Geographic Mobility and Redistribution*

Daniele Coen-Pirani [1]

University of Pittsburgh

January 24, 2021

Abstract

I study the effect of progressive taxation on internal migration and welfare using a quantitative dynamic model of geographic mobility. The model, which is analytically tractable, predicts that a more progressive tax-transfer scheme reduces internal migration rates. The magnitude of this relationship is consistent with reduced-form evidence for OECD countries. The internal migration channel contributes to significantly lower optimal tax progressivity relative to the one-location version of the economy. The optimal sequence of tax progressivity along the transition features a relatively high degree of tax progressivity early on, followed by a declining path of tax progressivity over time.

Keywords: Geographic mobility, migration, dynamic optimization, progressive taxation, redistribution, welfare, transitional dynamics.

JEL: C61, D15, H20, J22, J24, R23

*Manuscript received December, 2018; Revised September, 2020;.Accepted, December, 2020.

*I would like to thank the Editor of this journal and its anonymous Referees for useful and detailed comments and suggestions. Thanks also to Chris Flinn, Laurence Ales, Ali Shourideh, Giuseppe Bertola, Marla Ripoll, Omer Moav, Yona Rubinstein, Giulio Zanella, Fabian Lange, Max Dvorkin, Arpad Abraham, Rodi Manuelli, Axelle Ferriere, Jean-Francois Richard, Grey Gordon, Marco Casari, and seminar participants at various institutions and conferences. The usual disclaimer applies.

1 Introduction

A large literature studies the effect of taxes and tax progressivity on labor supply, both positively (Prescott, 2004) and normatively (Diamond and Saez, 2011). A common argument is that high and progressive taxes may reduce labor supply, especially along the intensive margin. In this paper I study the effect of progressive national (in the U.S. Federal) taxes and transfers, on the *location* - rather than the quantity - of a household's labor supply. The basic idea is that, if migration is costly, a more progressive national tax-transfer system reduces individuals' incentives to migrate internally to locations where they are idiosyncratically more productive. I embed this mechanism in a new tractable model of internal migration. The estimated model is used to quantify the empirical relevance of this distortion and to study the welfare implications of tax reform.

The new model of internal migration is rich, as it allows for both ex-ante and ex-post heterogeneity in terms of age, moving costs, and labor income, while delivering closed-form solutions for individual-level migration and labor supply behavior. I use these to show that a higher degree of income tax progressivity tends to reduce migration rates. A higher degree of tax progressivity reduces the returns from migration by taxing away a portion of the earnings growth associated with moving to a new location. This reduction in the after-tax benefit of moving reduces the incentives to migrate in the presence of a positive moving cost. The magnitude of the migration cost parameter, therefore, has important implications for the policy counterfactuals. The moving cost and other key parameters of the model are identified by the frequency of interstate migration by age, and by the evolution of the cross-sectional mean and variance of earnings over the life-cycle. These moments are computed using data from the American Community Survey.

I test the quantitative predictions of the model on moments that are not targeted in the estimation and that are computed using two different datasets. First, I compare its predictions with estimates of the growth in earnings associated with interstate migration. These estimated gains are based on household-level panel data from the Survey of Income and Program Participation. The model correctly predicts that earnings gains from internal migration are proportionately larger for more educated households and tend to decline with age. Second, I evaluate the model's ability to account for the magnitude of the cross-sectional relationship between newly constructed measures of tax progressivity and internal migration rates for OECD countries. In the data tax progressivity is negatively associated with geographic mobility. The model provides an accurate quantitative account of this relationship.

I characterize analytically the welfare function and the trade-offs associated with higher tax progressivity using a simplified version of the model. In principle, the internal migration channel has ambiguous implications for optimal tax progressivity. A more progressive tax system distorts incentives to move towards locations in which individuals are more productive. However, migration is associated with increased levels of earnings inequality and higher tax progressivity mitigates its effect on consumption inequality.

The quantitative analysis with the full model shows that optimal tax progressivity is smaller than in an equivalent model with one location and no scope for mobility. Keeping constant the average tax rate, the average income-weighted marginal tax rate is about 8 percentage points lower in the benchmark model than in an otherwise similar one-location economy. Taking into account the transition from the initial to the final steady state contributes to make the optimal degree of tax progressivity significantly higher than if the planner cared only about steady state welfare. This is because at the time of the reform the productivity level of old agents is fixed. Consequently, a relatively high level of tax progressivity is less distortionary in the short-run than in the long-run when the migration choices of all agents are based on the new policy regime.

This intuition suggests that an optimal *time-varying* tax progressivity policy would involve a relatively high degree of tax progressivity in the early phase of the tax reform - when old agent's productivity levels are fixed - followed by a declining path towards the new steady state. I exploit the analytical tractability of the model to compute this policy and verify that this is indeed the case. The welfare gain of optimal time-varying policies is 1.09 percent in consumption equivalent units, relative to 0.71 for a tax reform involving a constant level of tax progressivity.

This paper adds to the recent literature on the effects of tax progressivity and subsidies on human capital investment. Papers in this literature focus on investment in either formal schooling or training.[1] Absent the geographic mobility element, my model would be similar to those in Benabou (2002) and Heathcote et al (2017, 2020a). Their functional form assumptions allow me to derive closed-form solutions to the agents' dynamic migration problem. Differently from papers in this literature, I focus on internal migration, a costly investment that can be observed in many publicly available datasets (Schultz, 1961).

Gentry and Hubbard (2004) explore empirically the effect of the level and progressivity of taxation on job-to-job transitions in the U.S. and find evidence that a more progressive tax system is associated with less job mobility. This is consistent with my empirical evidence across OECD countries. Hassler et al (2005) provide suggestive evidence that countries with more generous unemployment insurance systems are characterized by lower internal geographic mobility. They use a politico-economic model of endogenous unemployment insurance and migration to account for this fact, but, differently from my approach, their focus is not quantitative.

The importance of transitional dynamics considerations for the calculation of optimal (Ramsey) tax progressivity has been emphasized, in the context of different models, by a number of authors, e.g. Domeji and Heathcote (2004), Bakis et al (2015) and Krueger and Ludwig (2016) among others. My paper departs from these by explicitly considering the optimal path of tax progressivity over the transition. In a related paper, Dyrda and Pedroni (2018) study optimal time-varying linear capital and labor income taxes in the context of the Aiyagari model using a numerical approach similar to mine.

A number of recent papers study optimal spatial taxation in the context of static equilibrium models. Albouy (2009), Eeckout and Guner (2015) and Colas and Hutchinson (2020) point out that Federal taxation of *nominal* income might distort the spatial allocation of

[1] A partial list of papers includes Manovskii (2002), Bovenberg and Jacobs (2005), Erosa and Koreshkova (2007), Abbott et al (2019), Krueger and Ludwig (2013, 2016), Guvenen et al (2014), Badel et al (2020), Stantcheva (2017), Mustre-del-Rio and Hawkins (2017), Koeniger and Prat (2018), among others.

workers across locations. The argument advanced in these papers hinges crucially on the fact that locations are ex-ante *heterogeneous* in terms of productivity and amenities. In this paper, instead, I focus on the dynamic effect of national tax progressivity on gross internal migration flows of workers across local labor markets, instead of its effect on the geographic distribution of population across heterogeneous locations. In my model agents make *dynamic* mobility choices and progressive Federal taxation reduces the net gain from moving even if all locations are ex-ante *identical*. The tax distortion I study is the effect of progressive taxation on workers' incentives to locate in the labor market where they are idiosyncratically more productive.[2]

The paper is also related to the recent papers by Caliendo et al (2019) and Artuc et al (2010) on spatial and sectoral labor reallocation following trade shocks. These papers follow Kennan and Walker (2011) and generate gross flows of labor across sectors and regions through extreme-value location-specific shocks, which are usually interpreted as preference or moving cost shocks. While my model also relies on extreme-value shocks for analytical convenience, it interprets them as permanent shocks to households' productivity growth in each location. Consequently, in the model as in the data, geographic mobility is primarily motivated by job-related reasons, such as prospect of income growth.[3] Finally, a number of recent papers focus on individuals' incentives to migrate across jurisdictions in order escape from relatively high-tax locations (Akcigit et al (2016) and Moretti and Wilson (2017)) or take advantage of more generous state and local welfare systems (Gelbach, 2004).[4] The tax distortion I consider, instead, is due to the national – in the U.S. Federal – tax system instead of tax differences across political sub-units.

The rest of the paper is organized as follows. Section 2 introduces the model and Section

[2]Other related contributions on the topic of location-specific taxes are Fajgelbaum and Gaubert (2020) and Albouy et al (2019).

[3]The paper is also related to the growing migration literature in many areas of economics, such as urban (Ahlfeldt et al (2015), Diamond, 2016), labor (Auray et al, 2017, Gemici 2017, Monras, 2017), macro (Coen-Pirani, 2010, Bayer and Juessen, 2012, Lkhagvasuren, 2014, Kaplan and Schulhofer-Wohl, 2017), and development (Desmet and Rossi-Hansberg, 2017 and Lagakos et al (2017)).

[4]See also Fajgelbaum et al (2019)'s contribution arguing that heterogeneity in taxes across states is a cause of misallocation.

3 characterizes its analytical solution. The description of the model's estimation and data fit is contained in Section 4. Section 5 tests some of the quantitative implications of the model using moments that were not targeted in the estimation. Optimal policies and welfare are discussed in Section 6. Finally, Section 7 concludes and discusses future work. All proofs, details about the data, and discussions of extensions of the model are contained in the online Appendix.

2 Model

Geography, Technology, and Policy The economy is comprised of a finite number of local labor markets indexed by $k = 1, 2, .., K$. Local economies are assumed to be ex-ante identical in terms of their aggregate characteristics, such as amenities and productivity.[5] Thus, the reasons for migration from one labor market to another are purely idiosyncratic. Each local economy produces an homogenous good using the same constant returns to scale production function. The only production input is represented by efficiency units of labor whose marginal product is normalized to one without loss of generality. Time is discrete and infinite, starting at $t = 1$. For the rest of the analysis, the only source of exogenous variation across time in the economy pertains to the aggregate policy variables $\{\tau_t, \lambda_t, G_t\}_{t=1}^{\infty}$, where G_t denotes the government's public good provision while τ_t and λ_t are the parameters of the tax-transfer scheme that maps a household's market income y into its after-tax and transfer income $\widetilde{y}_t$:[6]

$$\widetilde{y}_t = \lambda_t y^{1-\tau_t}, \tag{1}$$

[5] Online Appendix 6.2 discusses the case of ex-ante heterogeneous locations and shows that the latter collapses to the economy with homogeneous locations under specific restrictions on how local prices and amenities vary in the cross-section of locations.

[6] The presence of a public good does not affect the positive properties of the model because it enters additively in utility (see equation 3 below). However, its presence affects the socially optimal choice of tax progressivity $\{\tau_t\}_{t=1}^{\infty}$, as shown by Heathcote et al (2017).

In this tax-transfer system, which has been employed by Benabou (2002) and Heathcote et al (2017) among others, the household pays net taxes $T_t(y) = y - \widetilde{y}_t$ and the parameter τ_t indexes tax progressivity while λ_t indexes the overall level of taxation. For example, when $\tau_t = 0$, then marginal and average taxes are the same $(T_t'(y) = T_t(y)/y)$ and the tax system features a proportional income tax with rate $1-\lambda_t$. By contrast, when $\tau_t = 1$ the government taxes away the entire marginal unit of income earned so $T_t'(y) = 1$.

Demographic Structure, Timing, and Migration Choices The economy is populated by a continuum of finitely-lived households of measure one. Households live from age $a = 1$ to $\overline{a}$. A measure $1/\overline{a}$ of households is born in each location at the beginning of each period. A household belongs to one of $\overline{r}$ possible types, indexed by $r = 1, 2, .., \overline{r}$. Household type is determined at birth and does not change over time. There is a measure ω_r of households of type r, with $\sum_r \omega_r = 1$. The initial distribution of households by age and type across locations is symmetric, so that a measure $\omega_r / (\overline{a}K)$ of households of type r and age a resides in each location at $t = 1$. A household begins life at age $a = 1$ with initial productivity $\exp(\alpha_{1,r} + z)$, where z is drawn from the density $f(z|1, r)$. The latter is exogenous, assumed to be the same in all locations and time periods, and allowed to vary by type r. Household productivity evolves exogenously with aging, according to the deterministic function $\alpha_{a,r}$ and endogenously, through migration.

At each age a the timing of actions is as follows. At the beginning of each period, a household characterized by labor productivity $\exp(z + \alpha_{a,r})$ resides in one of the economy's K locations. Then, migration occurs. Some households are exogenously relocated while others are in the position of making a moving choice. Migration of either type involves a utility cost κ_r. The household then supplies labor ℓ and earns labor income $y = \ell \exp(\alpha_{a,r} + z')$, where z' denotes the household's idiosyncratic level of z *after* migration. Then, redistribution takes place leaving the household with disposable income $\widetilde{y}$, given by equation (1), which is assumed to be entirely consumed, $c = \lambda_t y^{1-\tau_t}$.[7]

[7]The assumption that an agent cannot save or borrow reflects a form of asset market incompleteness. Differently from Heathcote et al (2017), this assumption is not without loss of generality due to the explicit

Migration incentives are driven by the location-dependent evolution of household productivity. Specifically, a household with current idiosyncratic log productivity z observes, at the beginning of period t, the evolution of its efficiency in *all* locations k. Its log productivity in location k is assumed to evolve as follows during period t:[8]

$$z'_k = \rho_r z + \eta_r \varepsilon_k \text{ for all } k = 1, .., K, \tag{2}$$

where the parameter $\rho_r \leq 1$ determines the persistence of productivity over time and the parameter $\eta_r < 1$ governs the importance of location-specific shocks $\boldsymbol{\varepsilon} \equiv \{\varepsilon_k\}_{k=1}^K$. The shocks ε_k are assumed to be distributed according to a type-1 extreme value distribution (Rust, 1987).[9] They are independently drawn both over time for a given household as well as in the cross-section of households. I allow shocks drawn from the household's current home location, denoted by h, to have a different distribution from shocks drawn elsewhere due, for example, to its superior information about the home location's labor market. Specifically, the distribution of ε_h is assumed to be a type 1 extreme-value with unit scale parameter and location parameter $\ln \delta_r$. Instead, the distribution of ε_k for $k \neq h$ is assumed to be a type 1 extreme-value with unit scale parameter and location parameter equal to zero. A value $\delta_r > 1$ implies that, on average, a shock drawn from the home location is larger than shocks drawn from the other locations.[10]

I distinguish between endogenous and exogenous migration in the following simple way.

consideration of age as an important determinant of household choices. In my setting, on average, a young agent would like to save in anticipation of aging and a diminished scope for migration later in life.

