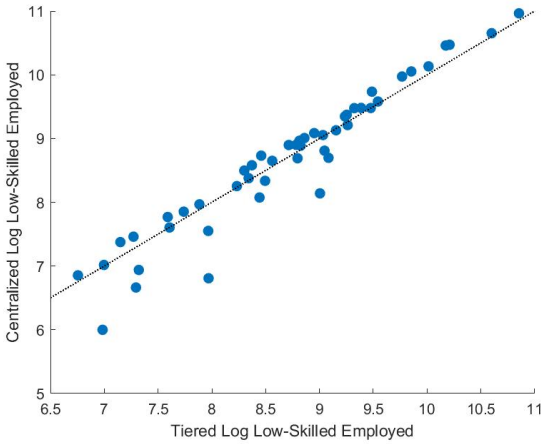
5.4 State Heterogeneity and Regional Redistribution

The minimum wage setting authority potentially impacts redistribution across states in addition to skill groups. Under centralization, the federal government has one policy lever that only directly affects the lowest productivity states. Higher federal uniform policies will therefore redistribute to more productive states. To some extent, this can improve total welfare because the more productive states are larger and have more low wage workers and unemployed, but a progressive government does not value transferring wealth from low to high income areas. Moving from centralization to the tiered system, there is an ambiguous effect on activity across states that depends on the relative progressivity of different governments.

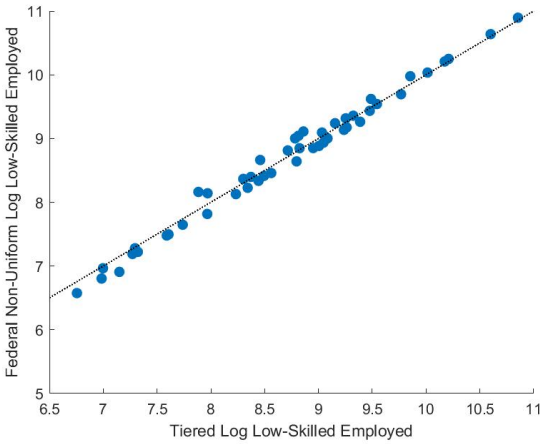
To measure how US states would be differentially affected under centralization compared to the current tiered system, the left three panels of Figure 7 show low-skilled employment, and high-skilled location decisions and wages change under these two regimes. Under centralization, Panel (a) shows that low-skilled employment is higher in most places. It mostly, but not exclusively, increases in states with low tiered minimum wages. The states experiencing the biggest decreases from centralization, represented by points far below the 45 degree line, are all relatively small, which reflects the government's trade off of redistributing to large workforces and low wage places. Without state minimum wages reducing the high-skilled outmigration externality, the low-skilled workers in the lowest productivity areas are hit the hardest.

There is very little state heterogeneity in the impact on high-skilled wages because these workers are imperfectly mobile. If states became too different, some workers would have an incentive to move. Panel (c) shows that wages are very close to the 45 degree line, but almost all are above. With only a single low federal minimum wage of \$7.05, aggregate low-skilled employment increases and drives up high-skilled wages. Centralization alone has only a small effect on the high-skilled migration externality and therefore allows for less redistribution from high to low-skilled workers, leading to higher average wages. However, since low-skilled employment decreases in some states, high-skilled workers' marginal products fall and they leave those affect low productivity areas. For this reason, Panels (a) and (e) look quite similar.

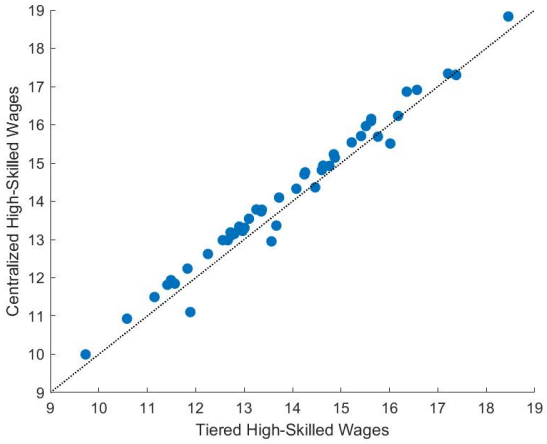
The regional distributional consequences of federal non-uniform over tiered policy setting are also ambiguous because states' preferences may be very different from the federal government. However, the optimal tiered minimum wages closely match the federal non-uniform, and so the



