Optimal Dynamic Spatial Policy [*]

Eric Donald[†] — *University of Pittsburgh*

Masao Fukui[‡] — *Boston University*

Yuhei Miyauchi[§] — *Boston University*

July 2025

Preliminary and Incomplete

Abstract

What is the optimal spatial distribution of population? Although being a central policy issue, existing answers are based on a static framework. We fill this gap by studying the constrained efficient allocation of dynamic spatial equilibrium models with frictional migration. The key constraint is that the location preference shocks are private information. We show that the impementation of the constrained efficient allocation only requires the allocation to depend on the history of living locations. We then provide a recursive formula that the constrained efficient allocation must satisfy, summarizing the dynamic trade-off between providing consumption insurance and incentivizing efficient migration. We apply our framework to the U.S. states to study welfare gains from moving to the constrained efficient allocation from a status quo economy with exogenously incomplete market. In the steady state, we find that substantial welfare gains can be achieved not only by reducing spatial inequality but also by reallocating population toward more productive states through dynamic incentives. This highlights the important role of dynamic incentives in overcoming the trade-off between spatial inequality and efficiently allocating the population across space. Along the transition in response to localized productivity shocks, the constrained efficient allocation features slower transitions relative to the status-quo economy with an exogenously incomplete market.

[*]We thank Costas Arkolakis, Lorenzo Caliendo, Jonathan Dingel, David Lagakos, Michael Peters, and Pascual Restrepo as well as seminar participants at Barcelona Summer Forum, Boston University, Green Line Macro Meeting, SED, University of Pittsburgh, and Yale for their helpful comments.

[†]University of Pittsburgh (e-mail: eric.donald@pitt.edu).

[‡]Boston University (e-mail: mfukui@bu.edu).

[§]Boston University (e-mail: miyauchi@bu.edu).

1 Introduction

What is the optimal spatial distribution of population? How does it differ from the observed equilibrium? And what policies can improve aggregate welfare? These questions lie at the center of ongoing academic and policy debates surrounding the design of place-based interventions (Fajgelbaum and Gaubert 2025). Yet, most of the existing literature addresses them in a static setting. This overlooks an important fact that migration is slow and subject to frictions (Caliendo, Dvorkin, and Parro 2019). The focus on static settings also limits the ability to speak about the optimal response to regional shocks. In this paper, we examine these questions through a dynamic perspective.

We study constrained efficient allocation in a dynamic spatial general equilibrium model in which agents make forward-looking migration decisions. In each period, individuals receive idiosyncratic location preference shocks and choose their location accordingly. A social planner allocates population to each location every period, subject to the constraint that the preference shocks are private information. We derive a recursive formulation of the constrained optimum, highlighting the core trade-off between consumption smoothing and efficient migration. When applied to the U.S. economy, our model predicts that the steady-state constrained efficient allocation features lower population and higher per capita consumption in less productive regions sustained through a dynamic incentive to move out from unproductive regions. We also examine the planner's dynamic response to regional productivity shocks and compare it with decentralized equilibrium outcomes.

We consider a general environment with many heterogeneous locations that differ in the paths of productivity, amenities, trade costs, and agglomeration externalities. In each period, households draw idiosyncratic preference shocks and choose their optimal residential location for the following period, following the framework of Artuç, Chaudhuri, and McLaren (2010) and Caliendo et al. (2019). A social planner allocates consumption and assigns locations, but must respect the fact that preference shocks are privately observed. As a result, the planner cannot directly control migration decisions and must design allocations that are incentive-compatible.

The main challenge in solving the constrained-efficient allocation arises from the high dimensionality of the state space and the choice variables. The planner must track the full distribution of agents by migration history and assign consumption and locations accordingly. We show that this complex problem can be decomposed into tractable subproblems: for each agent, the planner determines consumption and location assignments conditional only on their current location and promised continuation utility. Because each sub-problem depends only on each agent's current location and promised utility, this structure significantly reduces the dimensionality of the problem, yielding both analytical and computational tractability.

We derive a recursive formulation that fully characterizes the constrained-efficient allocation and formalizes the central trade-off facing the planner. On the one hand, the planner seeks to incentivize migration to more productive regions. This can be achieved by offering either higher contemporaneous consumption or by committing to higher future consumption conditional on remaining in those locations. On the other hand, the planner also values consumption smoothing, especially for agents who, due to their idiosyncratic shocks, remain in less productive areas where the marginal utility of consumption is higher.

Our formula reveals that optimal allocation is inherently dynamic: agents who relocate to productive regions initially receive lower consumption that increases over time, while those in unproductive regions receive higher initial consumption that declines over time. These dynamic incentives can support a steady-state allocation in which less productive regions feature lower population and higher per capita consumption, compared to the market equilibrium. Therefore, it is possible to achieve redistribution toward unproductive regions and reallocation toward productive ones simultaneously. We demonstrate that this pattern indeed arises in our model calibrated to the U.S. economy.

Our analysis highlights the central role of dynamic incentives in achieving the constrained-efficient allocation. To further illustrate this point, we consider an alternative planning problem in which consumption allocations depend solely on an agent's current location, rather than their full migration history – a setting we call the history-independent constrained-efficient allocation. We derive a corresponding recursive formulation, which reveals a more limited scope for dynamic incentives, as consumption can no longer be tailored based on past migration decisions. As a result, the planner faces a sharper trade-off between incentivizing migration to productive regions and smoothing consumption in unproductive ones. While suboptimal relative to the (history-dependent) constrained efficient allocation, the benefit of history-independent constrained-efficient allocation is less demanding to implement. Indeed, we show that, under the common assumption that migrating agents are hand-to-mouth, the history-dependent allocation can be implemented through location- and time-specific transfers.

We calibrate our model assuming that the status quo economy features an exogenously incomplete market in the spirit of Bewley-Huggett-Aiyagari (Bewley 1986, Huggett 1993, Aiyagari 1994, Imrohoroğlu 1989). We calibrate our economy to the 2017 US States to flexibly match bilateral trade and migration flows. We discipline the ability of households to smooth consumption by matching the empirical estimates of the marginal propensity to consume.

Armed with our calibration, we solve the steady state constrained efficient allocation. As noted earlier, we find that the constrained efficient allocation features less spatial inequality in consumption and more population concentration toward productive locations. The absence of trade-off between inequality and efficiency can be explained by the dynamic incentive. Consistent with our theoretical analysis, consumption profile is backloaded in the productive locations, incentivizing households to stay. In contrast, the consumption profile is front-loaded in unproductive locations, incentivizing households to leave. By appropriately choosing the level of consumption, the planner can simultaneously achieve low spatial inequality and efficient migration. We find welfare gains from moving from the status quo to constrained efficient allocation to be around 4%, measured in consumption equivalent units of utilitarian welfare.

We then turn to the history independent allocation. With the lack of dynamic incentives, we find the planner faces a substantial trade-off in spatial inequality and efficiency. In our calibration, the planner balances the trade-off by creating substantial spatial inequality to acheive the efficient migration decisions. This undermines the welfare gains by 0.5%.

In the final part of the paper, we study the constrained efficient response to the localized technology shocks. We first show that aggregate shocks can be tractably studied by focusing on the one-time shock that occurs with arbitrarily small probability. This is distinct from "MIT shock" in that the shock is anticipated so that the planner writes contingent plans in response to the shock. This circumvents the issue of time-inconsistency problem that arises if we were to study an unanticipated "MIT shock." We then show that the aggregate response can be obtained solely using the sequence space Jacobian (Auclert, Bardóczy, Rognlie, and Straub 2021) around the deterministic steady state.[1]

Relying on the above approach, we study a permanent negative productivity shock in one location. We find that the constrained efficient allocation involves more population reallocation in the short-run and less reallocation in the long-run, relative to the status-quo economy with exogenously incomplete market.

Related Literature First, we contribute to the literature of dynamic spatial general equilibrium models. These frameworks have been used for various applications, such as regional incidence of import competition, the rise of automation, immigration shocks, or climate change (see Desmet and Parro (2025) for a recent survey). Earlier work has models hand-to-mouth agents building on the framework of Caliendo et al. (2019), and more recent work has introduced agents' saving decisions (e.g., Bilal and Rossi-Hansberg 2023, Giannone, Li, Paixão, and Pang 2023, Greaney, Parkhomenko, and Van Nieuwerburgh 2025).

So far, the literature has remained largely silent about the optimal allocation and policies in that environment. An important recent exception is O'Connor (2024), who study the role of spatial transfers to address dynamic inefficiency resulting from wage rigidity. He provides an analytical formula for the optimal transfers for a two-period model and quantitatively imple-

[1]Similar approaches appear in the context of optimal risk-sharing contracts (Fukui 2020) and endogenous portfolio choice (Auclert, Rognlie, Straub, and Tapak 2024). Mukoyama (2021) clarifies the difference between the "MIT shock" and an anticipated shock with arbitrary small probability.

ments it with U.S. commuting zones using a linearized dynamic model.[2] Our paper is distinct in two important ways. First, instead of focusing on particular sources of externalities, we work with general infinite-horizon, fully-dynamic spatial equilibrium models and provide an analytical formula for its constrained efficient allocation. Second, besides externalities, our analytical results highlight the incentive-insurance trade-off resulting from information constraints for the planner.

Our paper also extends the literature of optimal policies in spatial equilibrium models. This literature focuses on how spatial policies can address externalities and redistribution (see Fajgelbaum and Gaubert 2025 for a recent review). In this static literature, our paper is particularly close to Gaubert, Kline, and Yagan (2021) and Guerreiro, Rebelo, and Teles (2023), who characterize the constrained efficient allocations subject to private information about households' skills and preference shocks, in the tradition of public finance literature (Mirrlees 1971). We extend these analyses to a dynamic environment.

We also contribute to the literature on macroeconomics and public finance that aims to characterize constrained efficient allocation of dynamic general equilibrium models under the planner's information constraints. In particular, our recursive formulation of constrained efficient allocation extends Atkeson and Lucas (1992), Farhi and Werning (2007), Veracierto (2022) to a dynamic discrete migration choice environment. We develop a computational method to handle many heterogeneous locations and implement them in the U.S. economy. Kurnaz, Michelini, Özdenören, and Sleet (2023) analyzes optimal tax design in a dynamic discrete choice choice with hand-to-mouth agents and characterizes its property in the steady state. Our paper differs in that we characterize constrained efficient allocations by taking a primitive approach of solving the planner's problem subject to information constraints and characterize both steady state and transition dynamics.

2 Model Environment

We consider an economy consisting of J locations. Time is discrete and the horizon is infinite, $t = 0, 1, \ldots, \infty$. There is a unit measure of household dynasty indexed by $h \in [0, 1]$, each of which is endowed with a unit of labor. This section defines the model environment. We introduce specific market structures later when we discuss implementations and calibration.