[8] I don't include a household or age-specific subindex to the variables in equation (2) not to complicate the notation. It should be kept in mind, however, that both z_k and ε_k vary idiosyncratically by household and over time for each household.

[9] It is straighforward to allow for some correlation among shocks, but their common component would not matter for migration decisions. The assumption that shocks are distributed as extreme-value is mostly for analytical convenience. It is not necessary to show that tax progressivity has a negative effect on geographic mobility.

[10] An interpretation is that a household of type r makes δ_r draws in its current location and only one in each of the away locations. The maximum of δ_r independent extreme-value draws is still distributed as extreme-value, so the distribution of shocks from the home location for a type r agent is an extreme-value with mean $\ln \delta_r + \gamma$, where γ is Euler's number. A draw from each of the other locations has a mean equal to γ.

With probability θ_r, upon observing the vector $\boldsymbol{\varepsilon}$, a household is able to select the location that provides the highest present discounted value of utility going forward. With probability $1-\theta_r$, instead, the household is exogenously relocated to one of the other $K-1$ locations with equal probability. If the household moves to a location k, either by choice or exogenously, it pays a utility cost κ_r and its productivity at the beginning of next period becomes $\exp\left(\alpha_{a+1,r}+z'_k\right)$.

Households' Preferences and Recursive Formulation Agents maximize their expected discounted utility net of moving costs. The discount factor is denoted by the parameter $\beta<1$. The household's static utility function is:

$$u\left(c,\ell,G\right)=\ln c-\zeta^{-1}\ell^{\zeta}+\chi\ln G, \tag{3}$$

where ℓ denotes labor supply, the parameter ζ determines the Frisch elasticity of labor supply (which equals $\left(1-\zeta\right)^{-1}$), and χ is the utility weight of the public good G, provided by the national government. The household's optimal labor supply maximizes equation (3) subject to the budget constraint $c=\lambda_t\left(\ell\exp\left(\alpha_{a,r}+z'\right)\right)^{1-\tau_t}$. It is straightforward to show that it is given by:[11]

$$\ell_t^*=\left(1-\tau_t\right)^{\frac{1}{\zeta}}. \tag{4}$$

Thus, a higher degree of tax progressivity reduces labor supply, as in Benabou (2002) and Heathcote et al (2017). Replacing ℓ_t^* back into (3), the static indirect utility function takes the log-linear form:

$$u_t^*\left(a,z';r\right)=\overline{u}_t^*\left(a;r\right)+\left(1-\tau_t\right)z', \tag{5}$$

[11] The logarithmic specification of utility in (3) simplifies the dynamic programming problem faced by each agent. While convenient, it has some well-known limitations (see e.g. Heathcote et al (2017)). For example, due to offsetting income and substitution effects, the labor supply choice is independent of an agent's productivity and of the overall level of taxation λ_t. In Online Appendix 6.1 I use a simplified version of the model to discuss the implications of a CRRA utility function for the effect of tax progressivity on migration .

where:

$$\overline{u}_t^* (a;r) \equiv \ln \lambda_t + (1-\tau_t) \ln \ell_t^* + (1-\tau_t)\, \alpha_{a,r} - \zeta^{-1} (\ell_t^*)^{\zeta} + \chi \ln G_t. \tag{6}$$

Each household faces a dynamic optimization problem that involves the choice of location. Let $V_t(a, z', k; r)$ denote the household's conditional value function. This is the maximum remaining lifetime utility attainable by a household of age a, with idiosyncratic productivity z', who is located in k *after* migration:

$$V_t(a, z', k; r) = \begin{cases} u_t^*(a, z'; r) \text{ if } a = \overline{a} \\ u_t^*(a, z'; r) + \beta E_{\boldsymbol{\varepsilon}'} \left[P_{t+1}(a+1, z', \boldsymbol{\varepsilon}', k; r) \right] \text{ if } a < \overline{a} \end{cases}, \tag{7}$$

where $E_{\boldsymbol{\varepsilon}'}[.]$ denotes the expectation taken with respect to the distribution of location-specific $t+1$ shocks, $\boldsymbol{\varepsilon}'$. The function P_{t+1} on the right-hand side of (7) represents the unconditional value function in $t+1$. The unconditional value function, denoted by $P_t(a, z, h, \boldsymbol{\varepsilon}; r)$, represents the lifetime utility attainable by a household of age a who, before migration, is located in h, has idiosyncratic productivity z, and observes the vector of location-specific shocks $\boldsymbol{\varepsilon}$:

$$\begin{aligned} P_t(a, z, \boldsymbol{\varepsilon}, h; r) &= \theta_r \max_k \{ V_t(a, \rho_r z + \eta_r \varepsilon_k, k; r) - I_{hk} \kappa_r \} + \\ &+ \frac{1-\theta_r}{K-1} \sum_{k \neq h} \{ V_t(a, \rho_r z + \eta_r \varepsilon_k, k; r) - \kappa_r \} \end{aligned} \tag{8}$$

where I_{hk} is an indicator function that takes a value of 1 if and only if $k \neq h$ and zero otherwise. Notice that the term multiplying θ_r in the first row of equation (8) represents the value of endogenous migration, while the term multiplying $(1-\theta_r)$ in the second row represents the expected value of exogenous migration. Let $M_t(a, \boldsymbol{\varepsilon}, h, k; r)$ denote the migration decision rule for households that are not exogenously relocated. In writing it in this way I anticipate the fact that it does not depend on z (see Proposition 2). Specifically, $M_t(a, \boldsymbol{\varepsilon}, h, k; r) = 1$ if a household of age a with a vector of shocks $\boldsymbol{\varepsilon}$ moves voluntarily at the beginning of time t from its current location h to location k, and $M_t(a, \boldsymbol{\varepsilon}, h, k; r) = 0$ if the household chooses not move to k. Notice that, by definition, $\sum_{k=1}^{K} M_t(a, \boldsymbol{\varepsilon}, h, k; r) = 1$.

Competitive Equilibrium Given the sequences $\{\tau_t, G_t\}_{t=1}^{\infty}$ and the pre-migration distributions $f(z|a, r)$ of household productivity by age and type at the beginning of $t = 1$, a competitive equilibrium for this economy is comprised of: a sequence $\{\lambda_t\}_{t=1}^{\infty}$; sequences of post-migration densities of household labor productivity by age and type $\{f_t^p(z'|a, r)\}_{t=1}^{\infty}$; sequences of unconditional and conditional value functions $\{P_t(a, z, \boldsymbol{\varepsilon}, h; r), V_t(a, z', k; r)\}_{t=1}^{\infty}$; sequences of decision rules $\{M_t(a, \boldsymbol{\varepsilon}, h, k; r)\}_{t=1}^{\infty}$ for geographic mobility; sequences of decision rules for labor supply $\{\ell_t^*\}_{t=1}^{\infty}$ such that:

1) The value functions $\{P_t(a, z, \boldsymbol{\varepsilon}, h; r), V_t(a, z', k; r)\}_{t=1}^{\infty}$ and decision rules $\{M_t(a, \boldsymbol{\varepsilon}, h, k; r)\}_{t=1}^{\infty}$ represent the solution to the agent's dynamic optimization problem (7)-(8), with ℓ_t^* taking the form in (4).

2) The value of λ_t is consistent with the government's balanced budget for all $t \geq 1$:

$$G_t = \frac{1}{\bar{a}} \sum_{a=1}^{\bar{a}} \sum_{r=1}^{\bar{r}} \omega_r \int \left(\ell_t^* \exp(\alpha_{a,r} + z') - \lambda_t \left(\ell_t^* \exp(\alpha_{a,r} + z') \right)^{1-\tau_t} \right) f_t^p(z'|a, r)\, dz'. \tag{9}$$

3) The densities $\{f_t^p(z'|a, r)\}_{t=1}^{\infty}$ are generated by the transition equation for productivity (2) combined with the optimal decision rules $\{M_t(a, \boldsymbol{\varepsilon}, h, k; r)\}_{t=1}^{\infty}$, starting from the initial densities $f(z|a, r)$.

The definition of competitive equilibrium reflects the model's symmetry in that all locations are characterized by the same distribution of population and productivity by age and type. Notice that the general equilibrium dimension of this economy stems from the government's budget constraint (9). At each point in time t, two of the variables in the triple (τ_t, G_t, λ_t) are given, and the remaining one has to satisfy equation (9). A *stationary equilibrium* of the model refers to a situation in which (τ, G) and all the endogenous objects that comprise a competitive equilibrium are constant over time. I will focus on a stationary equilibrium when estimating the model's parameters.

Discussion of modeling choices Before characterizing the households' decision rules, it is useful to briefly discuss the environment outlined above. The model describes households' geographic relocation in response to *idiosyncratic* productivity shocks. This formulation captures, realistically, the idea that an important portion of moves across local labor markets is due to job-related reasons. For example, the Current Population Survey asks respondents to select a reason for geographic moves. In 2000-2007 data, "new job or job transfer" is the single most-frequently selected answer among head of households ages 25–59 who moved across state lines in the previous year. A new job and other job-related reasons jointly account for about 42 percent of all interstate moves. In addition to being realistic, the presence of exogenous relocation allows the model to match the fact that the magnitude of income gains associated with migration is decreasing over the life cycle (see Section 5.1).

The model environment describes gross flows of workers across homogeneous locations. The assumption of homogeneous locations is a convenient simplification that allows me to focus on gross migration flows driven by idiosyncratic productivity shocks.[12] This approach complements much of the existing literature on internal migration (e.g. Diamond, 2016), which emphasizes idiosyncratic preference shocks as a source of gross migration flows. In the latter, localities are ex-ante different in terms of aggregate productivity and amenities but all workers with the same observable skills (e.g. education) earn the same wage if they work in the same labor market. In reality, observable skills such as those associated with schooling explain only a relatively small fraction of wage differences among individuals in the same local labor market. In my model instead, all labor markets are ex-ante identical from an aggregate perspective, but there is a substantial heterogeneity in wages among workers with the same observable skills and age operating within a labor market.[13]

[12] In the U.S. economy, gross migration flows are very large relative to net flows. For example, in the American Community Survey sample I use to estimate the model's parameters, the average U.S. state experiences combined inflows and outflows of nine households for each household it gains or loses in net terms in the course of one year. Moreover, in Coen-Pirani (2010) I show that workers who enter (gross inflows) and exit (gross outflows) the same local labor market are observationally very similar in terms of age, industry and occupation.

[13] In Online Appendix 6.2 I show that the model with homogeneous locations may be derived from one with ex-ante heterogeneous locations in which housing rents fully capitalize differences in local wages, taxes and amenities.

While the model above might also describe job-to-job transitions within a local labor market, I favor the migration interpretation because the household suffers the moving cost κ_r when it relocates. The internal migration literature (e.g. Kennan and Walker, 2011) has estimated sizeable moving costs and the findings of my paper are consistent with this evidence (see Section 4.5).

One of the paper's innovations is modelling the dynamics of location-specific productivity at the household level. In this specification, productivity gains in a location persist over time independently of whether the agent remains in that location or migrates to another one. In a way, locations provide idiosyncratic opportunities for a household to become more productive by moving there. This increase in productivity, which might depreciate over time depending on the persistence parameter ρ_r, is reflected in the household's human capital, i.e. it is portable across locations. An advantage of this specification is that each household is characterized by only one state variable, instead of K location-specific levels of productivity.

3 Analytical Derivation

In this section I characterize analytically the value function, the geographic mobility decision rule, and the dynamic evolution of household productivity. The results of this section provide valuable intuition on the main mechanisms of the model and form the basis for its structural estimation in Section 4.

3.1 Value Function and Decision Rules

Propositions 1 and 2 present, respectively, the value function that solves the Bellman equation (7)-(8) and the decision rule for migration. These closed-form solutions allow me to characterize analytically, in Proposition 3, how tax progressivity affects migration rates.

Proposition 1 (Value function) *Given policies* $\{\tau_t, G_t, \lambda_t\}_{t=1}^{\infty}$ *, the unique value function* *that solves the dynamic programming problem (7)-(8) takes the following form:*

$$V_t(a, z', k; r) = v_t^0(a; r) + v_t^1(a; r) z' \text{ for all } k, \tag{10}$$

where for $a < \overline{a}$:

$$v_t^1(a; r) = 1 - \tau_t + \sum_{k=1}^{\overline{a}-a} (\beta\rho_r)^k (1 - \tau_{t+k}), \tag{11}$$

while $v_t^1(\overline{a}; r) = 1 - \tau_t$. The age-dependent term $v_t^0(a; r)$ is defined recursively by:

$$\begin{aligned} v_t^0(a; r) &= \overline{u}_t^*(a; r) + \beta v_{t+1}^0(a+1; r) - \beta(1-\theta_r)\kappa_r \\ &\quad + \beta v_{t+1}^1(a+1; r)\eta_r(\gamma - \theta_r \ln(p_{t+1}(a+1; r)/\delta_r)), \end{aligned} \tag{12}$$

starting from $v_t^0(\overline{a}; r) = \overline{u}_t^*(\overline{a}; r)$. The term $p_{t+1}(a+1; r)$ in equation (12) represents the probability of choosing to remain in the same location at age $a+1$ and is formally defined in equation (14) below.

According to Proposition 1, the value function is linear in the household's log productivity.[14] The utility value of higher productivity depends on the age and time-dependent coefficient $v_t^1(a; r)$. Intuitively, the latter declines with the household's age and with the degree of tax progressivity that it faces in its remaining life span. Using the expression for the value function, it is possible to characterize the geographic decision rule of households in this economy.[15]

Proposition 2 (Geographic mobility) *A household of age a with shocks $\boldsymbol{\varepsilon}$ chooses to*

[14] The existence of a closed-form solution is due to the logarithmic specification of static utility and not to the extreme-value distribution of the ε_k shocks. A different distributional assumption on the ε_k's would be reflected in the term $v_t^0(a; r)$.

[15] Notice that the parameter λ, which determines the level of taxes, has no effect on moving choices because of offsetting income and substitution effects associated with logarithmic utility. An analogous result holds, for example, in Heathcote et al (2017, Proposition 2) where the skill investment depends only on tax progressivity and not on the level of taxes.

move from location h to a different location $k \neq h$ *(*$M_t(a, \boldsymbol{\varepsilon}, h, k; r) = 1$*) if and only if:*

$$\varepsilon_k - \varepsilon_h > \frac{\kappa_r}{v_t^1(a;r)\,\eta_r} \tag{13}$$

and $\varepsilon_k = \max_{l \neq h} \varepsilon_l$.