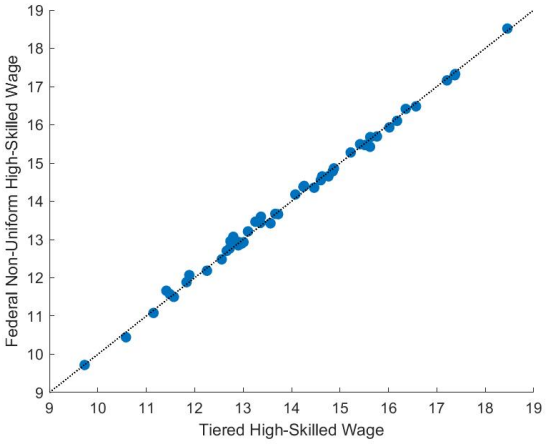
(a) Centralized Low-Skilled Employment



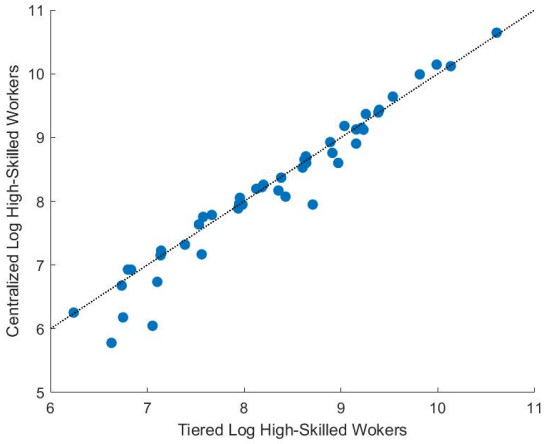
(b) Federal Non-Uniform Low-Skilled Employment



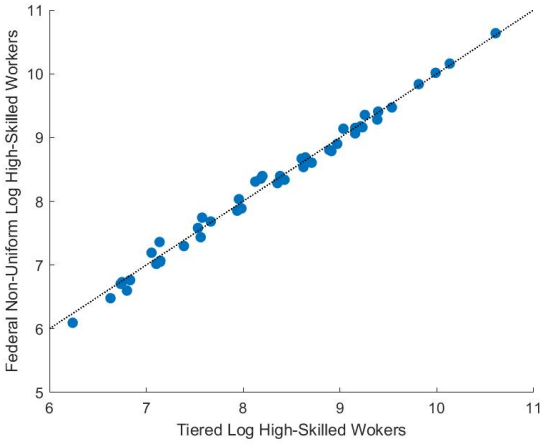
(c) Centralized High-Skilled Wages



(d) Federal Non-Uniform High-Skilled Wages



(e) Centralized High-Skilled Residencies



(f) Federal Non-Uniform High-Skilled Residencies

Figure 7: The Impacts of Minimum Wage Setting Authorities Across States

Note: The left three panels (a, c and e) compare tiered and centralized policy setting, while the right three (b, d, and f) compare the tiered and federal non-uniform systems. All figures contain a a dotted 45 degree line for reference.

right three panels of Figure 7 also show very little difference in low-skilled employment, and high-skilled location choices and wages.

5.5 A Brief Note on Decentralized Minimum Wage Setting

Competition between states for mobile high-skilled workers leads to low minimum wages under decentralized policy setting. Local governments are therefore less able to redistribute income from high to low-skilled workers. The amount of redistribution under decentralization compared with other systems importantly depends on regional heterogeneity in ν. Following the previous calibration for the federal government, we aim to estimate $\nu_j, j = \{1, ..., 48\}$, such that the observed 2015 minimum wage in state j maximizes j's welfare holding the other minimum wages fixed, so that the 2015 tiered policies are a Nash equilibrium. Since only 27 states set a binding minimum wage above \$7.25, ν_j is not point identified for every state.

In the calibrated model, there does not exist ν_j that replicates the state's 2015 policies. Under our assumptions and numerical values, each state would prefer a smaller minimum wage, holding all other states' and the federal government's policies constant. That is, there does not exist a progressivity level such that the current state minimum wages are best responses. There are several possible reasons why the model is unable to match the data. First, governments may be more sophisticated. If a states believes that others will set lower minimum wages too when it lowers its own policy, then there are smaller gains from doing so. The main advantage of a lower minimum wage is to attract more high-skilled workers, but the amount is a function of low-skilled wages everywhere. Second, states could face political frictions. They may not be able to set their policies on a continuum but are instead restricted to a small grid. The observed policies may be optimal on the grid but not globally. Finally, the model is potentially misspecified or, at least, states may have different beliefs about the implied elasticities. In particular, the model implications for the effect of the minimum wage on low-skilled unemployment and high-skilled wages may not be true, for example if workers have monoposony power or the high-skilled have larger moving costs. More work is needed in the future to better understand state competition and minimum wage setting.