Timing and Demographics. At the beginning of the period t, households living in location j consumes the location-specific consumption basket. Households then die with probability $1 - \omega$,

[2] Lhuillier (2023) also characterizes the optimal allocation with dynamic learning externality among workers in a stylized overlapping dynamic spatial general equilibrium model.

and they are replaced by newborns. The newborns are born in location j. Surviving households and newborns then draw idiosyncratic location preference shocks and decide where to live for the next periods. We now describe each step in details below.

Preferences. In each period t, households living in location j consume location-specific final good aggregator. We denote the consumption of goods produced in location j and time t as C_{jt}. The households' flow utility from the consumption in location j is given by $u_{jt}(C_{jt})$.

At the end of period t, households die with probability $1-\omega$. Whenever households die, they are replaced by newborns in the same location j. Then, both surviving households and newborns in location j draw a vector of idiosyncratic preference shocks and decide where to migrate. The location preference shocks $\epsilon_{jt} \equiv [\epsilon_{jkt}]_k$ are additively separable from $u_{jt}(C_{jt})$. These preference shocks $[\epsilon_{jkt}]_k$ capture factors determining migration decisions specific to households that are difficult to observe (e.g., find a new job, friends or relatives living in town). We do not impose particular assumptions on its distribution function $G_{jt}(\epsilon_{jt})$, such as the independence across alternative options k or extreme-value distribution, as commonly assumed in the literature (e.g., Artuç et al. (2010), Kennan and Walker (2011), and Caliendo et al. (2019)). However, we do assume that households draw of $[\epsilon_{jkt}]_k$ are serially uncorrelated, in line with virtually all existing works.[3] Notice also that the mean of ϵ_{jkt} can arbitrarily depend on origin j, destination k, and time t, which capture the migration utility costs that depend on location pairs and time. Throughout, we denote $\mathbb{E}_{it}$ as the expectation operator over ϵ_{it}, i.e., $\mathbb{E}_{it}[x] \equiv \int x dG_i$.

After observing the preference shocks, both surviving households and newborns can migrate. All households discount the future with discount factor $\beta \in (0,1)$. Let $\ell_t(h) \in \{1,\ldots,J\}$ be the location that household dynasty h lives, and $C_{jt}(h)$ be the final goods consumption of location j by household dynasty h at time t. The value function of household h living in location j at time t are recursively given by

$$v_{jt}(h) = u_{jt}(C_{jt}(h)) + \beta\omega\mathbb{E}_{jt}\left[\sum_j \mathbb{I}(\ell_t(h)=k)\{v_{kt+1}(h)+\epsilon_{jkt}\}\right]. \tag{1}$$

[3] Howard and Shao (2022) allows for time-series correlation in the form of generalized extreme value (GEV) distribution between time $t-1$ and t draw of preference shock for a given location. Given the nature of the GEV distribution, the distribution of preference draw at time t conditional on the choice of past location at time $t-1$ is still given by the GEV distribution that depends only on the choice of past location and calendar time t. Since we allow G_{it} to depend on location i and time t, our model nests their specification. XXX CHECK WITH GREG ABOUT THIS STATEMENT? XXX

The value function of newborns h born in location j at the end of period t is

$$v_{jt}^{n}(h) = \mathbb{E}_{jt}\left[\sum_{j} \mathbb{I}(\ell_t(h) = k)\{v_{kt+1}(h) + \epsilon_{jkt}\}\right]. \tag{2}$$

Technology. The location-specific final good aggregator is a composite of various goods, some of which can be tradable (e.g. food or manufacturing goods) or nontradable (e.g. housing or nontradable services). Instead of modeling each of these goods, we follow Adao, Costinot, and Donaldson (2017) and specify the reduced-form production technology for location-specific final good aggregators using labor services from various locations. Specifically, the non-tradable goods in each location j is produced using labor from various locations:

$$Y_{jt} = f_{jt}\left(\{l_{kjt}\}_k, \{L_{kt}\}_k\right). \tag{3}$$

where l_{kjt} is the labor service in k used for producing final goods in j at time t, and L_{kt} is the total population size of location k. We assume f_{jt} is constant returns to scale in $\{l_{kjt}\}_k$. The dependence on population size distribution $\{L_{kt}\}$ captures the agglomeration and congestion forces through which the population size affects productivity of a location.

Resource Constraints. The goods market clearing conditions are

$$\int_0^1 C_{jt}(h)dh = f_{jt}\left(\{l_{kjt}\}_k, \{L_{kt}\}_k\right), \tag{4}$$

and the labor market clearing conditions are

$$L_{jt} = \int_0^1 \mathbb{I}\left[\ell_t(h) = j\right] dh = \sum_k l_{jkt}. \tag{5}$$

Generality and Scope of the Environment. It should be clear that our model is general enough to nest many of the dynamic discrete choice literature in general equilibrium, for example, by Artuç et al. (2010) and Caliendo et al. (2019). We have not explicitly introduced endogenous amenity in our model, but we can interpret that a part of the final consumption goods in (3) consist of endogenous amenity that can be produced or affected by agglomeration and congestion forces. Note also that, although our main application is dynamic spatial models, our analysis equally applies to any other dynamic discrete choice models such as occupational and sectoral choice and the discrete choice of consumption basket (Mongey and Waugh 2025).

In order to focus on the dynamic migration margin, in the baseline model, we abstract from several considerations that appear in existing works. First, we abstract from capital accumulation in the baseline model. However, it is straightforward to introduce capital accumulation in each location, as we do in Appendix A.4. Second, we assume all households are ex-ante homogeneous, which we relax in Appendix A.5. Third, we assume agglomeration forces come from contemporaneous population distribution. In Appendix A.6, we allow for agglomeration forces to depend on lagged population distribution.

We introduce overlapping generation structure in order to induce stationary distribution in the constrained efficient allocation, as explained in Farhi and Werning (2007). Models with infinitely lived households are nested as a special case with $\omega = 1$, in which constrained efficient allocation features immiseration in the long-run (Atkeson and Lucas 1992).

3 Constrained Efficient Allocation

We first briefly discuss the first-best allocation with complete information as a benchmark. We then move to the case where idiosyncratic location preference shocks are private information. We set up the planner's problem and characterize the constrained efficient allocation. In all cases, we assume the planner seeks to maximize the weighted average of the expected utility of each generation:

$$\mathcal{W}_0 = \sum_{t=0}^{\infty} \frac{1}{R^t} \sum_{i=1}^{J} \Lambda_i v_{it}^n (1-\omega) \int_0^1 \mathbb{I}[\ell_t(h) = i] dh, \tag{6}$$

where v_{it}^n is the lifetime value of the newborns born at time t in location i, $(1-\omega)\int_0^1 \mathbb{I}[\ell_t(h) = i]dh$ corresponds to the mass of households born in location i at time t, Λ_i is the welfare weight attached to households born in i, and $R > 1$ is the social discount rate.

3.1 Complete Information Benchmark

We start from the complete information case as a theoretical benchmark. Suppose the planner observes the history of idiosyncratic preferences of households, $\epsilon^t \equiv (\epsilon_0, \ldots, \epsilon_t)$. The planner specifies the consumption and the location of living of each history of preference shocks, ϵ^t.

Following the large literature on dynamic public finance, we work with a recursive formulation of the planning problem using promised utility and continuation value instead of the sequence of allocation contingent on the history of preference shocks (e.g., Atkeson and Lucas 1992, Farhi and Werning 2007). At any given point in time, each household is identified by the promised utility v and the location of living i. Let $\phi_i(v)$ denote the measure of households in location i with promised utility v. The state variable of the planner is $\phi \equiv [\phi_i]_i$.

The planning problem in a recursive form is

$$\mathcal{W}_t(\phi) = \max_{\{C_{it}(v,\epsilon),\ell_i^n(\epsilon),\ell_i(v,\epsilon),v_i^n,v_{ij}^{n\prime}(\epsilon),v_{ij}^{\prime}(v,\epsilon),l_{ji},\phi',L_i\}} \sum_i \Lambda_i v_i^n (1-\omega) \int d\phi_i + \frac{1}{R}\mathcal{W}_{t+1}(\phi') \tag{7}$$

subject to

$$v_i^n = \beta \mathbb{E}_t \Big[\sum_j \mathbb{I}[\ell_i^n(\epsilon) = j] \left\{ v_{ij}^{n\prime}(\epsilon) + \epsilon_{i,j} \right\} \Big] \tag{8}$$

$$v = u_i(C_{it}(v)) + \beta\omega \mathbb{E}_t \Big[\sum_j \mathbb{I}[\ell_i(v,\epsilon) = j] \left\{ v_{ij}^{\prime}(v,\epsilon) + \epsilon_{i,j} \right\} \Big] \tag{9}$$

$$\int C_{it}(v) d\phi_i = f_{it}(\{l_{ki}\}_k\,; \{L_k\}_k) \tag{10}$$

$$\sum_j l_{ij} = \int d\phi_i \tag{11}$$

$$L_i = \int d\phi_i \tag{12}$$

and the law of motion of the distribution:

$$\begin{aligned} \phi_j'(\mathbb{V}) = \omega \sum_i \mathbb{E}_{it} \left[\phi_i(v_{ij}^{-1}(\mathbb{V},\epsilon)) \mathbb{I}[\ell_i(v_{ij}^{-1}(\mathbb{V},\epsilon),\epsilon) = j] \right] \\ + (1-\omega) \int d\phi_i \mathbb{E}_{it} \left[\mathbb{I}[\ell_i^n(\epsilon) = j] \mathbb{I}[v_{ij}^{n\prime}(\epsilon) \in \mathbb{V}] \right]. \end{aligned} \tag{13}$$

The first constraint (8) is the definition of the value of newborn born in location i. The second constraint (9) is the promise keeping constraint for existing generations. Note that the value of existing generations do not enter into the objective function because their values are predetermined in the past. The third and fourth constraints, (10) and (11), are the resource constraint of the final goods and the labor market clearing condition, respectively. The fifth constraint (12) defines the population size that governs the agglomeration/congestion forces.

While the original problem is quite complex because of infinite dimensionality of the state variable and the control variables, the problem dramatically simplifies by considering the Lagrangian of the above problem. We detail the derivations in Appendix A.1 and focus on the key results in the main text. Let P_{it}, w_{it}, and $w_{it}\alpha_{it}$ denote the Lagrangian multipliers of (10), (11), and (12), respectively. Let $\mathcal{S}_t(\phi)$ denote the associated Lagrangian. We then guess and verify that the Lagrangian $\mathcal{S}_t(\phi)$ is additively separable in (v, i):[4]

$$\mathcal{S}_t(\phi) = \sum_i \int S_{it}(v) d\phi_i + D_t, \tag{14}$$

where D_t is the term that is independent of ϕ. The constant term D_t solves

$$D_t = \max_{\{l_{ij}, L_i\}} \sum_i P_{it} f_{it}(\{l_{ki}\}_k; \{L_k\}_k) - \sum_i w_{it} \sum_j l_{ij} - \sum_i \alpha_{it} w_{it} L_i + \frac{1}{R} D_{t+1}. \tag{15}$$

The first-order optimality conditions are given by the following static conditions:

$$P_{it} \frac{\partial f_{it}}{\partial l_{ki}} = w_{kt} \tag{16}$$

$$\sum_i P_{it} \frac{\partial f_{it}}{\partial L_k} = w_{kt} \alpha_{kt}. \tag{17}$$

As one can imagine from the above expressions, and as we later show, once we consider decentralization of the planner's solution, the Lagrangian multiplier P_{it} would correspond to the price of final consumption goods in i, w_{it} corresponds to the wage in i, and α_{it} corresponds to the agglomeration elasticity in location i.