Notice that, absent moving costs ($\kappa_r = 0$), the household would simply pick the location with the highest growth rate of labor income, ε_k. A positive moving cost implies that the household might choose to remain in the same location even if the growth rate of its productivity in that location is smaller than in the rest of the economy. Conditional on moving to a different location, a household chooses the location with the highest growth rate of its productivity.

When equation (13) does not hold a household chooses not to migrate. The probability that this happens can be computed using the distribution of the random variable $\max_{l \neq h} \varepsilon_l - \varepsilon_h$. The assumption that the $\boldsymbol{\varepsilon}$ shocks are independently distributed as type 1 extreme-value random variables implies that $\max_{l \neq h} \varepsilon_l - \varepsilon_h$ is logistic with location parameter $\ln\left((K-1)\,\delta_r^{-1}\right)$ and unit scale parameter. It follows that the probability that a household of age a chooses to remain in the same location (when given the opportunity to choose) is:

$$p_t(a;r) = \frac{1}{1 + (K-1)\,\delta_r^{-1}\exp\left(-\kappa_r\left(\eta_r v_t^1(a;r)\right)^{-1}\right)}. \tag{14}$$

Recall that a household might migrate for exogenous reasons at the rate $1 - \theta_r$. Thus, the overall probability of moving away from a location is $1 - \theta_r p_t(a;r)$. The following proposition presents some important implications of the theory for the rate of geographic mobility.

Proposition 3 (Migration probabilities) *Assume that migration costs are positive and finite,* $\kappa_r \in (0, +\infty)$. *Then:*

1. A household's probability of migration declines with its age a.

2. Given a household's age a*, its migration probability declines with the degree of tax progressivity* $\{\tau_{t+k}\}_{k=0}^{\overline{a}-a}$ *that it faces in its remaining working life.*

3. The impact of tax progressivity on a household's probability of migration first rises and then falls with the moving cost parameter κ_r.

4. Exogenous relocations account for a growing fraction of all moves as an individual ages.

The intuition for these results is relatively straightforward. First, migration rates decline with age because the horizon over which the household can take advantage of the benefits of moving shrinks as it gets older. Second, the effect of tax progressivity on mobility is negative because the moving cost κ_r is positive and is not tax deductible. Thus, while the income gain from moving is subject to progressive taxation, the utility gain from *not* moving is not taxed. This asymmetry is at the core of the negative effect of τ_{t+k} on mobility.[16] Third, the effect of τ_{t+k} on mobility depends on the moving cost κ_r in a non-monotonic fashion. With costless mobility ($\kappa_r = 0$) the household always moves to the location that provides the highest before-tax income growth, independently of the degree of tax progressivity. As the moving cost increases, tax progressivity becomes distortionary, but eventually, if the moving cost is very large, there is no mobility and tax progressivity is again non-distortionary. Thus, the estimate of the moving cost parameter bears important implications for evaluating the effect of tax policy on migration and welfare. Last, as an agent grows older, it becomes more likely that a geographic move is due to exogenous reasons instead of being a choice because the probability of exogenous relocation is age-independent while ageing reduces the incentives to move. This result plays an important role in accounting for the fact that in the data the difference in earnings growth between movers and stayers tends to decline with age (see Section 5.1).

Notice that, while in this model progressive taxation has the potential to increase social welfare by providing insurance against shocks and initial conditions, it does not increase the propensity to migrate because agents make moving choices after observing the vector $\boldsymbol{\varepsilon}$. In principle, higher tax progressivity has the potential to increase migration by providing

[16] Thus, the assumption that the mobility cost is a utility cost as opposed to an income cost is important. For example, one could specify the migration cost as a loss of a fraction of human capital, in which case τ_{t+k} would have no effect on mobility choices because redistribution would affect symmetrically both the choice of staying and that of moving.

insurance against unanticipated negative income shocks. For example, following a geographic move, a household might experience a temporary decline in income or its income process may become more volatile.[17] I have tested these two hypothesis by analyzing the empirical properties of the income process upon migration and found no empirical support for either hypothesis (see Online Appendix 4.1.3). The panel data used to conduct this analysis is from the Survey of Income and Program Participation. It is described in more detail in Section 5.1 where it is used to measure the income gains associated with interstate migration.[18]

3.2 Households' Idiosyncratic Productivity Growth

In this section I characterize the evolution of the idiosyncratic productivity process for movers and stayers. This characterization provides the basis for comparing the model's predictions with its empirical counterparts in Section 5.1. I also rely heavily on the results of Proposition 4 to compute analytically the moments targeted by the model's estimation and (numerically) the elements of the government's budget constraint (equation (9)) and the various components of the welfare function.

Over time a household's labor income changes because of the age-dependent evolution of its productivity, as captured by $\alpha_{a,r}$, and because of idiosyncratic shocks to productivity and migration. If the household is able to make a migration choice, its idiosyncratic productivity evolves as:

$$z' = \rho_r z + \eta_r g_t (\boldsymbol{\varepsilon}, a; r) , \tag{15}$$

[17] For example, Manovskii (2002) shows that, in an incomplete markets framework, higher tax progressivity can increase human capital accumulation by reducing the cost of on-the-job training.

[18] My empirical analysis cannot entirely rule out the possibility of a positive effect of tax progressivity on migration. For example, even if the income process does not become more volatile upon migration, individuals might have better information about local than foreign shocks at the time of making moving choices. I formalize this idea in Online Appendix 6.3 using an Epstein-Zin utility function. I show that if the coefficient of relative risk-aversion is larger than one, this version of the model features a *positive* effect of tax progressivity on migration (in addition to the negative effect on which I focus).

where:

$$g_t(\boldsymbol{\varepsilon}, a; r) = \sum_{k=1}^{K} M_t(a, \boldsymbol{\varepsilon}, h, k; r)\, \varepsilon_k,$$

and $g_t(\boldsymbol{\varepsilon}, a; r)$ does not depend on h because of the symmetry of the model.[19] If the household is exogenously relocated to $k \neq h$, instead, its productivity evolves according to equation (2). The following proposition characterizes the distribution of $g_t(\boldsymbol{\varepsilon}, a; r)$.

Proposition 4 (Innovation to productivity growth with endogenous mobility) *The innovation $g_t(\boldsymbol{\varepsilon}, a; r)$ to the productivity process, conditional on choosing to stay in the same location, is distributed according to a type 1 extreme-value distribution with mean:*

$$E_{\boldsymbol{\varepsilon}}\left[g_t(\boldsymbol{\varepsilon}, a; r) \,|\, M_t(a, \boldsymbol{\varepsilon}, h, h; r) = 1\right] = \gamma + \ln \delta_r - \ln p_t(a; r), \tag{16}$$

and variance:

$$VAR_{\boldsymbol{\varepsilon}}\left[g_t(\boldsymbol{\varepsilon}, a; r) \,|\, M_t(a, \boldsymbol{\varepsilon}, h, h; r) = 1\right] = \pi^2/6. \tag{17}$$

The innovation $g_t(\boldsymbol{\varepsilon}, a; r)$ to the productivity process, conditional on choosing to migrate, is distributed according to a type 1 extreme-value distribution with mean:

$$E_{\boldsymbol{\varepsilon}}\left[g_t(\boldsymbol{\varepsilon}, a; r) \,|\, M_t(a, \boldsymbol{\varepsilon}, h, h; r) = 0\right] = \gamma + \ln \delta_r - \ln p_t(a; r) + \frac{\kappa_r}{\eta_r v_t^1(a; r)}, \tag{18}$$

and the same variance as in (17). The parameter γ in these equations denotes Euler's constant.

Notice the difference between the evolution of productivity with exogenous and endogenous mobility. In the former case there is no selection, and the innovation to productivity experienced by an agent who is exogenously relocated follows the density of shocks drawn from

[19] In other words, productivity growth does not depend on the specific "identity" h of the location where the household resides at the beginning of the period because the aggregate characteristics of each location are the same.

one of the "away" locations. By assumption, the latter is distributed as a type 1 extreme-value distribution with mean γ and variance $\pi^2/6$. For agents who can choose whether to migrate or not, movers are self-selected based on the evolution of productivity. In particular, the expressions in equations (16) and (18) reflect selection through migration.[20] The conditional mean in equation (16) reflects selection by agents who choose to stay put. For the latter to be the optimal choice, stayers must experience higher productivity growth, on average, than the set of households that are exogenously reallocated. This is indeed the case since it is always the case that $p_t(a;r) < 1 \leq \delta_r$. In particular, the lower the ex-ante chance of staying put – the lower $p_t(a;r)$ – the higher the average conditional productivity growth for stayers must be.[21] For agents who choose to move out of a location, instead, average productivity growth in (18) exceeds, on average, productivity growth conditional on staying by an amount that reflects the cost of moving. This is, again, a reflection of selection. For costly migration to be preferable to staying put, agents must experience, on average, an additional gain in productivity. We can summarize the discussion above in the following corollary.

Corollary 1 *Average productivity growth is largest for households who choose to relocate and smallest for households who are forced to relocate. Households who choose to stay in the same location experience an intermediate level of productivity growth.*

4 Empirical Implementation

In order to estimate the model's parameters I focus on the stationary equilibrium of the model with constant tax progressivity parameter τ. The following functional forms are assumed. The exogenous distribution of household productivity $f(z|1,r)$ at age one is taken to be lognormal with parameters (μ_r, σ_r^2). The exogenous component of the evolution of household

[20] See Nakosteen and Zimmer (1980) for an early application of Heckman's correction for selectivity bias to the income equations of internal migrants and non-migrants with normally-distributed disturbances.

[21] A low probability of staying reflects a relatively small moving costs. Thus, agents who choose to stay must experience a high income growth in the home location relative to what they would have earned if they had chosen to move.

productivity is a quadratic function of age:

$$\alpha_{a,r} = \overline{\alpha}_{0,r} a + \overline{\alpha}_{1,r} a^2. \tag{19}$$

The parameter vector is then:

$$\left\{\beta, \zeta, \overline{a}, \overline{r}, K, \tau, \chi, \left(\omega_r, \kappa_r, \theta_r, \overline{\alpha}_{0,r}, \overline{\alpha}_{1,r}, \eta_r, \delta_r, \mu_r, \sigma_r^2, \rho_r\right)_{r=1,\overline{r}}\right\}.$$

The empirical strategy has two parts. First, a number of parameters are set a-priori, including τ. Second, the remaining parameters are estimated by the Generalized Method of Moments (Hansen, 1982). I postpone setting χ until Section 6 as it only plays a role in welfare analysis.

4.1 Parameters Set A-Priori

The frequency of the model is one year. The years of working life are $\overline{a} = 35$, from age 25 to age 59 included. The annual discount factor $\beta = 0.97$. The labor supply parameter $\zeta = 3$, implying a Frisch elasticity of 0.5. These three numbers are the same as those selected by Heathcote et al (2020a). The number of locations K is set to 51 to capture mobility across U.S. states and the District of Columbia. The number of household types is set to $\overline{r} = 2$: households whose head has less than a college degree ($r = 1$) and households whose head has a college degree or more ($r = 2$). The fraction of $r = 1$ types in the American Community Survey 2000-07 data used to estimate the model is $\omega_1 = 0.6751$.[22] Finally, I set the parameters $\rho_r = 1$ a-priori, so that the idiosyncratic productivity process is a random walk.[23] This is consistent with numerous empirical studies of the income process (e.g. Meghir and Pistaferri, 2011).

[22] Notowidigdo (2011), Lkhagvasuren (2014) and Amior (2020) study differences in geographic mobility between workers of different education level. Empirically, there is a monotonic relationship between age-specific migration rates and schooling.

[23] I have tried to estimate ρ_r directly. In this case the estimation procedure converges to 1, so I have imposed it a-priori.

4.2 Measure of Tax Progressivity

The tax progressivity parameter τ is estimated using data on households' market income and income post-Federal taxes and transfers, following Heathcote et al (2017). Recall that, from equation (1), $(1-\tau)$ represents the elasticity of income post-Federal taxes and transfers to market income. This suggests an empirical specification of the form:

$$\ln \widetilde{y}_{pt} = a_t + b \ln y_{pt} + u_{pt}, \qquad (20)$$

where a_t represents year fixed-effects, b is the elasticity of post-redistribution income to market income (or $1-\tau$), and u_{pt} is an error term, assumed to be uncorrelated with $\ln y_{pt}$. The subscript p denotes percentiles of the household income distribution. The data used to estimate the parameter b are, in fact, percentiles of the distribution of post and pre-government household income from the Congressional Budget Office (2011, Table A-1). Data is available for percentiles $p = 20, 40, 60, 81, 91, 96, 99$. The CBO market income measure includes all cash income (taxable and tax-exempt), taxes paid by businesses and imputed to households such as corporate taxes and the employer's share of payroll taxes, and benefits, such as employer-paid health insurance premiums. The after-government income includes cash transfer payments (for example, unemployment insurance and welfare) and estimates of the value of in-kind benefits (Medicare, Medicaid, Children's Health Insurance Program, Supplemental Nutrition Assistance Program). It subtracts federal individual and corporate income taxes, payroll taxes and excise taxes.[24]

For consistency with the other data used in the paper, I restrict attention to the period 2000-2007. The estimated value of b is equal to 0.808 with a standard error 0.011 and $R^2 = 0.995$. This leads to a value of τ equal to 0.192, close to Heathcote et al (2017)'s estimate of 0.18 using the PSID for the period 2000–2006 and to Bakis et al (2015)'s estimate of 0.17 using March CPS data (1979–2009) and the NBER tax simulator. The Congressional

[24] State and local income taxes are not included in the CBO calculations. These taxes are significantly less progressive than Federal taxes so their exclusion is unlikely to significantly affect the estimate of τ. My benchmark estimate is, in fact, very close to Heathcote et al (2017)'s figure (see the main text). They include state and local income taxes in their calculations based on the PSID.