While we are unable to compare the decentralized and tiered equilibria explicitly, we aim to measure the relative trade-off of local and federal policy setting, holding preferences for redistribution fixed. Only 21 of the 48 states set binding minimum wages under decentralization when all governments have the same value of ν equal to 3. Without a binding policy in many states from the federal government, the high-skilled migration externality is large and states optimally set lower minimum wages. Figure 8 shows that the low-skilled wages in every state are lower under decentralized policy setting compared to the current US policies. However, total welfare under this version of decentralization is nearly the same as the tiered system. The loss from less redistribution is offset by the gains from no heterogeneity in ν.

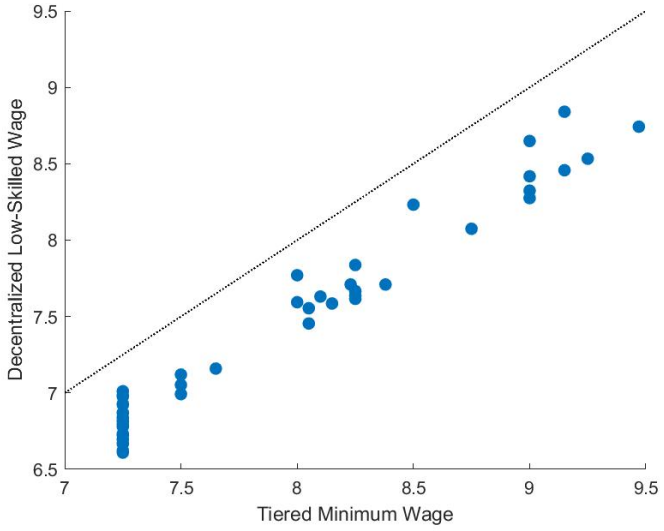


Figure 8: Decentralized Minimum Wage Setting

6 Conclusion

This paper provides a framework to think about the interaction of minimum wage setting and federalism. Previous optimal minimum wage research focuses on one jurisdiction, which misses the implications of government competition. We first build on the theoretical model of Lee and Saez (2012), extending it to a context in which high-skilled labor is mobile across regions. We present sufficient conditions for the desirability of a binding minimum wage, from the perspective of both a local and central government. Though minimum wages are not optimal in the presence of even a linear tax system in our model, their prevalence in the real world warrants consideration from an optimal policy perspective.

We calibrate a two-region model of a federation and compare minimum wage policies across four types of government structures: decentralization, centralization with uniform policy, a US-style combination of both, and a social planner. In our baseline model, the desirability of centralized policy setting is increasing in mobility; when mobility is shut down completely, decentralization obtains the social planners solution. However, when all movers face positive moving costs, and regions are heterogeneous, then the decentralized outcome leads to higher total welfare than federal uniform policy. Furthermore, suboptimal policy takes the form of minimum wages that are too low, such that the tiered system we observe in the US is weakly preferred to either decentralization or centralization exclusively. In the tiered system, the central uniform policy lessens the externalities from horizontal competition by local governments leading both types of governments to set policies more closely aligned with the social planner. Decentralized and centralized policy setting are strategic complements. When we extend our framework to US states, we find that the tiered system closely matches the social planner's solution. Our results indicate that higher levels of government should not forbid lower levels from setting their own minimum wages. These results

are consistent across social welfare functions, although the effect of heterogeneity may depend on the progressivity of the social welfare function.

Though the US system performs well relative to other potential systems, its ability to efficiently redistribute income still falls short of a simple system of linear taxes. From the perspective of our calibrated model, the existence of minimum wages and redistributive taxes in the US remains a puzzle. Further research remains to be done on the types of models, economic or political in nature, which might give rise to the joint optimality of both types of policies. It is likely that the income tax is not set optimally due to political frictions, which lead to demand for other redistributive policies. For now, our results suggest the policy makers should pay attention to the mobility implications of minimum wage policies, as well as their interactions with the existing tax system.