The dynamics comes from the following component planning problem, a problem for the household in location i with promised utility v:

$$\begin{aligned} S_{it}(v) = \max_{\{C_{it}, v'_{ij}(\epsilon), \ell_i(\epsilon)\}} & w_{it}(1 + \alpha_{it}) - P_{it} C_{it} + (1 - \omega) S^n_{it} \\ & + \frac{1}{R} \omega \mathbb{E}_{it} \sum_j \mathbb{I}[\ell_i(\epsilon) = j] S_{jt+1}\big(v'_{ij}(\epsilon)\big) \end{aligned} \tag{18}$$

subject to the promise keeping constraint:

$$v = u_i(C_{it}) + \beta \omega \mathbb{E}_{it} \Big[\sum_j \mathbb{I}[\ell_i(\epsilon) = j] \left\{ v'_{ij}(\epsilon) + \epsilon_{i,j} \right\} \Big]. \tag{19}$$

[4] Similar techniques appear in Atkeson and Lucas (1992), Farhi and Werning (2007, 2012), and Veracierto (2022, 2023) in the context of various different models.

The term S_{it}^n is the value function associated with newborns in location i, which solves

$$S_{it}^n = \max_{v_i^n, \{v_{ij}^{n\prime}(\epsilon), \ell_i^n(\epsilon)\}} \Lambda_i v_i^n + \frac{1}{R}\mathbb{E}_{it} \sum_j \mathbb{I}[\ell_i^n(\epsilon) = j] S_{jt+1}(v_{ij}^{n\prime}(\epsilon)) \tag{20}$$

$$\text{s.t.} \qquad v_i^n = \beta \mathbb{E}_{it} \sum_j \mathbb{I}[\ell_i^n(\epsilon) = j] \left\{ v_{ij}^{n\prime}(\epsilon) + \epsilon_{i,j} \right\}. \tag{21}$$

We let $\{C_{it}(v), v_{ij}^{\prime}(v,\epsilon), \ell_i(v,\epsilon)\}$ denote the policy functions associated with (18), and let v_i^n, $\{v_{ij}^{n\prime}(\epsilon), \ell_i^n(\epsilon)\}$ denote the policy functions associated with (20).

The objective function $S_{it}(v)$ has a clear economic interpretation as the net surplus associated with households in i with promised utility v (net of household utility). At each point in time, such a household adds the marginal product of labor, w_{it}, and the agglomeration benefit, $w_{it}\alpha_{it}$, while subtracting the cost of resource consumed by the household, $P_{it}C_{it}$. We can also interpret them as fiscal and technological externalities that each household generates in the equilibrium that decentralizes the planner's solution. A household may also die and drop newborns, which corresponds to the term $(1-\omega)S_{it}^n$. The continuation value takes into account that the household may move to the other locations.

Taking the first-order conditions and combining it with the envelope condition, we have the following characterization of the first-best allocation with complete information.

Proposition 1. *Under complete information, the following conditions must hold at the planner's solution. For each household, the consumption over time and across space satisfies*

$$\frac{u'(C_{it})}{P_{it}} = \beta R \frac{u'_{jt+1}(C_{jt+1})}{P_{jt+1}} \tag{22}$$

for all i, j, t. The migration decision solves

$$\ell_{it}(v,\epsilon) \in \arg\max_l \; \beta\omega \frac{P_{it}}{u'(C_{it})}[v_{ilt+1}(v,\epsilon) + \epsilon_{il}] + \frac{1}{R}\omega S_{lt+1}(v_{il}(v,\epsilon)). \tag{23}$$

Proposition 1 reveals a relatively intuitive property that the efficient complete information allocation must feature. Condition (22) says that the marginal utility of income must be equalized across space and over time after adjusting for discounting. If the marginal utilities were not equalized, the planner can improve welfare by reallocating resources from location or time periods with low marginal utility to those with high marginal utility. Note that condition (22) also shows that consumption, C_{jt+1}, is independent of the realization of the preference shock, conditional on the location of living, which comes from the additive separability of the preference shock. This implies that the continuation value, $v_{ijt+1}(v,\epsilon)$, is independent from ϵ as well, $v_{ijt+1}(v,\epsilon) = v_{ijt+1}(v)$. The second condition (23) says that the location choice must maximize the total surplus, a sum of worker utility and the net surplus.

Finally, we note that Proposition 1 is a dynamic analogue of the efficiency conditions in a static discrete choice models, presented in Mongey and Waugh (2024) and Donald, Fukui, and Miyauchi (2024). Relative to the static efficiency condition, there are two differences. First, the planner equalizes marginal utility of income not only across space but also over time. Second, the location choice is governed by a dynamic discrete choice system, rather than the static one.

3.2 Location Preference Shock as Private Information

The analysis in the previous section assumes that the planner observes the location preference shock. Although it serves as a theoretical benchmark, such a situation is hardly realistic. We now move to the case where the preference shocks are private information of the households. By revelation principle, we can focus on the direct revelation mechanism where households report their preference shock in each period. Due to the independence of preference shocks over time, we do not need to carry past realization or report of preference shocks as state variables.

It is instructive to first consider why households have an incentive to misreport their preferences in the complete information allocation. Since households only care about their utility, if households currently living in i could choose the location of living in the next period, they would solve

$$\max_{l} \beta\omega[v_{ilt+1}(v) + \epsilon_{il}], \tag{24}$$

instead of (23). Here we already imposed that the continuation value v_{ilt+1} does not depend on ϵ as explained earlier. This implies that households have an incentive to report their preferences so as to be assigned to location that solves (24) instead of (23). Intuitively speaking, from an individual household's perspective, they do not consider the fiscal surplus or the agglomeration/congestion forces that they generate in choosing where to live. In contrast, these forces are what the planner takes into account, and hence the misalignment arises.

We therefore need to impose the following incentive compatibility constraint such that the truth-telling is optimal from the household's perspective:

$$\epsilon \in \arg\max_{\hat{\epsilon}} \ \beta\omega \sum_{j} \mathbb{I}[\ell_i(v, \hat{\epsilon}) = j] \left\{ v'_{ij}(v, \hat{\epsilon}) + \epsilon_{ij} \right\}. \tag{25}$$

From the above expression, it is tempting to think that the planner should promise a low value of $v'_{ij}(v, \hat{\epsilon})$ to households with a high value of $\hat{\epsilon}_{ij}$, since such a household would anyway choose location j so the planner does not need to give a strong incentive. However, equation (25) tells us that such a policy is not incentive compatible. In fact, $v'_{ij}(v, \hat{\epsilon})$ cannot depend on the reported preference shock, $\hat{\epsilon}$, conditional on living location j. If $v'_{ij}(v, \hat{\epsilon})$ were a function of the reported preference shock, households would misreport the preference that gives the highest value of $v'_{ij}(v, \hat{\epsilon})$.

The above discussion implies that any incentive-compatible allocation must feature

$$v'_{ij}(v, \hat{\epsilon}) = v'_{ij}(v). \tag{26}$$

This says that conditional on location of living and the promised value, the planner cannot discriminate households. Such a property should be intuitive. If two households choose to live in Boston, it is impossible to tell from their behavior whether one likes to live in Boston more than the other.

This result also implies that the optimal policy is necessarily a placed-based policy, even though we have not restricted it to be. We later formally show that the constrained efficient allocation can be implemented by policies that depend on the history of living locations.

Based on the above observation, we can rewrite the incentive compatibility constraint as households directly choose the location of living to maximize their utility, instead of reporting preferences and receive assignment:

$$\ell_i(v, \epsilon) \in \arg\max_l \; \beta\omega \left[v'_{il}(v) + \epsilon_{il}\right] \tag{27}$$

Imposing this incentive compatibility constraint in the problem (18), we have the following component planning problem:

$$\begin{aligned} S_{it}(v) = \max_{\{C_{it}, v'_{ij}, \ell_i(\epsilon)\}} \; & w_{it}\left(1+\alpha_{it}\right) - P_{it}C_{it} + (1-\omega)S^n_{it} \\ & + \frac{1}{R}\omega\mathbb{E}_{it}\left[\sum_j \mathbb{I}[\ell_i(\epsilon) = j] S_{jt+1}\left(v'_{ij}\right)\right] \end{aligned} \tag{28}$$

$$\text{s.t.} \qquad v = u_i(C_{it}) + \beta\omega\mathbb{E}_{it}\left[\sum_j \mathbb{I}[\ell_i(\epsilon) = j]\left\{v'_{ij} + \epsilon_{i,j}\right\}\right] \tag{29}$$

$$\ell_i(\epsilon) \in \arg\max_l \; \beta\omega\left[v'_{il} + \epsilon_{il}\right]. \tag{30}$$

The only difference from the complete information case is the presence of the incentive compatibility constraint (30).

We can further simplify the above problem using the representation result from Hofbauer and Sandholm (2002). They show that any discrete choice problem can be equivalently represented as the maximization problem with respect to choice probability subject to appropriately defined cost function. Formally, we can rewrite the above problem as

$$S_{it}(v) = \max_{\{C_{it}, v'_{ij}, \mu_{ij}\}} w_{it}(1+\alpha_{it}) - P_{it}C_{it} + (1-\omega)S^n_{it} + \frac{1}{R}\omega \sum_j \mu_{ij} S_{jt+1}(v'_{ij}) \tag{31}$$

$$\text{s.t.} \qquad v = u_i(C_{it}) + \beta\omega \left[\sum_j \mu_{ij} v'_{ij} - \psi_{it}(\{\mu_{ij}\}_j)\right]. \tag{32}$$

$$\{\mu_{ij}\}_j \in \arg\max_{\{\tilde{\mu}_{ij}\}_j} \beta\omega \left[\sum_j \tilde{\mu}_{ij} v'_{ij} - \psi_{it}(\{\tilde{\mu}_{ij}\}_j)\right] \tag{33}$$

for some function ψ_{it} that only depends on the distribution function G_{it}[5] and $\mu_{ijt} \equiv \mathbb{E}_{it}\mathbb{I}[\ell_{it}(\epsilon) = j]$. Let $\{C_{it}(v), v'_{ij}(v), \mu_{ij}(v)\}$ denote the policy functions associated with the above Bellman equation.[6] Likewise, the net surplus of newborns, S^n_{it}, is given by

$$S^n_{it} = \max_{v^n_i, \{v^n_{ij}, \mu^n_{ij}\}} \Lambda_i v^n_i + \frac{1}{R}\sum_j \mu^n_{ij} S_{jt+1}(v^n_{ij}) \tag{34}$$

$$\text{s.t.} \qquad v^n_i = \beta \sum_j \left[\mu_{ij} v^{n\prime}_{ij} - \psi_{it}(\{\mu^n_{ij}\}_j)\right] \tag{35}$$

$$\{\mu^n_{ij}\}_j \in \arg\max_{\{\tilde{\mu}^n_{ij}\}} \beta \sum_j \left[\tilde{\mu}^n_{ij} v^{n\prime}_{ij} - \psi_{it}(\{\tilde{\mu}^n_{ij}\}_j)\right], \tag{36}$$

where $\mu^n_{ij} \equiv \mathbb{E}_{it}\mathbb{I}[\ell_{it}(\epsilon) = j]$.