Budget Office data are available at an yearly frequency starting in 1979. Interestingly, there is no evidence for a trend in τ over this period.[25] In Section 5.2 I discuss a different approach to measuring τ that can be used for all OECD countries. This approach derives τ from a comparison of a country's Gini coefficient of market income and the Gini coefficient of income after taxes and transfers. Following this alternative procedure yields an average estimate of τ equal to 0.21 for the U.S. in 2000–2007. Consistently with the evidence from the CBO data, also these estimates of τ are remarkably stable over time.[26]

4.3 Estimation and Identification of the Remaining Parameters

The remaining parameters are estimated by GMM. There are a total of 16 parameters to be estimated:

$$\boldsymbol{\phi}_r = \left\{\mu_r, \sigma_r^2, \kappa_r, \theta_r, \delta_r, \eta_r, \overline{\alpha}_{0,r}, \overline{\alpha}_{1,r}\right\}_{r=1,2}. \tag{21}$$

The parameter vector $\boldsymbol{\phi}_r$ is estimated separately by household type r. For each type, the estimation procedure targets the following $3\overline{a}$ moments: 1) the average migration rate by age ($\overline{a}$ conditions); 2) the cross-sectional mean of the distribution of log household earnings by age ($\overline{a}$ conditions); 3) the cross-sectional variance of the distribution of log household earnings by age ($\overline{a}$ conditions). The analytical results of Section 3 allow me to compute all moment conditions analytically (see Online Appendix 3.2), so the estimation procedure does

[25] Adding a linear interaction term between year and the coefficient on $\ln y_{pt}$ in equation (20) and using the entire sample period 1979-2007, yields an estimate of b equal to 0.817 and a statistically insignificant estimate of -0.0003 on the interaction term. Heathcote et al (2020b) also use CBO data to estimate τ and find that "the degree of tax progressivity has remained approximately constant over time."

[26] The estimates are in the range 0.19–0.23 for the period 1974-2012. In 1974, τ is equal to 0.19 while in 2012 it is equal to 0.21. Ferriere and Navarro (2018) use U.S. federal tax (but not transfer) data to estimate the tax progressivity parameter τ over time starting in the early 20th century. Their estimate suggests that tax progressivity in the U.S. has declined in the early-mid 1980s and has been relatively stable since then. Guner et al (2014) also estimate versions of the tax function (20) using tax (but not transfers) data and estimate a smaller value of b.

not rely on the use of simulations. The GMM estimator of $\boldsymbol{\phi}_r$ solves the following problem:

$$\widehat{\boldsymbol{\phi}}_r = \arg\min_{\boldsymbol{\phi}_r} \mathbf{m}'_r(\boldsymbol{\phi}_r)\, \mathbf{W}_r^{-1} \mathbf{m}(\boldsymbol{\phi}_r),$$

where $\mathbf{m}_r(\boldsymbol{\phi}_r)$ denote the moment conditions and $\mathbf{W}_r^{-1}$ is Hansen (1982)'s optimal weighting matrix.

These moment conditions are sufficient to identify the parameter vector. The geographic mobility data identify the parameters κ_r, θ_r, δ_r. In order to understand the role played by each of these parameters, consider the migration rate $1 - \theta_r p(a; r)$ at age a, where $p(a; r)$ is formally defined in equation (14).[27] As a household grows older, its incentives to voluntarily migrate decline because migration is costly ($\kappa_r > 0$) and its remaining time horizon ($\overline{a} - a$) shrinks. Therefore, the parameter θ_r is identified by the migration rate at older ages, when migration mostly occurs for exogenous reasons (see Proposition 3, point 4). To understand the different roles played by δ_r and κ_r, notice that if moving costs were zero ($\kappa_r = 0$), the migration rate would be a constant independent of age, because $p(a; r)$ would not depend on a (see equation 14). Thus, to account for the declining pattern of migration as a function of age, it is necessary that $\kappa_r > 0$. Given the discount factor β, the size of the moving cost relative to the importance of idiosyncratic shocks, κ_r/η_r, regulates the dependence of migration rates on age. Thus, κ_r/η_r is identified by the observed difference in average migration rates at younger and middle ages ($a = 1$ and, approximately, $a = 20$, based on Figure 1). Given κ_r/η_r, the parameter δ_r pins down the average migration rate of an agent through her lifecycle.

The age profile of mean wages for each education type identifies the parameters $(\overline{\alpha}_{0,r}, \overline{\alpha}_{1,r})$. The observed mean and variance of log income at age 25 ($a = 1$) identify, respectively, the mean μ_r and variance σ_r^2 of the initial productivity distribution.[28] As a cohort ages, the cross-

[27] Online Appendix 3.1 contains a formal illustration of this identification argument based on a version of the model with $\overline{a} = 3$. This is the minimum number of ages necessary to identify the three migration cost parameters.

[28] Notice that, since the productivity scale is arbitrary, I have normalized, without loss of generality, average income to equal one at age 25 for $r = 1$ agents. This normalization pins down μ_1.

sectional variance of log income in the model increases, due to the presence of idiosyncratic shocks. The slope of this relationship in the data identifies the parameter η_r, which regulates the importance of the ε shocks for households' productivity growth.

4.4 Sample Selection

The data set used to estimate the model is from the American Community Survey (ACS), years 2000-2007 (Ruggles et al (2017)). The main advantage of the ACS data is the relative large sample which allows one to accurately measure migration rates by age. The main disadvantage is that the ACS data is purely cross-sectional and does not allow one to measure wage growth for either migrants or non-migrants.[29] The ACS data sample consists of households whose head is 25–59 years old, and is not institutionalized or in school. I define a household's labor income as the sum of the wage, salary and business income of the household's members. I drop observations in the bottom 10 percent of yearly earnings. The geographic unit of observation is a U.S. state and each model period represents a year. Households' labor income and mobility data are purged of year, state, and sex effects by running regressions of each of these variables on year, state and sex dummies and using residuals to construct the moments of interest. The latter approach is consistent with the model's assumption that locations are ex-ante identical.

4.5 Estimated Parameters and Model's Fit for Targeted Moments

Table 1 reports the estimates of the model's parameters and their standard errors. The mobility cost κ_r, converted into consumption-equivalent units $(1 - \exp(-\kappa_r))$, corresponds to 90 and 99 percent of yearly consumption for $r = 1, 2$ agents respectively, a relatively large number but comparable to the figure estimated by Kennan and Walker (2011). The parameters θ_r are such that the probabilities of an exogenous move are 1.1 and 1.8 percent per year for households of type $r = 1, 2$, respectively. Therefore, exogenous relocation

[29] For this purpose I use longitudinal data from the Survey of Income and Program Participation, as discussed in Section 5.1.

Parameters set a priori			
		Value	Source
J	number of locations	51	U.S. states
ξ	Frisch elasticity parameter	3	Heathcote et al (2020a)
β	yearly discount factor	.97	Heathcote et al (2020a)
τ	degree of tax progressivity	.192	CBO
$\overline{a}$	duration of working life	35	Heathcote et al (2020a)
ω_1	measure of $r = 1$ (less than college) agents	0.6751	ACS, 2000-2007
ω_2	measure of $r = 2$ (college and above) agents	0.3249	ACS, 2000-2007
ρ_1	autocorrelation of productivity, $r = 1$	1.00	n.a.
ρ_2	autocorrelation of productivity, $r = 2$	1.00	n.a.
Estimated Parameters for Type $r = 1$ Agents (Less than College Degree)			
		Estimate	Standard Error
κ_1	mobility cost parameter	2.3095	0.1095
δ_1	frequency of local offers parameter	98.501	14.7234
η_1	importance of idiosyncratic shocks	0.0444	0.0004
$\overline{\alpha}_{0,1}$	age profile of earnings (linear)	−0.2000	0.0063
$\overline{\alpha}_{1,1}$	age profile of earnings (quadratic)	−0.0001	0.0000
θ_1	one minus probability of exogenous move	0.9893	0.0002
μ_1	mean of initial log productivity	−0.1558	0.0026
σ_1^2	variance of initial log productivity	0.3588	0.0014
Estimated Parameters for Type $r = 2$ Agents (College Degree and Above)			
		Estimate	Standard Error
κ_2	mobility cost parameter	5.7771	0.1212
δ_2	frequency of local offers parameter	2.2817	0.3074
η_2	importance of idiosyncratic shocks	0.0594	0.0004
$\overline{\alpha}_{0,2}$	age profile of earnings (linear)	−0.0556	0.0080
$\overline{\alpha}_{1,2}$	age profile of earnings (quadratic)	−0.0006	0.0000
θ_2	one minus probability of exogenous move	0.9828	0.0002
μ_2	mean of initial log productivity	0.2074	0.0039
σ_2^2	variance of initial log productivity	0.2925	0.0028

Table 1: Summary of model parameters and standard errors.

accounts for 56 and 48 percent of all moves for these two types. These figures are consistent with households' responses to the Current Population Survey's question about reasons for interstate migration cited in Section 2. College educated agents are characterized by a higher moving cost ($\kappa_2 > \kappa_1$) and by a lower frequency of offers in the home location ($\delta_2 < \delta_1$). As discussed in the identification section above and shown formally in Online Appendix 3.1, the lower frequency of home offers allows the model to account for the higher mobility of college-educated workers. Their higher moving cost gives rise to the faster decline between young and middle-ages in the geographic mobility rate observed for this group.[30] Figure 1 shows the age pattern of migration. At younger ages the migration rate of college-educated labor is about twice as large as the migration rate of those in the less skilled group. Such difference tends to disappear at older ages. Figure 2 represents the fit of the model with respect to the (log) earnings moments for each type of agents. Notice that the model captures very well the evolution of the cross-sectional pattern of the mean and the variance of earnings across the ages.

5 Model's Fit for Non-Targeted Moments

In this section I discuss the fit of the estimated model using two datasets that were not used to estimate its parameters. First, I ask whether the model can account for the association between migration and earnings growth at the household level. Second, I compare its prediction for the relationship between tax progressivity and aggregate migration rates with empirical evidence based on cross-country data.

[30] Amior (2020) studies the gap in migration rates across skill groups in a search model of migration. While his model and data are different from mine, he finds that more educated workers are more mobile because the wage returns from interstate migration are larger for this group. His estimated migration costs are also larger for more educated workers.

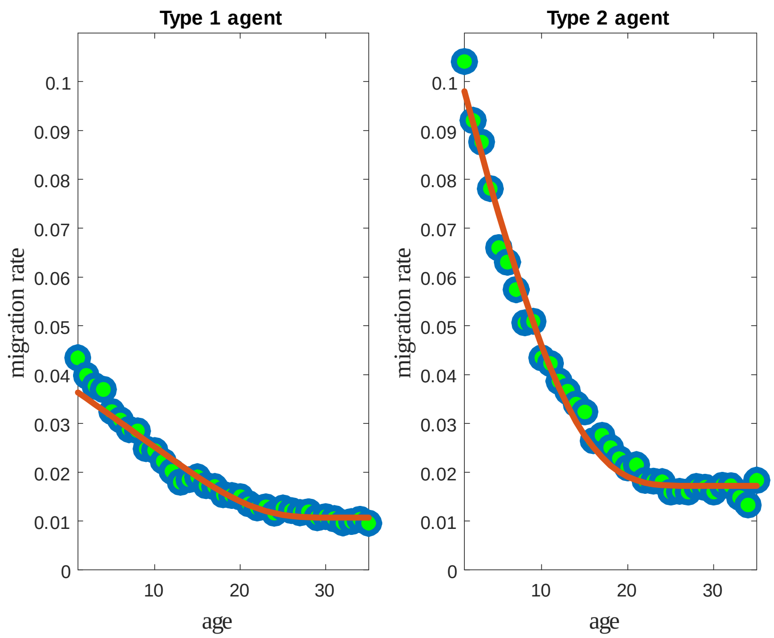

Figure 1: Average migration rate by age in the model (line) and in the data (circles) for type $r = 1, 2$ agents.

5.1 Earnings Growth and Migration

The model's parameters have been estimated using cross-sectional moments on household earnings and migration patterns by age. In this section I evaluate its implications for the difference in earnings *growth* between movers and stayers. This requires using longitudinal household-level data with a sufficiently large sample size because, on average, yearly migration rates are of the order of two percent per year.[31] The data I use to measure gains in labor income associated with inter-state migration are from the Survey of Income and Program Participation (SIPP User Guide, 2001).[32] The SIPP is a nationally representative

[31] I am aware of only a few empirical studies in this area. Yankow (2003, Table 8) uses panel data from the 1979 National Longitudinal Study of Youth to compare wage gains of internal migrants relative to stayers and finds wage gain gaps of around 9 percentage points, three to four years after the migration event for full-time workers with more than a high school degree. The corresponding gap estimated by Yankow for full-time workers with a high school degree or less is about 2 percentage points.

[32] I thank Jose' Mustre-del-Rio for pointing out the potential usefulness of the SIPP data to analyze the dynamics of gains in earnings upon migration. The advantage of the SIPP is that it has a much larger sample size than other longitudinal datasets such as the Panel Study of Income Dynamics or the National Longitudinal Study of Youth. Differently from the Current Population Survey, the SIPP attempts to locate original sample members even if they move to a new address.

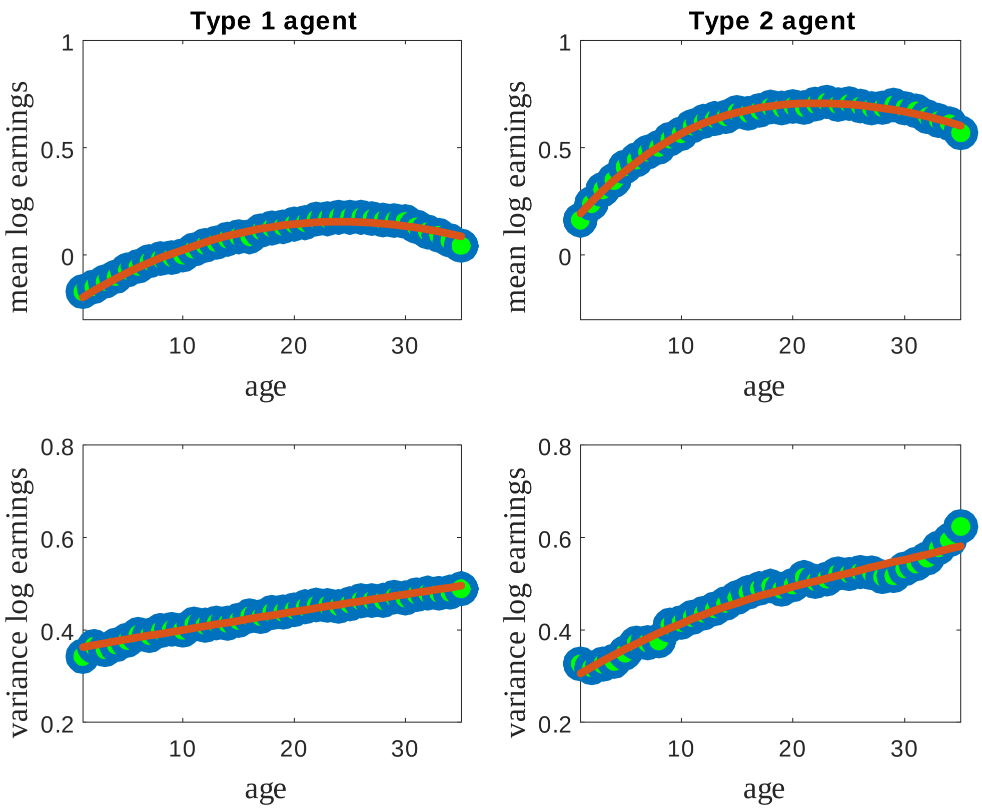

Figure 2: Cross-sectional mean and variance of log earnings by age in the model (line) and in the data (circles) for type $r = 1, 2$ agents.

longitudinal household survey. Starting with the 1996 redesign, every three or four years the SIPP selects a new panel of households and interviews them every four months for three or four years. In each interview (wave), the household provides information on a large number of variables, including earnings and state of residence, in each of the prior four months. I use the 1996, 2001, 2004, and 2008 panels of the SIPP.