7 References

Aaronson, Daniel, Eric French, Isaac Sorkin, and Ted To. "Industry dynamics and the minimum wage: a puttyclay approach." *International Economic Review* 59, no. 1 (2018): 51-84.

Allen, Stephen P. "Taxes, redistribution, and the minimum wage: a theoretical analysis." *The Quarterly Journal of Economics* 102, no. 3 (1987): 477-489.

Amior, Michael. "Education and geographical mobility: the role of the job surplus." CEP Discussion Paper, (2019).

Armenter, Roc, and Francesc Ortega. "Credible redistribution policy and skilled migration." *European Economic Review* 55, no. 2 (2011): 228-245.

Baicker, Katherine, Jeffrey Clemens, and Monica Singhal. "The rise of the states: US fiscal decentralization in the postwar period." *Journal of Public Economics* 96, no. 11-12 (2012): 1079-1091.

Banzhaf, H. Spencer, and B. Andrew Chupp. "Fiscal federalism and interjurisdictional externalities: New results and an application to US Air pollution." *Journal of Public Economics* 96, no. 5-6 (2012): 449-464.

Besley, Timothy, and Stephen Coate. "Centralized versus decentralized provision of local public goods: a political economy approach." *Journal of Public Economics* 87, no. 12 (2003): 2611-2637.

Blundell, Richard, and Thomas MaCurdy. "Labor supply: A review of alternative approaches." In *Handbook of labor economics*, vol. 3, pp. 1559-1695. Elsevier, 1999.

Boadway, Robin, and Katherine Cuff. "A minimum wage can be welfare-improving and employment-enhancing." *European Economic Review* 45, no. 3 (2001): 553-576.

Brown, Charles C., and Oates, Wallace E. "Assistance to the Poor in a Federal System." *Journal of Public Economics*, 32 (April 1987): 307-30.

Calabrese, Stephen M., Dennis Epple, Richard Romano. "Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout Competition. *The Review of Economic Studies*, vol. 79, no. 3, 2012, pp. 10811111.

Cadena, Brian C. "Recent immigrants as labor market arbitrageurs: Evidence from the minimum wage." *Journal of Urban Economics* 80 (2014): 1-12.

Caplan, Arthur J., Richard C. Cornes, and Emilson CD Silva. "Pure public goods and income redistribution in a federation with decentralized leadership and imperfect labor mobility." *Journal of Public Economics 77*, no. 2 (2000): 265-284.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. "The effect of minimum wages on low-wage jobs." *The Quarterly Journal of Economics* 134, no. 3 (2019): 1405-1454.

Clemens, Jeffrey, and Michael Wither. "The minimum wage and the Great Recession: Evidence of effects on the employment and income trajectories of low-skilled workers." *Journal of Public Economics* 170 (2019): 53-67.

Colas, Mark and Kevin Hutchinson. "Heterogenous Workers and Federal Income Taxes in a Spatial Equilibrium."

Cremer, Helmuth, and Pierre Pestieau. "Factor mobility and redistribution." In *Handbook of Regional and Urban Economics*, vol. 4, pp. 2529-2560. Elsevier, 2004.

Danziger, Eliav, and Leif Danziger. "The Optimal Graduated Minimum Wage and Social Welfare." In Transitions through the Labor Market: Work, Occupation, Earnings and Retirement, pp. 55-72. Emerald Publishing Limited, 2018.

Dixit, Avinash, and John Londregan. "Fiscal federalism and redistributive politics." *Journal of Public Economics* 68, no. 2 (1998): 153-180.

Dvorkin, Maximiliano A. "Fiscal Federalism and Optimal Income Taxes." (2017): 319-336.

Dube, Arindrajit, T. William Lester, and Michael Reich. "Minimum wage shocks, employment flows, and labor market frictions." *Journal of Labor Economics* 34, no. 3 (2016): 663-704.

Economides, George, and Thomas Moutos. "Minimum wages in the Presence of in-kind Redistribution." (2017).

Epple, Dennis, and Thomas Nechyba. "Fiscal decentralization." In *Handbook of Regional and Urban Economics*, vol. 4, pp. 2423-2480. Elsevier, 2004.

—, and Glenn J. Platt. "Equilibrium and local redistribution in an urban economy when households differ in both preferences and incomes." *Journal of Urban Economics* 43, no. 1 (1998): 23-51.

—, and Thomas Romer. "Mobility and redistribution." *Journal of Political Economy* 99, no. 4 (1991): 828-858.