At this point, the problem has a similar structure to the optimal design of dynamic unemployment insurance by Hopenhayn and Nicolini (1997) and Veracierto (2022). There, the planner seeks to equalize the marginal utility of employed and unemployed over time, taking into account that doing so discourages the job search effort. Here, the planner seeks to equalize marginal utility across space and over time taking into account that doing so leads to inefficient spatial population distribution. Aside from the difference in context, our problem differs in that we derived the problem from the dynamic discrete choice problem and that our problem accommodates a general choice set as well as a general production structure with technological externalities. The above problem is also closely related to optimal wage tenure contracts between a firm and a worker, as in Burdett and Coles (2003), Shi (2009), Balke and Lamadon (2022), and Souchier (2022), where the firms seek to insure workers against productivity shocks while trying to retain workers. In fact, we show that our formula that characterize the constrained efficient allocation resembles

[5]See Hofbauer and Sandholm (2002) or Donald et al. (2024) for an explicit expression for ψ.

[6]In general, $S_{it}(v)$ may not be concave in v, in which case, the lottery is needed to ensure concavity in the value function (see e.g., Prescott and Townsend 1984, Balke and Lamadon 2022), For the sake of notational simplicity, we abstract from the use of lottery. In the quantitative exercise, we verify that $S_{it}(v)$ is concave, and therefore the lottery is not used even if available.

their formula below.

Solving the above problem gives the following condition that the constrained efficient allocation must satisfy.

Proposition 2. *In any constrained efficient allocation, the following equation must hold for each household living location i at time t:*

$$\mu_{ijt}\left[\beta R\frac{u_j'(C_{jt+1})/P_{jt+1}}{u_i'(C_{it})/P_{it}}-1\right]+\underbrace{\sum_k\frac{\partial\mu_{ik}}{\partial C_{jt+1}}\frac{S_{kt+1}(v_{ik}')}{P_{jt+1}}}_{\equiv\xi_{ijt}}=0. \tag{37}$$

for all i, j, t.

The above proposition characterizes the trade-off between equalization of marginal utility of income and distorting migration decisions. Without the term ξ_{ijt}, the above condition is identical to (22) in the complete information benchmark. The term ξ_{ijt} encompasses the additional consideration that the planner must navigate in equalizing the marginal utility. In particular, an increase in consumption at location j at time $t+1$, C_{jt+1}, for households living in location i at time t induces migration responses, which is associated with a change in net surplus. Note that migration responds not only at location j, but also at all locations $k=1,\ldots,J$.

First, consider the case with $\xi_{ijt}>0$. This means that increasing consumption at location j at time $t+1$ induces migration responses that are net positive in terms of net surplus. In this case, the planner has an incentive to increase the relative consumption at location j at time $t+1$ to consumption at i at t beyond the complete information benchmark. This is the case where the planner back-loads the consumption profile to incentivize households to migrate to good locations. The opposite is true in the case of $\xi_{ijt}<0$. In this case, an increase in consumption at location j at time $t+1$ induces migration responses that are net negative. In this case, the planner lowers the relative consumption at location j at time $t+1$ to consumption at i at time t relative to the complete information benchmark. This is the case where the planner front-load consumption in order to discourage migration toward bad locations.

As mentioned earlier, Proposition 2 has a close connection with the formulas in the context of optimal wage tenure contracts derived in Burdett and Coles (2003), Balke and Lamadon (2022), and Souchier (2022). There, a similar trade-off arises between smoothing consumption and retaining workers. Proposition 2 is also a dynamic analogue of the static optimal spatial transfer formula derived in Donald, Fukui, and Miyauchi (2024).

Proposition 2 implies the inverse Euler equation.

Corollary 1. *Any constrained efficient allocation must satisfy the inverse Euler equation:*

$$\frac{P_{it}}{u_i'(C_{it})} = \mathbb{E}_{it}\left[\frac{P_{jt+1}}{\beta R u_j'(C_{jt+1})}\right] \tag{38}$$

As observed by Diamond and Mirrlees (1978) and Rogerson (1985), constrained efficient allocation in an economy with private information generally satisfies the inverse Euler equation. Not surprisingly, our environment is not an exception.

The rest of the allocations and the Lagrangian multipliers are determined in the same way as the complete information case. In particular, $\{l_{ijt}, L_{it}, \phi_i', P_{it}, w_{it}, \alpha_{it}\}$ solve (10), (11), (12), (13), (16), and (17).

Decentralization. Now we briefly discuss how the constrained efficient allocation can be decentralized. There are many ways to implement the constrained efficient allocation, but the key is that the policies to be contingent (only) on the history of living locations. For example, suppose households are hand-to-mouth, or the government bans the private savings. In this case, the transfers contingent on the history of living locations can implement the constrained efficient allocation, as we formally show in Appendix A.3. A tax system that has more "decentralized" flavor involves non-linear capital income taxation contingent on the history of living locations, as discussed in Kocherlakota (2005) in the context of dynamic Mirleese environment.

3.3 History-Independent Constrained Efficient Allocation

As discussed earlier, the implementation of constrained efficient allocation requires policy instruments to condition on the past history of living locations. We now study a simpler and more restricted policies that can only depend on the current location of living. This exercise serves two purposes. First, comparing this history independent policy to the history dependent policy sheds light on the importance of dynamic incentives we emphasized earlier. Second, while a policy contingent on a history of living locations should be feasible in theory, it might be difficult to implement in practice. Therefore, a simple history-independent policy might be a its own interest from a pratical perspective.

We argue that the underlying principles in Proposition 2 carry over to more restricted policies. To make this point, we place the restriction that consumption in each location must be the same for all households residing in the same location. This is trivially satisfied in many of the existing dynamic spatial equilibrium literature that assumes hand-to-mouth households (Artuç, Chaudhuri, and McLaren 2010, Caliendo, Dvorkin, and Parro 2019). Consequently, the set of implementable allocations coincide with the allocation that can be implementable with transfers that only depend on the current locations in those models.

The restriction implies that all households currently living in the same location attain same utility going forward in expectation. Imposing these requirements in the original problem, we obtain the formula that any history independent optimal policy must satisfy.

Proposition 3. *Any constrained efficient and history independent allocation that must satisfy*

$$\sum_i L_{it}\left[\mu_{ij}\left[\beta R\omega\frac{u_j'(C_{jt+1})/P_{jt+1}}{u_i'(C_{it})/P_{it}}+\beta R(1-\omega)\Lambda_i\frac{u_j'(C_{jt+1})}{P_{jt+1}}-1\right]+\sum_k\frac{\partial\mu_{ikt}}{\partial C_{j+1}}\frac{S_{kt+1}}{P_{jt+1}}\right]=0, \tag{39}$$

where

$$S_{jt}=w_{jt}\left(1+\alpha_{jt}\right)-P_{jt}C_{jt}+(1-\omega)\Lambda_j v_j^n+\frac{1}{R}\sum_k\mu_{jkt}S_{kt+1} \tag{40}$$

An important observation here is that the formula in Proposition 3 is isomorphic to that in Proposition 2 except that now that it is a weighted average across migration origins and generations. Since the planner cannot discriminate households from different origin locations and across cohorts, the planner weighs insurance-incentive trade-offs of all these different households altogether.

An important conceptual difference from Proposition 2 is that the planner is no longer able to back-load or front-load the consumption profile for each household. If the planner would like to invite more households to a certain location, then the planner must increase the consumption in that location regardless of the previous consumption level. We highlight this difference in the quantitative exercise later.

Note that when households are infinitely lived, $\omega=1$, the above formula simplifies to

$$\sum_i L_{it}\left[\mu_{ij}\left[\beta R\frac{u'(C_{jt+1})/P_{jt+1}}{u'(C_{it})/P_{it}}-1\right]+\sum_k\frac{\partial\mu_{ikt}}{\partial C_{j+1}}\frac{S_{kt+1}}{P_{jt+1}}\right]=0, \tag{41}$$

and

$$S_{jt}=w_{jt}\left(1+\alpha_{jt}\right)-P_{jt}C_{jt}+\frac{1}{R}\sum_k\mu_{jkt}S_{kt+1}. \tag{42}$$

With the lack of history dependence, the stationary distribution in this environment exists even with infinitely lived households, avoiding the immiseration result by Atkeson and Lucas (1992). Formulae (41) and (42) might be on its own interest if one wishes to study the optimal dynamic spatial policy in an environment where households are hand-to-mouth (e.g., Caliendo et al. 2019) and the government can only condition on the current location of living.[7]

3.4 Extensions

Our baseline model deliberately abstracts from several considerations in order to transparently convey the main forces. We now discuss extensions and generalizations of the baseline model.

3.4.1 Capital Accumulation

Some existing work (e.g., Kleinman, Liu, and Redding 2023, D'Amico and Alekseev 2024) introduce capital accumulation into the dynamic spatial equilibrium model, which we have abstracted from so far. In Appendix A.4, we show that our environment can be straightforwardly extended to incorporate location-specific capital accumulation subject to capital adjustment costs. Importantly, such consideration does not meaningfully interact with the trade-off that we highlighted in Proposition 2. In fact, Proposition 2 and the underlying Bellman equations remain unchanged. Meanwhile, optimal investment and capital accumulation follow the standard q-theory of investment.

3.4.2 Ex-ante Heterogeneous Households

In the baseline model, we have assumed that all households are ex-ante homogeneous. In Appendix A.5, we extend our baseline environment to an environment with many ex-ante heterogeneous household types $\theta \in \{\theta_1, \ldots, \theta_M\}$ with arbitrary heterogeneity in preferences, location choice, and demographics. We also consider a general form of agglomeration/congestion forces that allow for spillover across different household types. There, we show that Proposition 2 and the underlying Bellman equations remain unchanged, except that now everything is indexed by θ.