I focus on households with heads aged between 25 and 59 years old.[33] For each household i, I use the SIPP data to construct a panel of monthly observations on household earnings and inter-state moves. Given the short length of each panel, conditional on ever moving across state lines, the great majority of households moves only once. Let M_i be an indicator that equals one if household i ever moves across state lines and zero otherwise. The earnings gain associated with an inter-state move are estimated from the following regression equation:

$$(22)\quad \ln y_{i,a,s,m} = M_i \times Post_{i,m} \times (\xi_1 + \xi_2 (a - 25)) + \xi_3 \ln n_{i,m}^{18-64} + \zeta_i + \zeta_a + \zeta_s + \zeta_m + \epsilon_{i,a,s,m},$$

where $y_{i,a,s,m}$ denotes household earnings and a, s and m are indices for household age, state-of-residence and month-year of the data. The variable $Post_{i,m}$ is an indicator that equals one if the household is observed in month-year m after an interstate move and zero else. The regression controls for the (log) number of households members between 18 and 64 ($n_{i,m}^{18-64}$), and for fixed effects for household (ζ_i), age (ζ_a), state-of-residence (ζ_s), and month-year of the data (ζ_m).

The effect of geographic mobility on earnings is captured by the term $\xi_1 + \xi_2 (a - 25)$, which represents the within-household proportional change in earnings after an interstate move. Notice that this specification allows the effect of geographic mobility on earnings to vary by age. I estimate the regression separately for households headed by an individual with a college degree and for households headed by an individual with less than a college degree. The estimates of these parameters are reported in Table 2.

Figure 3 provides a visual representation of these estimates by plotting the predicted gap in labor income growth between movers and stayers as a function of age in the data

[33] Online Appendix 4.1 provides further details on sample selection criteria.

Parameter	Less than college	College and above
ξ_1	0.105**	0.174***
	(0.051)	(0.051)
ξ_2	-0.008**	-0.006*
	(0.003)	(0.003)
Number of observations	1,209,340	652,242
Number of households	36,154	18,146

Table 2: Estimates of the impact of interstate migration on household earnings. Standard errors clustered at the household-level in parenthesis. Data source: SIPP 1996, 2001, 2004, 2008 panels. Regressions control for household, age, state, and month-year fixed effects. * denotes a p-value <0.10; ** <0.05 and *** <0.01.

$(\widehat{\xi}_1 + \widehat{\xi}_2\,(a-25))$ and in the model.

The estimates in Table 2 are qualitatively and quantitatively consistent with two basic predictions of the structural model. First, earning gains of migration are proportionally larger for more educated workers. Second, for both groups these gains tend to decline with age and may even be negative at older ages. The model is able to account for the second prediction due to the possibility of exogenous relocation. Migrants who *choose* to move are positively selected on earnings growth and the extent of positive selection *increases* in age because older households have a shorter remaining horizon to make up for the fixed cost of moving.[34] Therefore, absent exogenous relocation, the model would incorrectly predict that the earnings gains of migration increased with age. In the model, however, exogenous relocation accounts for a growing fraction of all moves as an individual ages and incentives to migrate voluntarily decline (Proposition 3, part 4). Since those who are relocated exogenously experience smaller average earnings growth than workers who choose not to migrate (see Corollary 1), this effect accounts for the observation that the earnings gains of migration decline with age.

[34] Formally, inspection of equations (16) and (18) in Proposition 4 reveals that the difference between earnings growth of movers and stayers (by choice) is given by the term $\kappa_r/v_t^1\,(a;r)$ which is increasing in age.

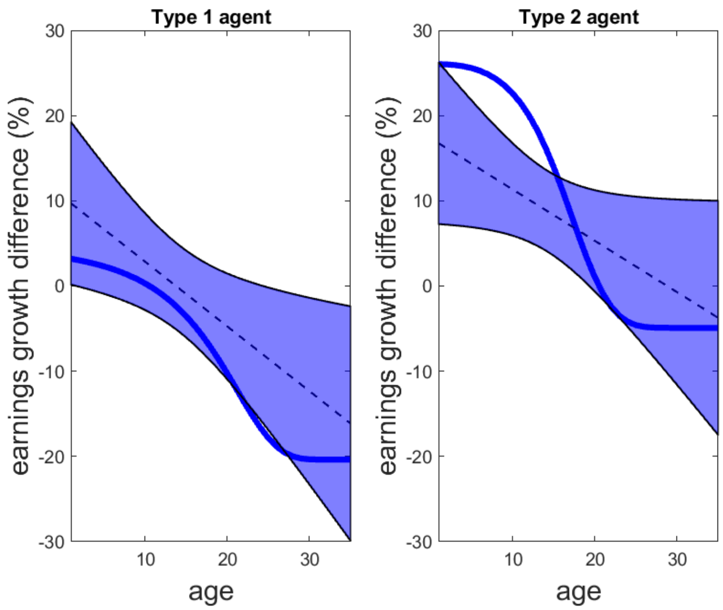

Figure 3: Average gap in labor income growth between movers and stayers for type $r = 1, 2$ agents as a function of age. Data generated by the model are denoted by the thick blue lines. Estimates based on the regression equation (22) are denoted by the dashed line (- -). 95% confidence intervals based on the regression estimates are also reported.

5.2 Tax Progressivity and Internal Migration Across-Countries

The model predicts that economies characterized by a higher degree of tax progressivity τ will display smaller internal migration rates. In this section I test this prediction quantitatively using OECD data on tax progressivity and internal migration that was *not* used in the estimation of the model's parameters. Data on internal migration comes from the OECD (2000, 2005, 2013) and refers to the period 1980–2010. The territorial unit in each country is the first administrative tier of sub-national governments, such as a state for the U.S. and a Lander for Germany.[35] Since there is no readily-available estimate of the parameter τ across countries, I measure the latter using OECD data on the distribution of household income before (y) and after ($\widetilde{y}$) taxes and transfers for the period 1980–2013. Specifically, for

[35] These are called Territorial Level 2 (TL2) units by the OECD. I focus on the 21 countries with more than five TL2s because internal migration in countries with very few TL2s (Belgium, Denmark, Finland, Slovakia, Slovenia, Iceland) is likely to be qualitatively different from the rest of the sample. See Online Appendix 4.2 for the full list of countries.

each country and year I compute the value of τ that rationalizes the ratio between the Gini coefficient of income after taxes and transfers and the Gini coefficient of market income.[36] Since measures of tax progressivity and of internal migration may be subject to measurement error, I average each of these variables for a given country over the sample period.

Figure 4 represents a scatter plot of average migration rates and estimates of tax progressivity, τ. The slope of the regression line is -5.15 with a standard error equal to 2.06.[37] This implies that an increase in tax progressivity by a cross-sectional standard deviation (0.07) is predicted to reduce a country's average migration rate by about 0.35 percentage points, or slightly less than one-half of the cross-sectional standard deviation of migration rates (see Table A.3 in the Online Appendix).

Interestingly, the partial correlation between migration and tax progressivity is robust to controlling for differences in the *level* of taxes across countries. For example, including average personal income taxes (net of cash transfers) relative to the average wage in the regression for internal migration yields a coefficient on τ equal to -5.58 (s.e. 1.38). The measure of average taxes enters with a *positive* and marginally statistically significant sign in this regression.[38]

While the evidence presented in this section is suggestive of a negative effect of tax progressivity on internal migration, caution has to be used in interpreting the correlation in Figure 4. Countries may vary along other, potentially unobserved, dimensions in addition to the degree of progressivity of their tax system. For example, countries in which geographic mobility is low for cultural reasons might choose a more progressive tax system to provide additional insurance against idiosyncratic income shocks (Hassler et al (2005)).[39]

[36] The income data is measured at the household level for the population 18–65 and is adjusted for household size with an equivalence scale. The procedure to compute τ is described in detail in Online Appendix 4.2. Table A.3 in this appendix presents summary statistics on the estimates of tax progressivity across countries.

[37] Online Appendix 4.2 contains details on the regression results discussed in this section and on a number of robustness exercises.

[38] See the regression results in Online Appendix 4.2, Table A.4. Taking into account the cross-sectional standard deviations of these variables and the size of the regression coefficients, the effect of the level of taxation on migration is, in absolute terms, about two-thirds of the effect of tax progressivity. Online Appendix 6.1 shows that, with a general CRRA utility function, the effect of the level of taxes on the incentive to migrate might be either positive or negative.

[39] A related, but different, interpretation of the data is that in a country in which the internal migration

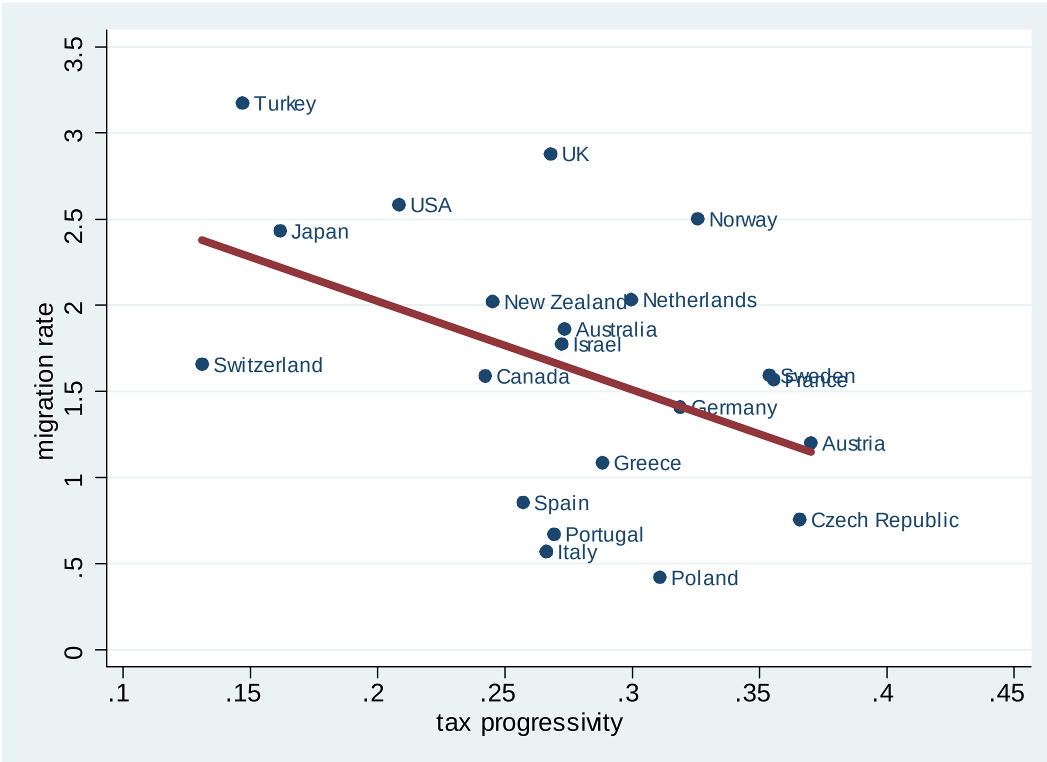

Figure 4: Scatterplot of country-level migration rates against country-level tax progressivity τ. The regression line has slope -5.15 (s.e. 2.06) with an $R^2 = 0.21$.

Some of these concerns might be attenuated by exploiting the panel dimension of the data. Unfortunately, the available data is an unbalanced panel as neither observations on internal migration nor on Gini coefficients are collected on a yearly basis. There are only 32 country-year pairs of observations for the post-1980 period for 21 countries. A panel regression of internal migration rates on measured tax progressivity with year and country fixed effects yields an estimate of -7.51 (s.e. 6.94). While this is not statistically different from zero, the magnitude of the point estimate is comparable with the cross-sectional one. Ultimately, questions of causality are difficult to settle without quasi-experimental variation in τ across countries and over time.

In what follows I sidestep the causality question and, instead, ask whether the quantitative

rate is low for reasons other than taxes, a utilitarian government might choose a more progressive tax system because it does not need to worry about distorting geographic mobility choices. Such interpretation is consistent with the hypothesis advanced in this paper.