Flinn, Christopher J. "Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates." *Econometrica* 74, no. 4 (2006): 1013-1062.

Gerritsen, Aart, and Bas Jacobs. "Is a minimum wage an appropriate instrument for redistribution?." (2016).

Gould, Elise. "Wages for workers with a high school degree or less rose the fastest over the last year." Economic Policy Institute. https://www.epi.org/blog/wages-for-workers-with-a-high-school-degree-or-less-rose-the-fastest-over-the-last-year/ (accessed March 1, 2018)

Gordon, Roger H. "An optimal taxation approach to fiscal federalism." *The Quarterly Journal of Economics* 98, no. 4 (1983): 567-586.

—, and Julie Berry Cullen. "Income redistribution in a federal system of governments." *Journal of Public Economics* 96, no. 11-12 (2012): 1100-1109.

Guesnerie, Roger, and Kevin Roberts. "Minimum wage legislation as a second best policy." *European Economic Review* 31, no. 1-2 (1987): 490-498.

Hamilton, Jonathan, and Pierre Pestieau. "Optimal income taxation and the ability distribution: Implications for migration equilibria." *International Tax and public finance* 12, no. 1 (2005): 29-45.

Hoynes, Hilary W., and Erzo FP Luttmer. "The insurance value of state tax-and-transfer programs." *Journal of Public Economics* 96, no. 11-12 (2012): 1110-1128.

Hungerbühler, Mathias, and Etienne Lehmann. "On the optimality of a minimum wage: New insights from optimal tax theory." *Journal of Public Economics* 93, no. 3-4 (2009): 464-481.

Janeba, Eckhard, and John Douglas Wilson. "Optimal fiscal federalism in the presence of tax competition." *Journal of Public Economics* 95, no. 11-12 (2011): 1302-1311.

Katz, Lawrence F., and Kevin M. Murphy. "Changes in relative wages, 19631987: supply and demand factors." *The Quarterly Journal of Economics* 107, no. 1 (1992): 35-78.

Keen, Michael J., and Christos Kotsogiannis. "Does federalism lead to excessively high taxes?." *American Economic Review* 92, no. 1 (2002): 363-370.

—. "Tax competition in federations and the welfare consequences of decentralization." *Journal of Urban Economics* 56, no. 3 (2004): 397-407.

Kennan, John, and James R. Walker. "The effect of expected income on individual migration decisions." *Econometrica* 79, no. 1 (2011): 211-251.

Kleven, Henrik, Camille Landais, Mathilde Muoz, and Stefanie Stantcheva. Taxation and Migration: Evidence and Policy Implications. No. w25740. National Bureau of Economic Research, 2019.

Kline, Patrick, and Enrico Moretti. "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs." *Annual Review of Economics* 6, no. 1 (2014): 629-662.

"Labor Force Statistics from the Current Population Survey." Bureau of Labor Statistics. https://www.bls.gov/cps/ (accessed March 1, 2018)

Ladd, Helen F., and Doolittle, Fred C. "Which Level of Government Should Assist the Poor?" *National Tax Journal*, 35 (September 1982): 323-36.

Lavecchia, Adam. "Minimum Wage Policy with Optimal Taxes and Unemployment." McMaster University Department of Economics Working Paper 2019-03 (2019).

Lee, David, and Emmanuel Saez. "Optimal minimum wage policy in competitive labor markets." *Journal of Public Economics* 96, no. 9-10 (2012): 739-749.

Lockwood, Ben. "Distributive politics and the costs of centralization." *The Review of Economic Studies* 69, no. 2 (2002): 313-337.

Luttmer, Erzo FP. "Does the minimum wage cause inefficient rationing?." *The BE Journal of Economic Analysis & Policy* 7, no. 1 (2007).

Moretti, Enrico. "Real wage inequality." *American Economic Journal: Applied Economics* 5, no. 1 (2013): 65-103.

Neumark, David. "The Econometrics and Economics of the Employment Effects of Minimum Wages: Getting from Known Unknowns to Known Knowns." NBER Working Paper, (2018).

—, and William L. Wascher. "Minimum wages and employment." Foundations and Trends in Microeconomics 3, no. 12 (2007): 1-182.

Oates, Wallace E. Fiscal Federalism. New York: Harcourt Brace Jovanovich, 1972.

—. "An economist perspective on Fiscal federalism-in: The Political Economy of Fiscal Federalism (ed. by. WE Oates), 1977." (1977): 3-19.