3.4.3 Lagged Agglomeration/Congestion Forces

In the baseline model, we have assumed that the agglomeration and congestion forces arise from contemporaneous population size. Some existing work like Allen and Donaldson (2020) allows the agglomeration/congestion forces to depend on the lagged population size. In Appendix A.6, we extend our environment to allow agglomeration/congestion forces to depend on arbitrarily long lags of population size distribution. The only material difference from our baseline model is the term α_{it} in (31) now includes the agglomeration/congestion forces not only today but also of future dates.

[7] This was the focus of the earlier version of this paper.

4 Quantification

In this section, we calibrate our model assuming the status quo is generated by an exogenously incomplete market in the style of Bewley-Hugget-Aiyagari. We then solve the constrained efficient allocation to quantify its deviation from the status quo economy.

4.1 Status Quo Economy with Exogenously Incomplete Market

We assume an exogenously incomplete financial market in the spirit of Bewley-Hugget-Aigayari in the status quo economy (Bewley 1986, Huggett 1993, Aiyagari 1994, Imrohoroğlu 1989). The only available asset in the economy is the state non-contingent bonds in zero net supply. Let $1 + r_t$ be the real interest between time t and $t + 1$. All households face a common exogenous borrowing limit and the minimum asset level is given by $\underline{a}$. We also introduce spatial transfers implemented by the government. Households living in location j at time t receive transfers T_{jt}.

The Bellman equation of the household living in location j at time t with bond holding a_t is

$$v_{jt}(a_t) = \max_{C_{jt},\{\mu_{jk}\}_k,a_{t+1}} u_{jt}(C_{jt}) + \beta\omega\left[\sum_k \mu_{jk}v_{kt+1}(a_{t+1}) - \psi(\{\mu_{jk}\})\right] \tag{43}$$

$$\text{s.t.} \quad P_{jt}C_{jt} + a_{t+1} = (1 + r_{t-1})a_t + w_{jt} + T_{jt} \tag{44}$$

$$a_{t+1} \geq \underline{a}, \tag{45}$$

where P_{jt} is the price index, w_{jt} is the wage, and $\underline{a}$ is the exogenous minimum asset level. Here, we have already imposed the representation result by Hofbauer and Sandholm (2002) and Donald et al. (2024) to write the dynamic discrete choice problem. Let $C_{jt}(a)$, $a_{jt+1}(a)$, and $\mu_{jk}(a)$ denote the policy functions associated with the above value functions. We assume that when households die, they leave accidental bequests for their offspring. Note that our model nests commonly used hand-to-mouth households as a special case with $\underline{a} = 0$.

There is a representative firm in each location that imports factor services from other regions and produces non-traded financial goods. The representative firm in location j solves

$$\max_{\{l_{kjt}\}_k} P_{jt}f_{jt}(\{l_{kjt}\}, \{L_{kt}\}_k) - \sum_k w_{kt}l_{kj}, \tag{46}$$

taking $\{L_{kt}\}_k$ as given. Therefore, agglomeration/congestion forces are externalities that are not internalized by private agents.

We assume that the government runs a balanced budget. The government budget constraint is

$$\sum_k T_{kt} = 0. \tag{47}$$

Let $\varphi_j(\mathbb{A})$ denote the measure of households with bond $a \in \mathbb{A}$ in location j. The goods market clearing condition is

$$\int C_{jt}(a) d\varphi_j = f_{jt}(\{l_{kjt}\}_k, \{L_{kt}\}_k). \tag{48}$$

The factor market clearing condition is

$$\sum_j l_{kjt} = \int d\phi_k. \tag{49}$$

The consistency of population size requires

$$L_{kt} = \int d\phi_k. \tag{50}$$

The distribution evolves according to the following law of motion:

$$\varphi_{jt+1}(\mathbb{A}) = \sum_k \mu_{kjt}(a_{jt+1}^{-1}(\mathbb{A}))\varphi_{kt}(a_{jt+1}^{-1}(\mathbb{A})) \tag{51}$$

The decentralized equilibrium of the status quo economy consists of value and policy functions $\{v_{jt}(a), C_{jt}(a), a_{jt+1}(a), \mu_{jk}(a)\}$, factor contents of trade, $\{l_{kjt}\}$, population distribution, $\{L_{kt}\}_k$, spatial transfers, $\{T_{kt}\}$, distribution over assets in each location, $\{\varphi_{jt}\}$, and prices $\{w_{kt}, P_{kt}, r_t\}$ such that: (i) given prices $\{w_{kt}, P_{kt}, r_t\}$ and policy $\{T_{kt}\}$, the value and the policy functions $\{v_{jt}(a), C_{jt}(a), a_{jt+1}(a), \mu_{jk}(a)\}$ solve the households problem (43); (ii) given prices $\{w_{kt}, P_{kt}\}$ and population size $\{L_{kt}\}$, the factor contents of trade $\{l_{kjt}\}$ solve the firm's problem (46); (iii) the government sets the transfer that satisifes (47); (iv) markets clear (48), (49); (v) the population size is consistent (50); and (vi) the distribution $\{\varphi_{jt}\}$ evolves according to (51).

We note that there is no reason to expect that the status quo economy achieves first-best nor constrained efficient allocation for two reasons. First, private agents do not internalize agglomeration or congestion in making agglomeration decisions. Second, the market is incomplete in the sense that there is no market that insures agents against the uncertainty in location of living or preference shocks. In what follows, we quantify these deviations through the lens of the calibrated model.

4.2 Calibration

Table 1 summarizes our calibration. We calibrate our status-quo economy to match the data on the US 2017, assuming that the US is in its steady state in 2017. One period is a year. We consider 48 states in the US as a geographical unit, which excludes Alaska and Hawaii from 50 states. We choose Alabama's labor as numeraire and set its wage to one. Since our calibration assumes steady state, we drop the subscript t whenever there is no risk of confusion.

We first parameterize the utility function as a CRRA utility function,

$$u_j(C_{jt}) = \frac{C_{jt}^{1-\gamma} - 1}{1-\gamma}, \tag{52}$$

and we set $\gamma = 1$, corresponding to log utility, a standard specification in the literature (e.g., Caliendo et al. 2019). We also set $\omega = 1 - 1/75$ so that the average life expectancy is 75 years. The production function is assumed to take the constant elasticity of substitution form:

$$f_j(\{l_{kj}\}, \{L_k\}_k) = \left[\sum_k (\mathcal{A}_{kj}(L_k)\, l_{kj})^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{53}$$

where $\sigma > 1$ corresponds to the trade elasticity, and $\mathcal{A}_{kj}(L_k)$ is the productivity shifter of goods shipped from location k to j that depends on the population size of location k. The productivity is an iso-elastic function of population size of the origin location:

$$\mathcal{A}_{kj}(L_k) = A_{kj} L_k^{\alpha}. \tag{54}$$

This specification, together with (16) and (17), implies

$$\alpha_{jt} = \alpha \quad \text{for all } j, t. \tag{55}$$

For the baseline model, we assume $\alpha = 0.04$. This value corresponds to an intermediate estimate in the literature summarized in Rosenthal and Strange (2004) and Combes and Gobillon (2015). We set the trade elasticity to $\sigma = 5$, as in Costinot and Rodríguez-Clare (2014).

We assume that the migration cost function takes the following form:

$$\psi_i(\mu_{ijt}) = \frac{1}{\theta} \sum_j \mu_{ij} \ln(\mu_{ij}/\chi_{ij}), \tag{56}$$

which would imply logit dynamic discrete choice system, as in Artuç et al. (2010) and Caliendo, Parro, Rossi-Hansberg, and Sarte (2018). This is equivalent to assuming $\{\epsilon_j\}_j$ follows an indepen-

Table 1: Parameter Values

Parameter	Description	Value	Source/Target
A. Assigned Parameters			
γ	Risk aversion	1	Standard
σ	Trade elasticity	5	Costinot and Rodríguez-Clare (2014)
θ	Migration elasticity	0.5	Caliendo et al. (2019)
ω	Surviving probability	0.987	Life expectancy 75 years
α	Agglomeration elasticity	0.04	Baseline
B. Internally Calibrated Parameters			
β	Private discount factor	0.988	Real interest rate 2%
$\underline{a}$	Borrowing limit	-0.60	MPC 0.3
$\{A_{ij}\}$	Productivity shifter	-	Trade flows and real output
$\{\chi_{ij}\}$	Migration cost shifter	-	Migration flows
$\{\varkappa_i\}$	Net transfer rate	-	Net transfer from the government
C. Parameters for Social Welfare Function			
$1/R$	Social discounting	0.988	Private discount factor
$\{\Lambda_j\}$	Location welfare weights	-	Equal weight

Note: The table shows the parameters used in our quantitative exercise. Parameter values for $\{A_{ij}\}$, $\{\chi_{ij}\}$, and $\{\varkappa_i\}$ are chosen to exactly match the data moments described in the main text.

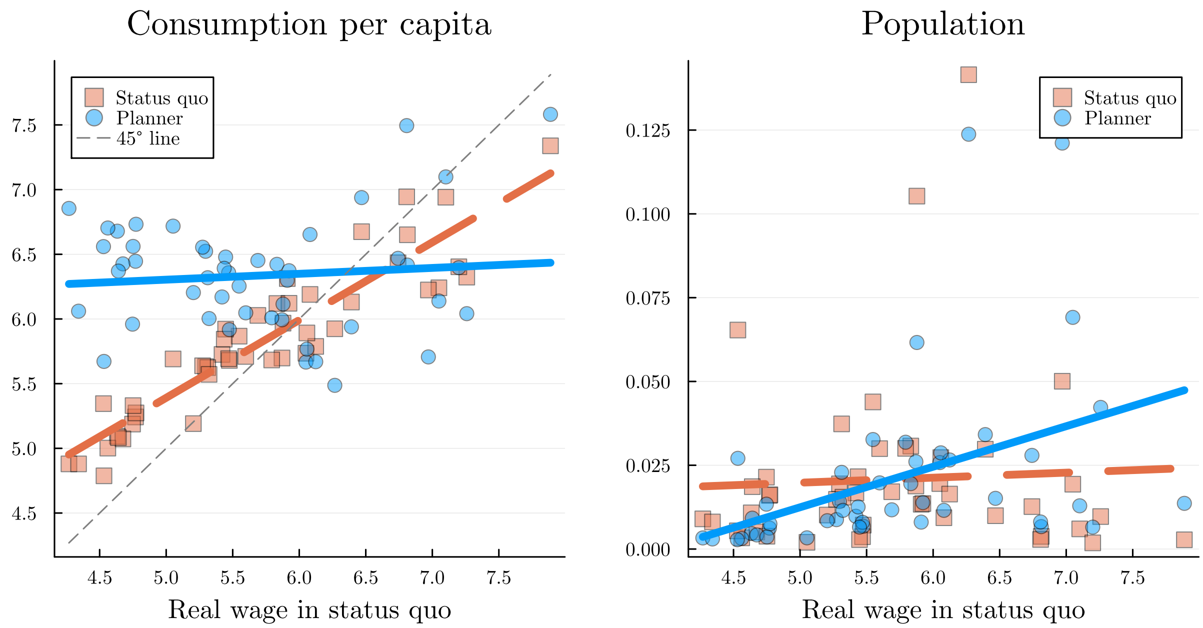
dent Type-I extreme value distribution, as originally shown by Anderson, De Palma, and Thisse (1988). The parameter θ governs the migration elasticity, and χ_{ij} represents the bilateral migration cost shifter. We set the value of migration elasticity to $\theta = 0.5$, in line with the estimates in Caliendo et al. (2019).