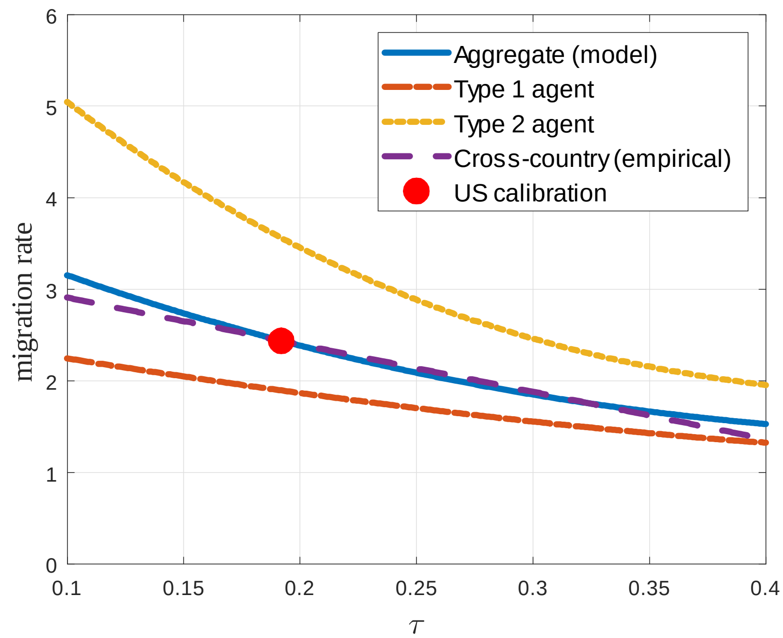

Figure 5: Aggregate and type-specific migration rates across steady states for different values of the tax progressivity parameter τ. The dot denotes the model-predicted average migration rate for the U.S. in the benchmark calibration.

model can account for this cross-country evidence. To do so, I compute the model-implied steady state migration rate for various degrees of tax progressivity τ.[40] Figure 5 plots the steady state migration rate for the economy as a whole (solid line) and for each type of agent as the parameter τ varies. The benchmark calibration is denoted by the point "US" in the figure. For comparison, the figure also represents (dashed line) the cross-country regression represented in Figure 4. Notice that I have adjusted its intercept so that it passes through the point "US" in order to facilitate comparison with the model. The quantitative prediction of the model is consistent with the cross-country evidence. Specifically, an increase in τ from its benchmark value of 0.192 to 0.262 - a cross-country standard deviation - is predicted to reduce internal mobility by 0.41 percentage points in the model and by 0.35 percentage points in the data.

[40] The steady state assumption is the counterpart of the fact that the cross-country empirical evidence is based on comparisons of country-average internal migration rates and tax progressivity over a period of about 30 years.

6 Welfare Analysis

In this section I use the model to analyze how a utilitarian planner would optimally set tax progressivity. I begin by considering a simplified version of the model to show analytically how internal migration affects the planner's trade-offs concerning tax progressivity. I then consider the full quantitative model and compute optimal policies numerically.

6.1 Policy and Welfare: Analytical Results

A simplified version of the model allows me to study analytically the welfare implications of varying the degree of tax progressivity. In the simplified version of the model households live forever, ($\overline{a} \rightarrow +\infty$) and the exogenous productivity term $\alpha_{a,r}$ is set to zero, without loss of generality. In addition, I restrict attention to economies with one type of agent ($\omega_1 = 1$); no exogenous mobility ($\theta_r = 1$); i.i.d. shocks ($\rho = 0$); two locations ($K = 2$); and symmetric shock distributions ($\delta_r = 1$). The lack of persistence in z implies that, without loss of generality, I can focus on a tax policy that is constant over time.[41]

To discuss the welfare implications of fiscal policy in this economy, I assume that the social welfare function is utilitarian. The latter is equal to the sum of the lifetime utilities of all agents, taking moving costs into account:[42]

$$W = (1 - \beta) \int P(\boldsymbol{\varepsilon})\, g(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon}, \tag{23}$$

where $P(\boldsymbol{\varepsilon})$ is the unconditional value function (see equation (8)).

Under these assumptions, W can be computed analytically as a function of the two policy variables available to the planner, i.e., tax progressivity and the supply of the public good.[43]

Proposition 5 (Social welfare function) *Under the assumptions of this section, the so-*

[41] The assumption of i.i.d. shocks allows me to derive the welfare function analytically because it implies that the stationary distribution of productivity $f_t^p(z')$ can be easily inferred from the distribution of shocks $g_t(\boldsymbol{\varepsilon})$ characterized in Proposition 4.

[42] Notice that I multiply the social welfare function by $(1 - \beta)$ to simplify the notation.

[43] Recall that given these two policy variables, the third one, λ, is implied by the government's budget constraint, equation (9).

cial welfare function takes the form

$$W(\varphi,\tau) = W^{ra}(\varphi,\tau) + \tag{24}$$
$$+ (1+\chi)\ln\left(p^{1-\eta} + (1-p)^{1-\eta}\right) - \kappa(1-p) -$$
$$- \left(\ln\Gamma(1-(1-\tau)\eta) - \eta\gamma(1-\tau) + \ln(p + (1-p)\exp(\kappa)) - \kappa(1-p)\right),$$

where $W^{ra}(\varphi,\tau)$ *is welfare in the representative-agent economy, defined as:*

$$W^{ra}(\varphi,\tau) = \ln(1-\varphi) + \chi\ln\varphi + (1+\chi)\ln\ell^* - \zeta^{-1}(\ell^*)^{\zeta} + (1+\chi)\ln\Gamma(1-\eta), \tag{25}$$

φ *denotes the share of public goods in aggregate output,* Γ *is the gamma function, and* p *is the probability of not migrating:*

$$p = \frac{1}{1+\exp(-\kappa/((1-\tau)\eta))}. \tag{26}$$

The first row of equation (24) represents the welfare of a representative-agent version of this economy, in which each agent's productivity is exogenous and equal to the average level of productivity in the case of no migration. In this case, welfare depends only on the allocation of output between private and public consumption and on the quantity of output produced by exerting effort, net of the disutility of labor supply. As discussed by Heathcote et al (2017, Proposition 5), and easily shown here as well, the welfare of the representative is maximized with a share of public good equal to $\varphi^* = \chi/(\chi+1)$ and by subsidizing labor supply by setting tax progressivity to $\tau = -\chi$. The labor subsidy is needed because agents do not internalize the effect of their labor supply choices on the provision of public goods.

Consider now the second row of equation (24). It represents the net social welfare derived from the migration process. The latter has the potential to increase aggregate productivity by locating workers where they are idiosyncratically more productive. Specifically, the term multiplying $(1+\chi)$ in the second row of equation (24) is the logarithm of average productivity in the economy. This term is multiplied by $(1+\chi)$ because the planner, but not the agents, takes into account the effect of higher geographic mobility on the provision of public goods. The term $\kappa(1-p)$, instead, represents the aggregate utility costs of migration. The level of τ that optimally trades-off the social benefits of the productivity gains from migration and aggregate mobility costs satisfies the following condition:

$$(1+\chi)\frac{(1-\eta)\left((1-p)^{-\eta}-p^{-\eta}\right)}{p^{1-\eta}+(1-p)^{1-\eta}}=\kappa. \tag{27}$$

The left-hand side of equation (27) represents the marginal social cost of reducing migration. This depends on the average productivity difference between movers and stayers, represented by the numerator on the left-hand side of (27). Notice that the average productivity difference declines with η due to a selection effect that mitigates the productivity loss due to a higher τ. With a higher τ, fewer agents move, but those who do so must enjoy a higher average productivity gain (see Proposition 4).[44]

The term in the third row of equation (24) represents the welfare cost of consumption inequality. Consumption dispersion is due to both the direct effect of exogenous idiosyncratic shocks and to agents' endogenous migration responses to these shocks. To minimize the welfare costs of consumption inequality, the social planner would like to set tax progressivity at its maximum level, $\tau=1$, pooling consumption among all agents.

Putting all these effects together, the first-order condition of W with respect to τ is:

$$\begin{aligned}\frac{\partial W(\varphi,\tau)}{\partial\tau}=&\underbrace{\frac{\partial W^{ra}(\varphi,\tau)}{\partial\tau}-\eta\left\{\frac{\Gamma'(1-(1-\tau)\eta)}{\Gamma(1-(1-\tau)\eta)}+\gamma\right\}}_{\text{trade-offs in the one location version of the model }(K=1)}+\\ &+\underbrace{\left\{\frac{\exp(\kappa)-1}{p+(1-p)\exp(\kappa)}-(1+\chi)(1-\eta)\frac{(1-p)^{-\eta}-p^{-\eta}}{p^{1-\eta}+(1-p)^{1-\eta}}\right\}\frac{\partial p}{\partial\tau}}_{\text{trade-offs associated with geographic mobility}}=0,\end{aligned} \tag{28}$$

where $\partial p/\partial\tau>0$ by equation (26). For ease of interpretation, I have re-arranged terms so that the first-row of equation (28) reflects the trade-offs associated with the one-location

[44] When $\eta=1$, this selection effect fully offsets the direct productivity losses of a higher τ and the left-hand side of equation (27) becomes zero.

version of the model. If $K = 1$, there is no migration and agents experience purely exogenous variation in their productivity over time. The relevant trade-offs are those associated with the representative agent version of the model and the social benefits from reducing consumption inequality (term in curly brackets in the first row of equation (28)). Let $\widehat{\tau}$ denote the level of tax progressivity that maximizes welfare in the one-location version of the model.[45]

The second row of equation (28) represents, instead, the contribution that geographic mobility makes to the socially optimal level of τ. Since the two terms $\kappa(1-p)$ involving mobility costs in equation (24) cancel out, a marginal increase in τ produces only two effects.[46] The first-term in the curly brackets represents the welfare benefit of lowering consumption dispersion by reducing internal migration. The second term in the curly brackets represents the social cost associated with lower average productivity brought about by a higher τ. An important question is whether these additional trade-offs associated with geographic mobility contribute to push the welfare-optimal τ above or below $\widehat{\tau}$. The answer to this question depends on the model's parameters. The following proposition provides sufficient conditions under which each of the two effects described above prevails.[47]

Proposition 6 (The role of geographic mobility) *The welfare function $W(\varphi, \tau)$ is such that:*

(a) If $\widehat{\tau} \geq 0$ and $\chi/(1+\chi) > \eta$, then:

$$\frac{\partial W(\varphi, \widehat{\tau})}{\partial \tau} < 0.$$

[45] Formally, $\widehat{\tau}$ is such that the first row of equation (28) is equal to zero. It can be shown that, if $K = 1$, the function W is globally concave in τ.

[46] The reason the terms $\kappa(1-p)$ cancel out is that the aggregate moving costs, a net welfare loss, are exactly offset by the increase in average log earnings enjoyed by movers, a net gain. The latter point emerges clearly in equation (18) of Proposition 4.

[47] It is important to emphasize that these conditions are only sufficient, and they are quite restrictive. Even when $\chi = 0$ and so condition (a) is not satisfied, the planner might have an incentive to reduce τ relative to the one-location economy. Unfortunately, less restrictive sufficient conditions are not as simple as those in Proposition 6.

(b) If $\widehat{\tau} \leq 0$ *and* $\chi/(1+\chi) < \eta$*, then:*

$$\frac{\partial W(\varphi, \widehat{\tau})}{\partial \tau} > 0.$$

These conditions have an intuitive interpretation. Suppose that in the one-location version of the model the optimal tax system is progressive ($\widehat{\tau} > 0$). Proposition 4 then implies that, in the two-location version of the model ($K = 2$), the average productivity difference between movers and stayers is relatively large because movers need to overcome the tax hurdle to be induced to migrate. All else equal, this increases the marginal productivity benefit of reducing τ below $\widehat{\tau}$. This effect is larger the smaller is η, as was discussed after equation (27). By contrast, a higher χ contributes to magnify this effect because of the existence of a positive gap between the social and private benefits of increasing productivity. If χ is large relative to η, the social value of increasing productivity by reducing τ exceeds the loss associated with a higher cross-sectional dispersion in consumption. An analogous reasoning explains why, if the one-location tax system system is regressive ($\widehat{\tau} < 0$) and η is large relative to χ, the planner would like to increase τ above $\widehat{\tau}$.[48]

The main take-away from Proposition 6 is that while tax progressivity has unambiguous effects on labor mobility, the new endogenous mobility channel highlighted in this paper might contribute to either reduce or increase the welfare-optimal τ relative to the one-location version of the model. In the next section I use the full quantitative model to study the optimal degree of tax progressivity and the contribution of the labor mobility channel to its determination.[49]

[48] Since the value of $\widehat{\tau}$ is itself a function of the model's parameters, I have verified numerically that there exist different sets of parameter values such that cases (a) and (b) apply.

[49] In Online Appendix 6.4 I study the implications of expanding the set of policy tools to include a migration subsidy. In this situation the planner would set the subsidy rate so as to equalize the social marginal benefits and costs of migration, while setting $\tau = \widehat{\tau}$, the optimal degree of tax progressivity in the one-location version of the model.

6.2 Quantitative Analysis

I now study optimal tax progressivity and public good provision policies in this economy starting from an initial steady state corresponding to a constant $\tau = 0.192$, the value estimated from U.S. data. At $t = 1$, the social planner chooses tax progressivity to maximize a utilitarian welfare criterion. The latter includes both the remaining lifetime utility of agents that are alive at $t = 1$ as well as the lifetime utility of all cohorts that are not born yet, discounting their utility using the agents' own discount factor β. The welfare criterion W takes the following form:

$$
(29) \qquad W = \underbrace{(1-\beta)\,\bar{a}^{-1}\sum_{r=1}^{\bar{r}}\omega_r\sum_{a=1}^{\bar{a}}E_z\left\{E_{\boldsymbol{\varepsilon}}\left[P_1\left(a,z,\boldsymbol{\varepsilon};r\right)|z,a,r\right]\right\}}_{\text{remaining lifetime utility of agents alive at } t=1} + \underbrace{(1-\beta)\,\bar{a}^{-1}\sum_{r=1}^{\bar{r}}\omega_r\sum_{t=2}^{\infty}\beta^{t-1}E_z\left\{E_{\boldsymbol{\varepsilon}}\left[P_t\left(1,z,\boldsymbol{\varepsilon};r\right)|z,1,r\right]\right\}}_{\text{discounted lifetime utility of agents born in } t\geq 2}.
$$

Notice that social welfare is therefore computed taking into account the entire transition of the economy from its initial steady state to its final one. The planner is subject to the government's budget constraint (9), which requires a period-by-period budget balance. Since the government's budget imposes a relationship among the objects of the triple (τ_t, λ_t, G_t) at each point in time, the planner needs to optimize only with respect to two of them, taking (9) into account. In what follows I express the problem as that of finding the optimal level or sequence of τ_t and the share of aggregate income spent on public goods, defined as $\varphi_t \equiv G_t/Y_t$. It turns out that, given the log utility specification (3), these two dimensions of optimization are independent of one another.[50] In particular, the optimal share of public expenditures can be obtained analytically, as in Heathcote et al (2017).

Proposition 7 (Optimal public good provision) *The welfare maximizing share of pub-*

[50] It follows that the optimal value of φ_t is the same independently of whether the planner chooses a unique value of τ or the entire sequence $\{\tau_t\}_{t=1}^{\infty}$.

lic goods in aggregate income is constant over time and given by:

$$\varphi^* = \frac{\chi}{1+\chi}. \tag{30}$$

In order to set the parameter χ, I assume that public good provision in the U.S. is set optimally, according to (30). The data counterpart of G/Y for this economy is the share of the Federal government's consumption expenditures in the sum of the latter and personal (non-durable) consumption expenditures. The average share for the period 2000-2007 is 0.082, which implies a value $\chi = 0.089$.