Phillips, David C. "Do low-wage employers discriminate against applicants with long commutes? Evidence from a correspondence experiment." *Journal of Human Resources* (2018): 1016-8327R.

Roback, Jennifer. "Wages, rents, and the quality of life." *Journal of political Economy* 90, no. 6 (1982): 1257-1278.

Rohlin, Shawn M. "State minimum wages and business location: Evidence from a refined border approach." *Journal of Urban Economics* 69, no. 1 (2011): 103-117.

Rosen, Sherwin. (1979). Wage-based indexes of urban quality of life. Baltimore: Johns Hopkins University.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 9.0 American Community Survey. Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D010.V9.0

Stigler, George J. "The Tenable Range of Functions of Local Government." In Federal Expenditure Policy for Economic Growth and Stability, by the Joiny Economic Committee. Washington: Government Printing Office, 1957.

Stiglitz, Joseph E. "Self-selection and Pareto efficient taxation." *Journal of Public Economics* 17, no. 2 (1982): 213-240.

Surez Serrato, Juan Carlos, and Owen Zidar. "Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms." *American Economic Review* 106, no. 9 (2016): 2582-2624.

Wildasin, David E. "Labor-market integration, investment in risky human capital, and fiscal competition." *American Economic Review* 90.1 (2000): 73-95.

—. "Income distribution and redistribution within federations." *Annales d'Economie et de Statistique* (1997): 291-313.

Wilson, John D. "A theory of interregional tax competition." *Journal of Urban Economics* 19, no. 3 (1986): 296-315.

—. "Trade, Capital Mobility, and Tax Competition." *Journal of Political Economy.* 95 (August 1987): 835-56.

—. "Tax competition in a federal setting." *Handbook of Multilevel Finance* (2015): 264.

—. "Optimal linear income taxation in the presence of emigration." *Journal of Public Economics*

18, no. 3 (1982): 363-379.

Zodrow, George R., and Peter Mieszkowski. "Pigou, Tiebout, property taxation, and the underprovision of local public goods." *Journal of Urban Economics* 19, no. 3 (1986): 356-370.

A Optimal Linear Taxes and Minimum Wage

When the federal government has access to linear income taxes, it will always prefer linear taxes to a combination of a minimum wage and linear taxes under the assumptions of our calibrated model.

Consider a binding minimum wage and tax policy $\bar{w}, t \geq 0$. Under the assumptions of our model, mainly a perfectly competitive firm, this set of policy will induce additional unemployment over the competitive equilibrium so that $l(\bar{w}, t)$ low-skilled workers are employed. The minimum wage reduced labor demand while the tax reduced labor supply. The workers earn $\bar{w}$ and high-skilled workers earn $w_h(l(\bar{w}, t))$. The high-skilled wage depends on our structural parameters and the stock of low-skilled employed workers. Now consider $t^\star$ such that $l(0, t^\star) = l(\bar{w}, t)$. It must be the case that $t^\star \geq t$, as a binding minimum wage always reduces low-skilled employment, as long as $l(\bar{w}, t) > 0$. Since l is the same, w_l, w_h, and total output are equal across the policies. However, $t^\star \geq t$ gives that the tax alone raises more revenue. The additional tax revenue is, on net, redistributed to lower consumption agents through the demogrant, increasing total welfare since the social welfare function is concave. As the tax increases, low-skilled labor supply drops, increasing the low-skilled wage and decreasing the high-skilled wage; when it reaches 70%, only 34% of the low-skilled workers are employed, and the low-skilled wage is higher. We may be worried that since the low-skilled face a working cost, they can be high-income but low-consumption and the additional redistribution is harmful for social welfare. However, at that point, income tax collected is redistributed on net to the unemployed from the employed. Since the consumption of the unemployed is always bounded above by the consumption of the employed, even with this tax and demogrant, the tax policy always redistributes to the lower consumption agents.

We need to show that $t^\star$ exists. As long as the low-skilled labor supply is not perfectly inelastic, it will exist because as $l \to 0$, the marginal product goes to infinity, but increasing the tax rate will weakly lower labor supply. When the labor supply is perfectly inelastic (because the low-skilled workers do not face costs from working), then the government can implement a levelling tax, since high-skilled labor supply is also perfectly inelastic to the economy. The value of the levelling tax is maximized when output and labor demand is maximized, i.e. when there is no binding minimum wage.

So the government will always prefer only a linear income tax when both a linear income tax and minimum wage are available in our model.