We briefly describe the calibration of other parameter values and relegate the details to Appendix B.1. We choose $\{A_{ij}\}_{i\neq j}$ and $\{\chi_{ij}\}_{i\neq j}$ to match the bilateral trade and migration flows at the state level. We choose $\{A_{ii}\}$ to match the real wage level in each state, and we normalize $\chi_{ii} = 1$ for all i. We set the discount factor β so that the real interest rate is 2%. We choose the value of $\underline{a}$ so that the average annual marginal propensity to consume (MPC) is 0.3, which corresponds to the estimate by Orchard, Ramey, and Wieland (2023).[8] Finally, we parameterize $T_j = \varkappa_j w_j + \bar{T}$ and choose $\{\varkappa_j\}$ to match the net transfer from the government to income ratio at the state level. We obtain net transfer from the government from the Bureau of Economic Analysis (BEA). We adjust the term $\bar{T}$ to ensure the government budget constraint holds.

Finally, we need to take a stance on the parameters that govern the social welfare function, $\{R, \Lambda_j\}$. We choose the social discounting $1/R$ to be the same as the private discount factor β. The planner puts equal weight on households born in different regions, which we normalize to

[8] This is at the lower end of the MPC estimates in the literature. We view this choice as conservative since higher MPC calibration would give less of a chance of risk sharing in the status quo economy.

Figure 1: Steady State Consumption and Population: Status Quo vs. Planner



Note: The left panel plots the average consumption per capita in each state against the real wage in the status quo economy. The square dot correspond to the status quo economy, and the circle dot corresponds to the planner's solution. The dashed red line is the best linear fit for the status quo economy. The solid blue line is the best linear fit for the planner's solution. The right panel plots population size against the real wage in the status quoeconomy and is analogous to the left panel.

one, $\Lambda_j = 1$ for all j.

4.3 Steady State: Status Quo vs. Planner

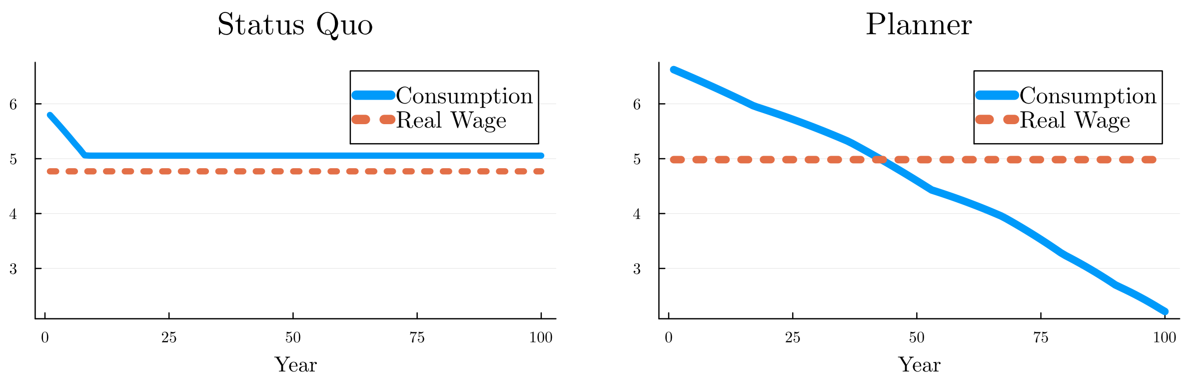
Armed with our calibrated parameters, we solve the constrained efficient allocation and compare to the status quo economy. Our first goal is to understand the steady state properties.

The left panel of Figure 3 compares the average consumption per capita. We plot them against the real wage in the status quo economy, which proxy how "good" the location is.[9] The square dots plot the average consumption in each state against the real wage in the status quo economy. If the economy were in financial autarky, they would lie exactly on the 45 degree line. The availability of state non-contingent bonds and government transfers in the status quo economy helps smooth consumption, and consequently the slope of the relationship is slightly below one. The circle dots show the consumption per capita in the planner's solution, and the relationship is nearly flat. This suggests that the constrained efficient allocation features substantially more spatial equality.

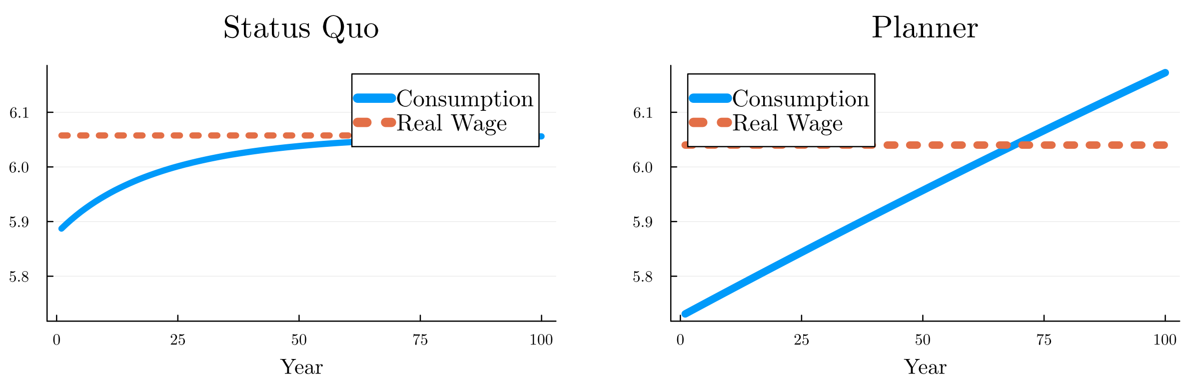
[9]In fact, Appendix Figure C.2 that the real wage in the status quo economy is strongly positively related to the average net surplus S_j. Appendix Figure C.1 shows that the real wage in the planner's solution are tightly related to that in the status quo economy.

Figure 2: Consumption and Real Wage over the Life-Cycle for Stayers

(a) Alabama



(b) Washington



Note: The figure plots the consumption and real income profile of households born in Alabama (panel (a)) and in Washington (panel (b)). We focus on the households who keep staying in the same location. The left panel shows the status quo economy, and the right panel shows the planner's solution. In both cases, the initial condition is the average of the households born in each location.

The right panel of Figure 3 compares population. The slope of the planner's solution is higher than the status quo economy. Therefore, perhaps surprisingly, we see more population reallocation toward states with higher real wage in the planner's solution, despite their average consumption being lower than the status quo economy. In a static economy, there is generally a trade-off between equalizing consumption and distorting reallocation of population to the productive location (Donald, Fukui, and Miyauchi 2024). Here, it appears that there is no such trade-off.

The key here is that the dynamic incentives goes a long way in overcoming the tradeoff between spatial inequality and distorting migration decisions. The planner not only controls the average consumption level in each state, but also the entire path of consumption conditional on the history of living locations. The planner can incentivize households to stay in good locations by back-loading the consumption profile, even with a low average consumption. Likewise, the planner can incentivize households to leave bad locations by front-loading the consumption profile.

	Welfare (C equivalent)
Status quo	1.000
Planner (second-best)	1.042
Complete information (first-best)	1.060
History independent	1.029

Table 2: Steady State Welfare of Newborns

Note: The table shows the utilitarian welfare of newborns. They are measured in consumption equivalent units, $C^{eq} = u^{-1}((1-\beta\omega)W)$, where W is the utilitalian welfare. We normalize by the welfare of status quo economy to one.

Figure 2 demonstrates this by plotting life-cycle consumption and income profile in Alabama ("bad" location) and in Washington ("good" location). We consider households born in these locations and end up staying there for their lifetime. In the status quo economy, households born in Alabama borrow to consume more than the real wage. This immediately drives down the asset and eventually hits the borrowing constraint. After hitting the borrowing constraint, consumption remains flat. In contrast, in the planner's solution, the consumption profile is front loaded throughout the lifetime with initially higher level of consumption than the status quo economy. In this way, the planner can effectively insure households born in Alabama while strongly incentivizing them to leave Alabama. Washington example illustrates the polar opposite case. In the status quo economy, as households accumulate savings, consumption increases over the lifecycle but at a moderate pace. In the planner's solution, the consumption initially starts from a low level but is steeply backloaded. This strongly incentivizes households to stay in Washington while sharing the risk of where to be born.

What are the welfare gains from moving to the constrained efficient allocation? The second row of Table 2 shows that the utilitarian welfare of newborns is 4.2% higher in consumption equivalent units. For comparison, the third row of Table 2 shows that moving to the first-best allocation with complete information, which we characterized in Proposition 1, achieves 6.03% of welfare gains. Therefore, even with private information, the constrained efficient allocation achieves the bulk ($70\% = 4.2/6.0$) of welfare gains that are possible under complete information.

4.4 History-Independent Constrained Efficient Allocation

We have highlighted the importance of dynamic incentives in simultaneously achieving consumption smoothing and efficient migration decisions. This critically relies on the planner's ability to set the consumption profile as a function of the history of living locations.

We now contrast the history contingent allocation with a simple history-independent allocation that we characterized in Proposition 3. With history independent allocation, the planner is no longer able to use dynamic incentives, as the planner has to give the same consumption to two households currently living in the same location but differ in their history of living locations. For this reason, the comparison sheds light on the importance of dynamic incentives. As before, we focus on the steady-state allocations.

Figure 3 shows the per-capita consumption per capita and population that is analogous to Figure 3. The left panel shows that the planner creates substantial spatial consumption inequality with a higher slope than one (45-degree line). Despite inducing substantial spatial inequality, population is less concentrated in the productive locations than the constrained efficient allocation. These two results highlight the importance of dynamic incentives in overcoming the trade-off. Without dynamic incentives, the planner faces a strong trade-off in spatial inequality and efficient migration decisions.

The third row of Table 3 shows that the welfare gains from history-independent policy is 2.9% relative to the status-quo economy, measured in consumption equivalent units. Therefore, the welfare gains from the optimal history-independent allocation is 30% lower than the constrained efficient allocation and less than half of the optimal complete information allocation. As highlighted in Figure 3, the welfare gain predominantly comes from the reallocation of population at the expense of creatin dispersion in the marginal utility across space.

4.5 Robustness to Alternative Parameterization

To be written.