6.2.1 Optimal Tax Progressivity

I first consider the case in which the planner announces at $t = 1$ a degree of tax progressivity τ which remains constant over time, while the economy goes into the transition phase towards its new steady state (e.g. Domeij and Heathcote (2004) and Krueger and Ludwig (2016)). The first column of Table 3 reports the welfare maximizing value $\tau^* = 0.307$ for the benchmark case.

Benchmark version		One location version ($K = 1$)	
τ^*	τ^{ss}	τ^*	τ^{ss}
0.307	0.113	0.397	0.397

Table 3: Optimal one-shot tax progressivity policy taking the transition into account (τ^*) and considering only steady state welfare (τ^{ss}).

As discussed in Section 6.1, in selecting the optimal τ, the planner trades-off the costs and benefits of redistribution. On the benefits side, higher tax progressivity provides insurance against idiosyncratic location-specific shocks $\boldsymbol{\varepsilon}$ and the initial ($a = 1$) productivity draw z. On the cost side, a more progressive tax system reduces incentives to supply labor and to move to locations and jobs with higher earnings, a distortion to a dynamic choice. There are two important features of the full model that are absent in the simplified version considered in Section 6.1: finite lifetimes and persistent idiosyncratic shocks. These two elements together imply that the incentives to migrate decline with age. The fact that older agents are relatively immobile reduces the distortions associated with redistribution in the early periods of a tax reform because the short-run elasticity of average productivity to changes in tax progressivity is smaller than the associated long-run elasticity.

The two plots in the top row of Figure 6 illustrate the impact of the one-shot reform on migration rates by age. The reform increases τ relative to its initial value, so migration rates for both types and for all model ages less than (about) $a = 20$ fall.[51] The dynamic path of average productivity for each agent type is plotted in the bottom row of Figure 6.[52] As the agents who are alive at $t = 1$ age and are replaced by young agents, the decline in age-specific migration rates produces a decline in average productivity over time until the economy settles down at a lower average productivity in the final steady state.

This discussion suggests that taking the transition into account plays a potentially important role in the planner's trade-offs between incentives and redistribution. In the short-run, at the time the reform is implemented, agents older than $a = 20$ have basically stopped moving geographically by choice. Their productivity is therefore unaffected by the tax reform. As a consequence, the planner has an incentive to set τ at a relatively high value because it understands that this margin will not be distorted. To provide a sense of the importance of the transition, consider the alternative exercise in which the planner maximizes steady state welfare, ignoring the transition all together (as in, e.g., Erosa and Koreshkova, 2007)). The optimal degree of tax progressivity in this case is $\tau^{ss} = 0.113$ (see Table 3, second column), a much smaller degree of progressivity than obtained when taking the transition into account.[53]

[51] Notice that there is no variation over time in migration rates by age because τ is a constant in the one-shot reform considered in this section.

[52] Average productivity at a point in time is the average of individual wages $\exp(z' + \alpha_{a,r})$ across all ages for a given type r and time period t.

[53] This comparison is consistent with the findings of Heathcote et al (2017, Section 6.3), who also find a smaller optimal degree of tax progressivity in an economy with reversible skill investments (a steady state

The last two columns of Table 3 show the welfare-maximizing τ in the one-location version of the model ($K = 1$). Notice that in this case taking or not the transition into account is irrelevant for the optimal τ because, absent migration, productivity evolves exogenously over time. In the one-location version of the model, the optimal τ is 9 percentage points higher than in the benchmark model. This implies that while the two economies share the same average tax rate φ^*, the average income-weighted marginal tax rate is about 8 percentage points higher in the one-location economy than in the benchmark.[54] Thus, the endogenous geographic mobility mechanism has a considerable effect on the optimal τ. These results are in accordance with those in Proposition 6, part (a), according to which, if χ is sufficiently larger than η, optimal tax progressivity in the benchmark model is smaller than in the one-location version. Notice, however, that, as discussed above, transitional dynamics considerations reduce the distortions associated with high tax progressivity relative to the steady state analysis of Section 6.1.

6.2.2 Optimal Path of Tax Progressivity

In this section I extend the analysis by considering the optimal policy *sequence* $\{\tau_t\}_{t=1}^{\infty}$ and the additional welfare gains from a time-varying policy. The existence of such gains had been hypothesized, but not verified, by Krueger and Ludwig (2016, p.96), among others.[55] The thought experiment now is as follows. Starting from the steady state with the benchmark $\tau = 0.192$, at $t = 1$ the planner select, unexpectedly, a policy sequence $\{\tau_t, \lambda_t, G_t\}_{t=1}^{\infty}$ to maximize utilitarian welfare as defined in equation (29).

The optimal sequence $\{\tau_t^*\}_{t=1}^{\infty}$ is represented in Figure 7 by the solid (purple) line. The main feature of the optimal time-varying policy is that at $t = 1$ the planner finds it optimal to set tax progressivity at a relatively high level, with τ_1^* approximately equal to 0.43, while

comparison) than in one with irreversible investment (i.e., taking the transition into account).

[54] As shown in Online Appendix 5.1, the average income-weighted marginal tax rate in the economy is $(1 - (1 - \tau)(1 - \varphi^*))$ while the average net tax rate is simply φ^*.

[55] Meghir (2016, p. 101) proposes an approach along these lines in his discussion of Krueger and Ludwig (2016)'s paper. Online Appendix 5.3 describes the numerical algorithm used to solve for the optimal time-varying policy. Dyrda and Pedroni (2018) use a similar method to study optimal time-varying taxation in the context of the Aiyagari model.

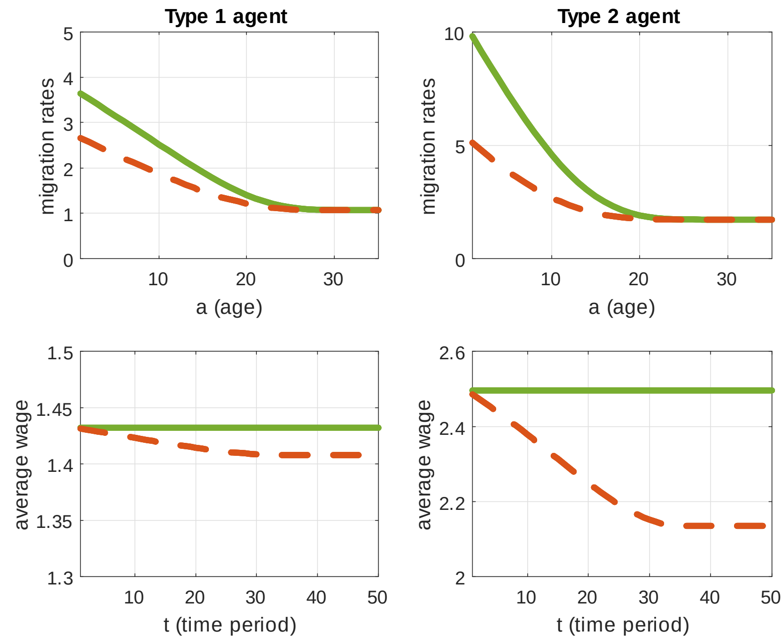

Figure 6: Average migration rates by age (top row) and average wages over time (bottom row) following the optimal one-shot tax progressivity reform. The solid (green) line corresponds to the initial steady state. The dotted (red) line corresponds to the optimal τ^*.

announcing a declining path of τ_t^* for subsequent periods. The optimal path declines until it reaches its final steady state value around $t = 40$. The intuition behind this declining path hinges crucially on the observation that high degrees of tax progressivity in the early periods are less distortionary, in terms of reduced incentives for geographic mobility, than the expectation of high degrees of tax progressivity in subsequent periods. Higher levels of τ_t increase taxes on high income households and increase subsidies to low income ones. Since income is the product of labor effort and productivity, relatively rich households may respond in the short-run by reducing their work effort but cannot escape higher taxes by reducing their productivity. In other words, at $t = 1$, the distribution of productivity in the population is fixed and redistribution from high to low productivity households has a lump-sum component akin to unexpectedly taxing physical capital in the neoclassical growth model. The planner however understands that keeping high levels of τ_t over time is suboptimal because it induces households *not* to pursue job opportunities that would increase their income over time. This reduced incentive to move would lead to a decline in average productivity and therefore in

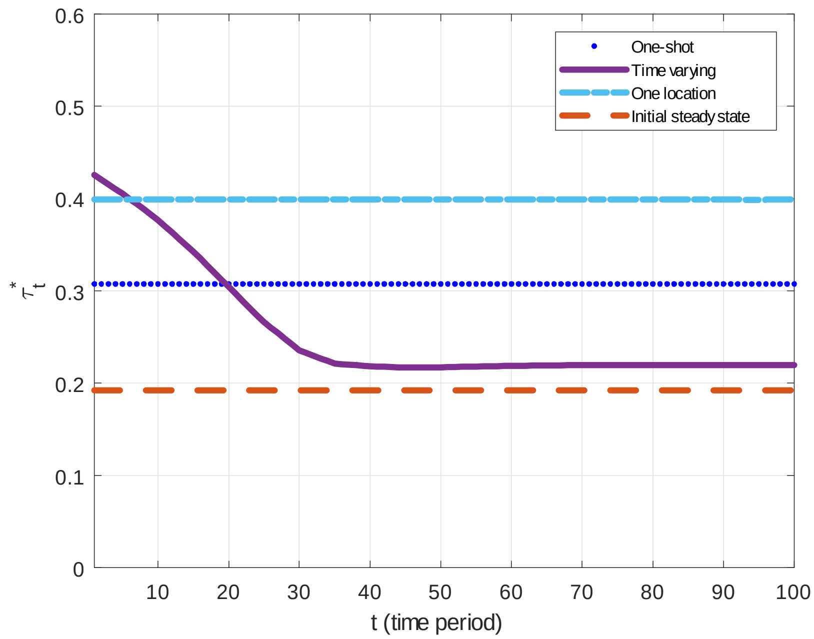

Figure 7: Optimal tax progressivity reform in various scenarios. The solid (purple) line represents the optimal time-varying policy $\{\tau_t\}_{t\geq 1}$. The dotted (dark blue) line represents the optimal one-shot policy. The dash-dotted (light blue) line represents the optimal one-shot and time-varying policies (they are the same) for the version of the model with one location ($K = 1$). The dashed (red) line represents the initial steady state $\tau = 0.192$.

the economy's tax base. Hence, the declining pattern of τ_t^* over time.[56]

As already discussed in the previous section, the intuition for this result is confirmed if one considers the version of the model with one location ($K = 1$), shutting down all migration opportunities. In this circumstance, the optimal policy reform calls for a constant τ because the evolution of productivity is exogenous and independent of the degree of progressivity. The planner does not face any dynamic trade-off and its optimization problem becomes a static choice between the costs of labor supply distortions and the benefits of redistribution. This case is represented by the solid horizontal line around 0.4 in Figure 7 (labelled "One location"). For reference, the figure also presents the one-shot optimal policy $\tau^* = 0.307$ (dotted line) discussed in the previous section and the calibrated $\tau = 0.192$ in the initial steady state (dashed line).[57]

Figure 8 represents the evolution of life-cycle profiles of migration for a number of cohorts over the transition with time-varying tax progressivity. For reference I have included the profile corresponding to the initial steady state as a dashed line starting at $t = 1$. Notice how the reform reduces, on impact, mobility rates of the cohort born in period $t = 1$, especially type 2 agents. Smaller mobility translates into a lower life-cycle profile of productivity and wages. Over time as optimal tax progressivity falls migrations rates increase at each age and so do lifecycle profiles of wages. However, they remain depressed relative to the initial steady state because optimal tax progressivity remains asymptotically larger than its value in the original steady state.

[56] In a similar way, in the neo-classical growth model physical capital is more elastic in the long-run than in the short-run so a planner might want to tax it more heavily earlier than later. See Straub and Werning (2015) for a recent discussion of the optimal capital taxation literature. Notice that in the optimal taxation literature in the Ramsey tradition (Jones et al (1993)), upper bounds on capital tax rates have to be imposed early on to prevent the planner from effectively taxing capital in a lump-sum fashion in the early period of the reform. In my model, the presence of elastic labor supply endogenously reduces the planner's incentives to redistribute in the early periods of the reform.

[57] Figure A.1 in Appendix 5.2.2 presents the optimal tax progressivity sequence for different values of the model's parameters. In all these cases, the endogenous migration mechanism is active and the shape of optimal tax progressivity over time is qualitatively similar to the one in the benchmark, although it might be shifted upwards of downwards according to the specific parameter that is changed.

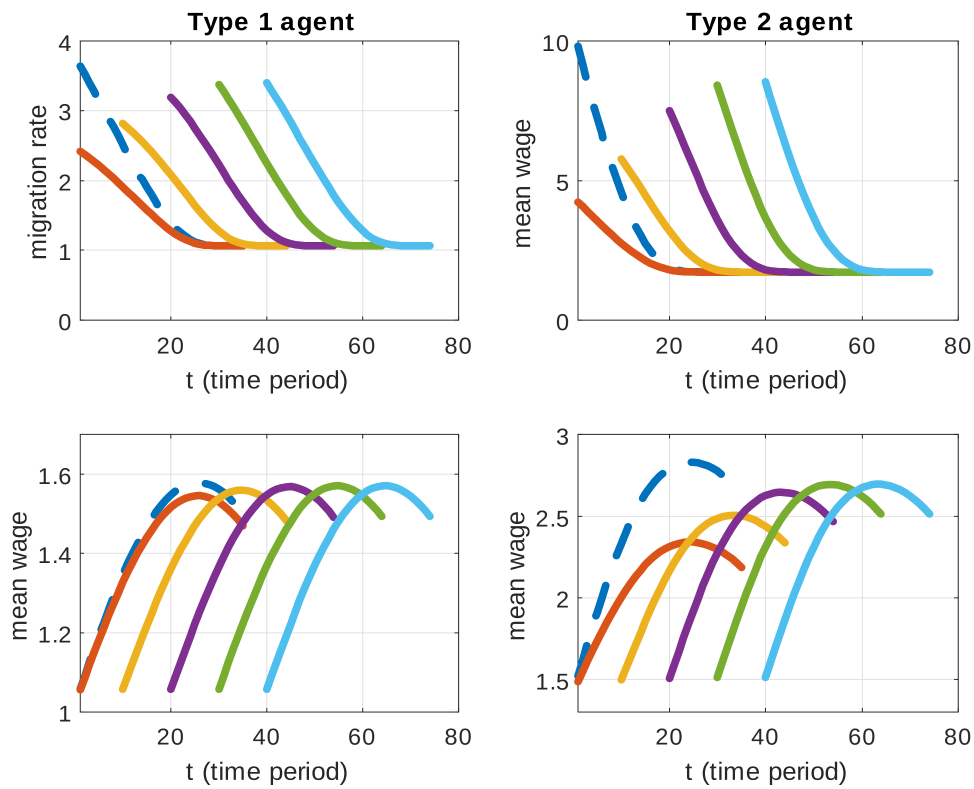

Figure 8: Average migration rates and average wages by age and cohort in the transition following the optimal time-varying tax progressivity reform. The figures represent profiles over the life cycle for selective cohorts born in periods $t = 1, 10, 20, 30, 40$. For reference I present the pattern corresponding to the initial steady state as a dotted (– –) line.