5 Transitions in Response to Aggregate Shocks

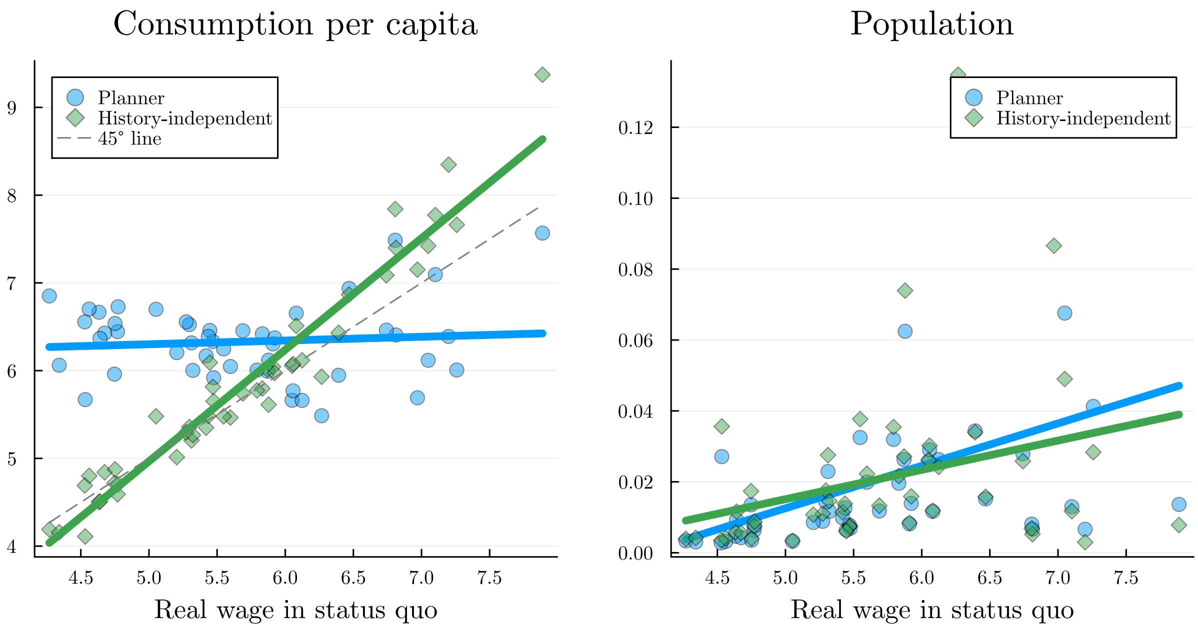
So far, we have focused on the steady state. Now we introduce aggregate shocks to our baseline economy to study the transition dynamics.

5.1 Introducing Aggregate Shock

We now study the response to aggregate shocks. We consider a one-time shocks to technologies. Let $x_t = 0, 1, \ldots$ denote the time elapsed since the arrival of the aggregate shock, where $x_t = 0$ indicates the before the arrival of the shock. If the shocks have not occurred, the technology evolves according to

$$f_{jt}(\{l_{kjt}\}_k, \{L_{kt}\}_k) = \begin{cases} f_j^1(\{l_{kjt}\}_k, \{L_{kt}\}_k) & \text{with prob. } p \\ f_j^0(\{l_{kjt}\}_k, \{L_{kt}\}_k) & \text{with prob. } 1-p \end{cases}, \tag{57}$$

Figure 3: Steady State Consumption and Population: Planner vs. History-Independent



Note: The left panel plots the average consumption per capita in each state against the real wage in the status quo economy. The square dots correspond to the constrained efficient economy, and the circle dot corresponds to the history independent solution. The solid line is the best linear fit. The right panel plots population and is analogous to the left panel.

where f_j^0 is the technology before the realization of the shock that is constant over time, and f_j^1 is the technology immediately after the realization of the shock. After the arrival of the aggregate shock, the technology is given by the deterministic sequence $\{f_j^x\}_{x=1}^{\infty}$, and we assume the sequence is convergent:

$$f_j^x \to f_j^\infty \quad \text{as} \quad x \to \infty. \tag{58}$$

The Bellman equation before the realization of the aggregate shock is

$$S_{it}^0(v) = \max_{C_{it}, \{v_k^{x\prime}, \mu_{ij}^x\}_{j, x\in\{0,1\}}} w_{it}^0\left(1+\alpha_{it}^0\right) - P_{it}^0 C_{it} + (1-\omega) S_{it}^{n,0} \tag{59}$$

$$+ (1-p)\left[\frac{1}{R}\sum_k \mu_{ij}^0 S_{jt+1}^0(v_k^{0\prime})\right] + p\left[\frac{1}{R}\sum_j \mu_{ij}^1 S_{jt+1}^1(v_j^{1\prime})\right] \tag{60}$$

$$\text{s.t.} \quad v = u_i(C_{it}) + \beta\omega\left\{(1-p)\left[\sum_j \mu_{ij}^0 v_j^{0\prime} - \psi(\{\mu_{ij}^0\}_j)\right] + p\left[\sum_j \mu_{ij}^1 v_j^{1\prime} - \psi(\{\mu_{ij}^1\}_j)\right]\right\} \tag{61}$$

$$\{\mu_{ij}^x\}_j \in \arg\max_{\tilde{\mu}_{ij}^x} \sum_j \tilde{\mu}_{ij}^x v_j^{x\prime} - \psi(\{\tilde{\mu}_{ij}^x\}_j) \quad \text{for } x \in \{0,1\}, \tag{62}$$

where the variables with superscript $x = 0, 1, \ldots$ denote those after x periods have passed since the arrival of the shock. The Bellman equation for newborn S^0_{jt} is likewise given by

$$S^n_{it} = \max_{\{v^{n,x}, v^{n,x}_{ij}, \mu^{n,x}_{ij}\}} (1-p)\left[\Lambda_i v^{n,0} + \frac{1}{R}\sum_j \mu^{n,0}_{ij} S_{jt+1}(v^{n,0}_{ij})\right] \tag{63}$$

$$+ p\left[\Lambda_i v^{n,1} + \frac{1}{R}\sum_j \mu^{n,1}_{ij} S_{jt+1}(v^{n,1}_{ij})\right] \tag{64}$$

$$\text{s.t.} \qquad v^{n,x} = \beta\left[\sum_j \mu^{n,x}_{ij} v^{n,x}_{ij} - \psi(\{\mu^{n,x}_{ij}\}_j)\right] \quad \text{for } x \in \{0,1\} \tag{65}$$

$$\mu^n_{ij} \in \arg\max_{\{\tilde{\mu}^n_{ij}\}_j} \sum_j \tilde{\mu}^{n,x}_{ij} v^{n,x}_{ij} - \psi(\{\tilde{\mu}^{n,x}_{ij}\}_j). \tag{66}$$

After the realization of aggregate shock, the Bellman equations are analogous to (**??**) and (**??**) indexed with $x = 1, 2, \ldots$

Importantly, we make following two assumptions. First, we assume that the probability that the aggregate shock hits the economy is arbitrarily small:

$$p \to 0. \tag{67}$$

Note that this assumption is distinct from the "MIT shock" (a one-time unanticipated shock). Here, the shock is anticipated, and the planner writes contingent plans in response to the shock.[10] This distinction is important because if shocks were unanticipated, the planner needs to re-optimize in response to the shock. If the planner could re-optimize, then the planner faces the time inconsistency problem that the new plan is not necessarily optimal from the viewpoint of the pre-shock plan. Second, we study the first-order approximation with respect to the size of the aggregate shock.

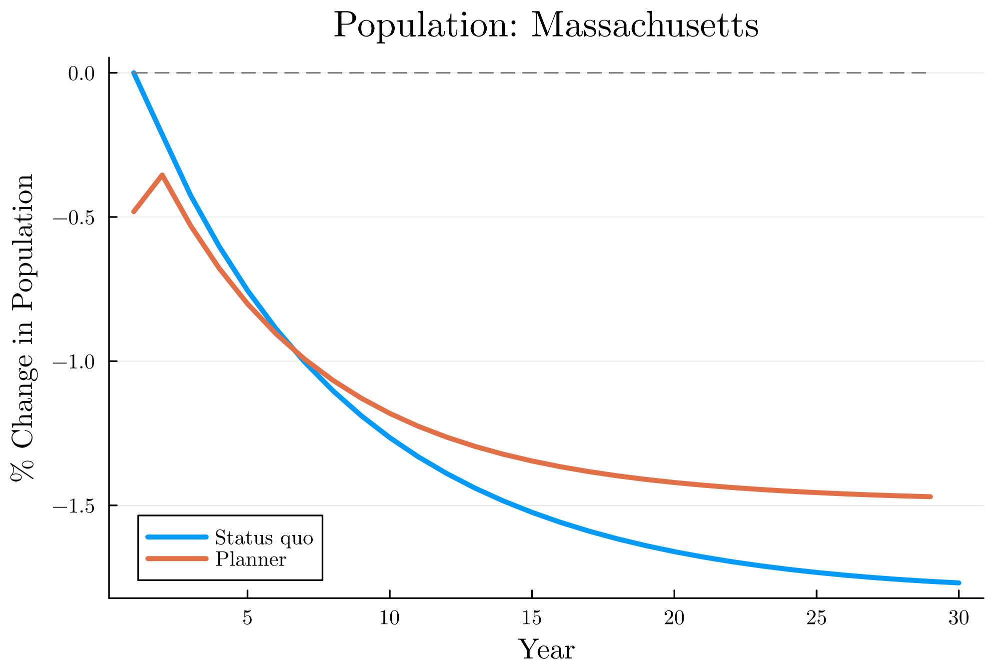
These two assumptions have two implications that dramatically simplify our analysis. First, the economy is in a deterministic steady state before the realization of the aggregate shock. Second, the response to an aggregate shock can be studied using the sequence space Jacobian (Auclert et al. 2021) around the deterministic steady state.[11] We describe the details of the computational algorithm in Appendix B.3.

We compare the response of constrained efficient allocation to those in the status quo economy. We continue to assume that the only available asset in the status quo economy is state non-contingent bonds. This implies that the response to the above shock is the same as the re-

[10] Mukoyama (2021) clarifies this difference in detail.

[11] Similar approaches appear in the context of optimal risk-sharing contracts (Fukui 2020) and endogenous portfolio choice (Auclert et al. 2024).

Figure 4: Population Response to 1% Reduction in Massachusetts Productivity



Note: The figure plots the response of Massachusets population in response to 1% permanent negative productivity shock to Massachusetts. The blue line is the status quo economy, and the red line is the constrained efficient allocation.

sponse to an "MIT shock," and therefore, the first-order response can be obtained by applying the sequence space Jacobian methodology in Auclert et al. (2021).

Figure 4 shows the result in response to 1% permanent negative productivity shock to Massachusetts. We find that the constrained efficient allocation involves more population reallocation in the short-run and less reallocation in the long-run, relative to the status-quo economy with exogenously incomplete market.

5.2 Spectrum Analysis and the Speed of Convergence

6 Concluding Remarks

To be written.

References

Adao, R., A. Costinot, and D. Donaldson (2017): "Nonparametric Counterfactual Predictions in Neoclassical Models of International Trade," *American Economic Review*, 107, 633–689.