6.2.3 Welfare Effects

In this section I discuss the welfare gains from tax progressivity reform. Welfare gains in consumption equivalent units are defined as the proportional increase in households' period consumption in the benchmark economy (with $\tau = 0.192$) that makes social welfare the same as under the optimal policy. Formally, given the logarithmic specification of utility, the equivalent variation is

$$s^* = \exp\left(W^* - W^{\text{bench}}\right) \tag{31}$$

where W^* denotes the welfare level corresponding to the optimal degree of tax progressivity under consideration and W^{bench} welfare in the model's original steady state. The first row of Table 4 reports s^* corresponding to the one-shot, time-varying, and steady-state-optimal tax progressivity reform, respectively.

	Optimal Policy		
	One-shot $\tau^* = 0.307$	Time-varying $\{\tau_t^*\}_{t=1}^{\infty}$	Steady-state $\tau^{ss} = 0.113$
s^*	0.706%	1.090%	−1.505%
s^{ss}	−1.775%	−0.352%	0.508%

Table 4: Summary of welfare effects of tax progressivity reform. The equivalent variation s^* is defined in equation (31). The equivalent variation s^{ss} is computed in a similar way, except that it only compares social welfare across initial and final steady states.

In order to compare with the previous literature, the second row of Table 4 also reports the equivalent variation, s^{ss}, obtained by comparing initial and final steady states, ignoring the transition.

The welfare gain from the optimal one-shot policy is 0.706 percent. By contrast, the welfare gain of the optimal time-varying policy is equivalent to 1.090 percent of lifetime consumption, or about 54 percent larger than the gain from the optimal one-shot policy. Consistent with the previous literature (Krueger and Ludwig, 2016), adopting a policy that is optimal when comparing steady states may lead to welfare losses once the transition is taken into account. The welfare loss in this case is 1.505 percent of lifetime consumption. As shown in the second row of Table 4, this same policy increases welfare by 0.508 percent when comparisons are restricted to steady states. The table also illustrates how, adopting policies that are optimal taking the transition into account may lead to welfare losses when comparing steady states.

Who gains and who loses from the optimal policy during the transition? In order to answer this question Proposition 1 in Online Appendix 5.4 shows how to decompose the expression for s^* into a weighted average of type-age-cohort specific equivalent variations. This decomposition reveals that, relative to the benchmark economy, both optimal policies (one-shot and sequence of tax progressivity) favor type 1 agents and hurt type 2 agents independently of their age and birth cohort. In addition, the one shot reform tends to produce more evenly distributed welfare gains and losses across various age groups and cohorts than the reform that optimizes over the transition. The latter produces larger gains and losses for the cohorts who are alive at the time of the reform and smaller welfare effects for the cohorts that are born after period 10.

7 Conclusions and Future Work

I have constructed an analytically tractable, yet rich, dynamic model of internal migration to study quantitatively the relationship between tax progressivity and the geographic mobility of labor. Higher degrees of tax progressivity reduce households' incentives to undertake costly migration to take advantage of higher earnings opportunities. I use the model to compute the optimal tax progressivity policy and show that the migration channel plays a quantitatively important role in determining both the level and the time-path of optimal tax progressivity relative to a one-location economy.

A more general contribution of the paper is to introduce a new dynamic model of internal migration that is analytically tractable. The framework proposed here can be extended in a relatively straightforward fashion to allow for heterogeneity in productivity and amenities across locations, and for additional geographic details, such as moving costs that depend on distance. This flexibility makes the framework suitable to investigate the implications of other policies and shocks that lead to geographic reallocation of labor. A particularly interesting application of this framework would be to study the aggregate implications of heterogeneity in the level of taxes across U.S. states (see Fajgelbaum et al, 2015).

It would be interesting, but not straightforward, to relax some of the assumptions that make the model analytically tractable and evaluate their impact on optimal tax progressivity. For example, if agents could self-insure against labor income shocks, the welfare benefits of government-provided insurance might be smaller, leading to a larger reduction in optimal tax progressivity relative to the one location version of the model. For the same reason, allowing self insurance would also change the optimal path of tax progressivity, inducing the planner to select smaller degrees of tax progressivity in the early periods following a tax reform. Another extension would consider the extensive margin of labor supply, allowing one to evaluate the impact on migration of policies, such as disability payments, that are tied to non-participation in the labor market. I leave these and other extensions to future research.

Bibliography

Abbott B., Gallipoli G., Meghir C. and G. Violante, "Education Policy and Intergenerational Transfers in Equilibrium," *Journal of Political Economy*, 127 (2019), 2569–2624.

Ahlfeldt G., Redding S., Sturm D, and Wolf N., "The Economics of Density: Evidence from the Berlin Wall," *Econometrica,* 83 (2015), 2127-2189.

Akcigit U., Baslandze, S. and Stantcheva S., "Taxation and the International Mobility of Inventors," *American Economic Review*, 106 (2016), 2930–2981.

Albouy D., "The Unequal Geographic Burden of Federal Taxation," *Journal of Political Economy*, 117 (2009), 635–667.

Albouy D., Behrens K., Robert-Nicoud F., Seegert N., "The Optimal Distribution of Population across Cities," *Journal of Urban Economics*, 110 (2019), 102-113.

Amior M., "Education and Geographical Mobility: The Role of the Job Surplus," mimeo, Hebrew University of Jerusalem, 2020.

Artuc, Erhan, Shubham Chaudhuri, and John McLaren, "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100 (2010), 1008–45.

Auray S., Fuller D., Lkhagvasuren D., and Terracol A., "Dynamic Comparative Advantage, Directed Mobility Across Sectors, and Wages," CREST working paper n. 2017–59, 2017.

Badel A., M. Hugget, and W. Luo (2020): "Taxing Top Earners: A Human Capital Perspective," *The Economic Journal*, Vol. 130, Issue 629, pages 1200–1225.

Bakis O., Kaymak B., and M. Poschke, "Transitional Dynamics and the Optimal Progressivity of Income Redistribution," *Review of Economic Dynamics*, 18 (2015), 679–693.

Bayer C. and F. Juessen, "On the Dynamics of Interstate Migration: Migration Costs and Self-Selection," *Review of Economic Dynamics*, 15 (2012), 377–401.

Benabou R., "Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 46 (2002), 263-303.

Bovenberg L. and Jacobs B., "Redistribution and Education Subsidies are Siamese Twins," *Journal of Public Economics* 89 (2005), 2005–2035.

Caliendo L., Dvorkin M., Parro F., "Trade and Labor Market Dynamics: General Equilibrium Dynamics of the China Trade Shock," *Econometrica*, 87 (2019), 741-835.

Coen-Pirani D., "Understanding Gross Worker Flows across U.S. States" *Journal of Monetary Economics*, 57 (2010), 769-784.

Colas M. and K. Hutchinson, "Heterogeneous Workers and Federal Income Taxes in a Spatial Equilibrium," forthcoming *American Economic Journal: Economic Policy,* 2020.

Congressional Budget Office, *Trends in the Distribution of Household Income Between 1979 and 2007,* The Congress of the United States, 2011.

Desmet K., Nagy D.K., and Rossi-Hansberg E., "The Geography of Development," *Journal of Political Economy*, 126 (2018), 903-983.

Diamond P. and E. Saez, "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 25 (2011), 165–190.

Diamond R., "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106 (2016), 479–524.

Domeij D. and Heathcote J., "On the Distributional Effects of Reducing Capital Taxes," *International Economic Review*, 45 (2004), 523-554.

Dyrda S and M. Pedroni, "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks," mimeo, University of Toronto, 2018.

Eeckout J. and N. Guner, "Optimal Spatial Taxation: Are Big Cities Too Small?," IZA Discussion Paper No. 8781, 2015.

Erosa, A. and T. Koreshkova, "Progressive Taxation in a Dynastic Model of Human Capital," *Journal of Monetary Economics*, 54 (2007), 667-85.

Fajgelbaum P., and Gaubert, "Optimal Spatial Policies, Geography, and Sorting," *The Quarterly Journal of Economics*, 135 (2020), 959–1036.

Fajgelbaum P., Morales E., Serrato J.C. and Zidar O., "State Taxes and Spatial Misallocation," *Review of Economic Studies*, 86 (2019), 333–376.

Ferriere A. and G. Navarro, "The Heterogeneous Effects of Government Spending: It's All About Taxes," International Finance Discussion Papers 1237, Board of Governors of the

Federal Reserve System, 2018.

Gelbach J. (2004): "Migration, the Life Cycle, and State Benefits: How Low Is the Bottom?" *Journal of Political Economy*, Vol. 112, No. 5, pp. 1091-1130.

Gemici A., "Family Migration and Labor Market Outcomes," mimeo Royal Holloway College London, 2017.

Gentry W. and G. Hubbard, "The Effects of Progressive Income Taxation on Job Turnover," *Journal of Public Economics*, 88 (2004), 2301– 2322.

Guner N., Kaygusuz R., and G. Ventura, "Income Taxation of U.S. Households: Facts and Parametric Estimates," *Review of Economic Dynamics*, 17 (2014), 559–581.

Guvenen F., Kuruscu B., and S. Ozcan, "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," *Review of Economic Studies*, 81 (2014), 818–850.

Hansen L. P., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50 (1982), 1029–1054.

Hassler J., Rodriguez-Mora J.-V., Storesletten K, Zilibotti F., "A Positive Theory of Geographic Mobility and Social Insurance," *International Economic Review,* 46 (2005), 263–303.

Heathcote J., Storesletten K., and G. Violante, "Optimal Tax Progressivity: An Analytical Framework," *Quarterly Journal of Economics*, 132 *(*2017), 1693–1754.

Heathcote J., Storesletten K., and G. Violante, "Optimal Progressivity with Age-Dependent Taxation," forthcoming, *Journal of Public Economics*, (2020a).

Heathcote J., Storesletten K., and G. Violante, "How Should Tax Progressivity Respond to Rising Income Inequality?" *Journal of the European Economic Association,* 18 (2020b), 2715–2754.

Jones L., Manuelli R., and P. Rossi, "Optimal Taxation in Models of Endogenous Growth," *Journal of Political Economy*, 101 (1993), 485–517

Kaplan G. and S. Schulhofer-Wohl, "Understanding the Long-Run Decline in Interstate Migration," *International Economic Review*, 58 (2017), 57–94.

Kennan J. and R. Walker, "The Effect of Expected Income on Individual Migration

Decisions," *Econometrica*, 79 (2011), 211–251.

Koeniger W. and J. Prat, "Human Capital and Optimal Redistribution," *Review of Economic Dynamics*, 27 (2018), 1–26.

Krueger D. and A. Ludwig, "Optimal Progressive Labor Income Taxation and Education Subsidies When Education Decisions and Intergenerational Transfers Are Endogenous," *American Economic Review*, 103 (2013), 496–501.

Krueger D. and A. Ludwig "On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrium," *Journal of Monetary Economics*, 77 (2016), 72–98.

Lagakos D., Mobarak M., and Waugh M., "The Welfare Effects of Encouraging Rural-Urban Migration," mimeo, University of California San Diego 2017.

Lkhagvasuren D., "Education, Mobility and the College Wage Premium," *European Economic Review*, 67 (2014), 159–173.

Manovskii I., "Productivity Gains from Progressive Taxation of Labor Income," mimeo, University of Pennsylvania 2002.

Meghir C., "A comment: 'On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrium' by Dirk Krueger and Alexander Ludwig," *Journal of Monetary Economics* 77 (2016), 99–102.

Meghir C. and L. Pistaferri, "Earnings, Consumption and Life Cycle Choices," *Handbook of Labor Economics*, Elsevier, 4 (2011), Part B, 773–854

Monras J., "Economic Shocks and Internal Migration," mimeo, Universitat Pompeu Fabra, 2017.

Moretti E. and Wilson D., "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists," *American Economic Review*, 107 (2017), 1858–1903.

Mustre del-Rio J. and Hawkins W., "Financial Frictions and Occupational Mobility," Federal Rserve Bank of Kansas City Research Working Paper 12-06, 2017.

Nakosteen R. and M. Zimmer, "Migration and Income: The Question of Self-Selection," *Southern Economic Journal*, 46 (1980), 40–851.

Notodiwigdo M., "The Incidence of Local Labor Demand Shocks," NBER Working Paper 17167, 2011.

Organization for Economic Cooperation and Development, *Employment Outlook*, June 2000.

Organization for Economic Cooperation and Development, *Employment Outlook*, July 2005.

Organization for Economic Cooperation and Development, *Regions at a Glance*, December 2013.

Prescott E., "Why do Americans Work so much more than Europeans?," Federal Reserve of Minneapolis Quarterly Review, July 2004.

Ruggles S., Genadek K., Goeken R., Grover J., and M. Sobek. Integrated Public Use Microdata Series: Version 7.0 [dataset]. Minneapolis, MN: University of Minnesota, 2017.

Rust, J., "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55 (1987), 999–1033.

Schultz T., "Investment in Human Capital," *American Economic Review*, 51 (1961), 1–17.

Stantcheva, V., "Optimal Taxation and Human Capital Policies over the Life Cycle," *Journal of Political Economy*, 125 (2017), 1931–1990.

Straub L. and I. Werning, "Positive Long-Run Capital Taxation: Chamley-Judd Revisited," mimeo, MIT 2015.

Survey of Income and Program Participation Users' Guide, Third Edition, Washington, D.C. 2001.

Yankow J., "Migration, Job Change, and Wage Growth: A New Perspective on the Pecuniary Return to Geographic Mobility," *Journal of Regional Science*, 43 (2003), 483–516.