Aiyagari, S. R. (1994): "Uninsured idiosyncratic risk and aggregate saving," *The Quarterly Journal of Economics*, 109, 659–684.

Allen, T. and D. Donaldson (2020): "Persistence and Path Dependence in the Spatial Economy," Working Paper 28059, National Bureau of Economic Research.

Anderson, S., A. De Palma, and J.-F. Thisse (1988): "A Representative Consumer Theory of the Logit Model," *International Economic Review*, 461–466.

Artuç, E., S. Chaudhuri, and J. McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100, 1008–1045.

Atkeson, A. and R. Lucas (1992): "On Efficient Distribution with Private Information," *The Review of Economic Studies*, 59, 427–453.

Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2021): "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models," *Econometrica*, 89, 2375–2408.

Auclert, A., M. Rognlie, L. Straub, and T. Tapak (2024): "When do endogenous portfolios matter for hank," *Unpublished manuscript*.

Balke, N. and T. Lamadon (2022): "Productivity Shocks, Long-Term Contracts, and Earnings Dynamics," *American Economic Review*, 112, 2139–77.

Bewley, T. (1986): "Stationary monetary equilibrium with a continuum of independently fluctuating consumers," *Contributions to mathematical economics in honor of Gérard Debreu*, 79.

Bilal, A. and E. Rossi-Hansberg (2023): "Anticipating Climate Change Across the United States," Working Paper 31323, National Bureau of Economic Research.

Burdett, K. and M. Coles (2003): "Equilibrium Wage-Tenure Contracts," *Econometrica*, 71, 1377–1404.

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

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018): "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," *The Review of Economic Studies*, 85, 2042–2096.

Carroll, C. D. (2006): "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics letters*, 91, 312–320.

Combes, P.-P. and L. Gobillon (2015): "The empirics of agglomeration economies," in *Handbook of regional and urban economics*, Elsevier, vol. 5, 247–348.

Costinot, A. and A. Rodríguez-Clare (2014): "Trade theory with numbers: Quantifying the consequences of globalization," in *Handbook of international economics*, Elsevier, vol. 4, 197–261.

Desmet, K. and F. Parro (2025): "Spatial Dynamics," Tech. rep., National Bureau of Economic Research.

Diamond, P. A. and J. A. Mirrlees (1978): "A model of social insurance with variable retirement," *Journal of Public Economics*, 10, 295–336.

Donald, E., M. Fukui, and Y. Miyauchi (2024): "Unpacking Aggregate Welfare in a Spatial Economy," *Working Paper*.

D'Amico, L. and M. Alekseev (2024): "Capital Market Integration and Growth Across the United States," Tech. rep., Working Paper.

Fajgelbaum, P. and C. Gaubert (2025): "Optimal spatial policies," Tech. rep., National Bureau of Economic Research.

Farhi, E. and I. Werning (2007): "Inequality and Social Discounting," *Journal of Political Economy*, 115, 365–402.

——— (2012): "Capital taxation: Quantitative explorations of the inverse Euler equation," *Journal of Political Economy*, 120, 398–445.

Fukui, M. (2020): "A theory of wage rigidity and unemployment fluctuations with on-the-job search," *Working Paper*.

Gaubert, C., P. Kline, and D. Yagan (2021): "Place-Based Redistribution," Working Paper 28337, National Bureau of Economic Research.

Giannone, E., Q. Li, N. Paixão, and X. Pang (2023): "Unpacking Moving," *Working Paper*.

Greaney, B., A. Parkhomenko, and S. Van Nieuwerburgh (2025): "Dynamic Urban Economics," Tech. rep., National Bureau of Economic Research, Inc.

Guerreiro, J., S. Rebelo, and P. Teles (2023): "Foreign Residents and the Future of Global Cities," Tech. rep., National Bureau of Economic Research.

HOFBAUER, J. AND W. SANDHOLM (2002): "On the Global Convergence of Stochastic Fictitious Play," *Econometrica*, 70, 2265–2294.

HOPENHAYN, H. A. AND J. P. NICOLINI (1997): "Optimal unemployment insurance," *Journal of political economy*, 105, 412–438.

HOWARD, G. AND H. SHAO (2022): "Internal Migration and the Microfoundations of Gravity," *Working Paper*.

HUGGETT, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of economic Dynamics and Control*, 17, 953–969.

IMROHOROĞLU, A. (1989): "Cost of business cycles with indivisibilities and liquidity constraints," *Journal of Political economy*, 97, 1364–1383.

KENNAN, J. AND J. WALKER (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79, 211–251.

KLEINMAN, B., E. LIU, AND S. REDDING (2023): "Dynamic Spatial General Equilibrium," *Econometrica*, 91, 385–424.

KOCHERLAKOTA, N. R. (2005): "Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation," *Econometrica*, 73, 1587–1621.

KURNAZ, M., M. MICHELINI, H. ÖZDENÖREN, AND C. SLEET (2023): "Tax Design in Dynamic Discrete Choice Economies," .

LHUILLIER, H. (2023): "Should I Stay or Should I Grow?" .

MIRRLEES, J. (1971): "An exploration in the theory of optimum income taxation," *The review of economic studies*, 38, 175–208.

MONGEY, S. AND M. E. WAUGH (2024): "Discrete choice, complete markets, and equilibrium," Tech. rep., National Bureau of Economic Research.

——— (2025): "Pricing Inequality," Tech. rep., National Bureau of Economic Research.

MUKOYAMA, T. (2021): "MIT shocks imply market incompleteness," *Economics Letters*, 198, 109666.

ORCHARD, J., V. A. RAMEY, AND J. F. WIELAND (2023): "Micro MPCs and macro counterfactuals: the case of the 2008 rebates," Tech. rep., National Bureau of Economic Research.

O'Connor, D. G. (2024): "Revitalize or Relocate: Optimal Place Based Transfers for Local Recessions," .

Prescott, E. C. and R. M. Townsend (1984): "General competitive analysis in an economy with private information," *International Economic Review*, 1–20.

Rogerson, W. P. (1985): "Repeated moral hazard," *Econometrica*, 69–76.

Rosenthal, S. S. and W. C. Strange (2004): "Evidence on the nature and sources of agglomeration economies," in *Handbook of regional and urban economics*, Elsevier, vol. 4, 2119–2171.

Shi, S. (2009): "Directed search for equilibrium wage–tenure contracts," *Econometrica*, 77, 561–584.

Souchier, M. (2022): "The pass-through of productivity shocks to wages and the cyclical competition for workers," Tech. rep., Working Paper.

Veracierto, M. (2022): "Moral Hazard, Optimal Unemployment Insurance, and Aggregate Dynamics," .

——— (2023): "Aggregate dynamics in mirrlees economies: The case of persistent shocks," *Working Paper*.

A Theory Appendix

A.1 Lagrangian of the Recursive Planning Problem

Let P_{it}, w_{it}, and $w_{it}\alpha_{it}$ denote the Lagrangian multipliers of (10), (11), and (12), respectively. Let $\mathcal{S}_t(\phi)$ denote the associated Lagrangian of the problem (7). It is given by

$$\begin{aligned}\mathcal{S}_t(\phi) = \max_{\{C_{it}(v,\epsilon),\ell_i^n(\epsilon),\ell_i(v,\epsilon),v_j^{n\prime}(\epsilon),v_{ij}'(v,\epsilon),l_{ji},\phi',L_i,P_{it},w_{it},\alpha_{it}\}} & \sum_i \Lambda_i v_i^n (1-\omega) \int d\phi_i \\ & + \sum_i P_{it} f_{it}(\{l_{ki}\}_k ; \{L_k\}_k) - \sum_i P_{it} \int C_{it}(v) d\phi_i \\ & + \sum_i w_{it} \int d\phi_i - \sum_i w_{it} \sum_j l_{ij} \\ & + \sum_i \alpha_{it} w_{it} \int d\phi_i - \sum_i L_i \\ & + \frac{1}{R} \mathcal{S}_{t+1}(\phi')\end{aligned} \tag{A.1}$$

subject to

$$v_i^n = \beta \mathbb{E}_{it}\Big[\sum_j \mathbb{I}[\ell_i^n(\epsilon) = j] \left\{v_{ij}^{n\prime}(\epsilon) + \epsilon_{i,j}\right\}\Big] \tag{A.2}$$

$$v = u_i(C_{it}(v)) + \beta\omega \mathbb{E}_{it}\Big[\sum_j \mathbb{I}[\ell_i(v,\epsilon) = j] \left\{v_{ij}'(v,\epsilon) + \epsilon_{i,j}\right\}\Big] \tag{A.3}$$

and the law of motion of the distribution

$$\begin{aligned}\phi_j'(\mathbb{V}) = \omega \sum_i & \mathbb{E}_{it}\left[\phi_i(v_{ij}^{-1}(\mathbb{V},\epsilon))\mathbb{I}[\ell_i(v_{ij}^{-1}(\mathbb{V},\epsilon),\epsilon) = j]\right] \\ & + (1-\omega) \int d\phi_i \mathbb{E}_{it}\left[\mathbb{I}[\ell_i^n(\epsilon) = j]\mathbb{I}[v_{ij}^{n\prime}(\epsilon) \in \mathbb{V}]\right].\end{aligned} \tag{A.4}$$

We now guess and verify that the value function takes the following form:

$$\mathcal{S}_t(\phi) = \sum_i \int S_{it}(v) d\phi_i + D_t. \tag{A.5}$$

First, observe that the flow value in (A.1) is additively separable in ϕ_i and i. Second, under the guess, the continuation value can be rewritten as

$$\begin{aligned}\frac{1}{R}\mathcal{S}(\phi') = \frac{1}{R}\omega \sum_i \mathbb{E}_{it}\left[\sum_j \mathbb{I}[\ell_i(v,\epsilon)=j]S_{jt}(v'_{ij}(v))\right] \\ + \frac{1}{R}(1-\omega)\sum_i \mathbb{E}_{it}\left[\sum_j \mathbb{I}[\ell_i^n(\epsilon)=j]S_{jt}(v_{ij}^{n'})\right]\end{aligned}. \tag{A.6}$$

With these observations, it is immediate to see that the guess satisfies the Bellman equation (A.1) with $S_{it}(v)$ solving (18) and D_t solving (15).

A.2 Proofs

A.2.1 Proof of Proposition 1

Let Ξ_i be the Lagrangian multiplier on the promise-keeping constraint (19). The first-order condition with respect to $v'_{ij}(\epsilon)$ is given by

$$\frac{1}{R}\partial_v S_{jt+1}(v'_{ij}(\epsilon)) + \beta\Xi_i = 0. \tag{A.7}$$

The first-order condition with respect to C_{it} is

$$P_{it} = u'_i(C_{it})\Xi_i. \tag{A.8}$$

The envelope condition is

$$\partial_v S_{it}(v) = -\Xi_i. \tag{A.9}$$

Combining the above three expressions, we have

$$-\frac{1}{R}\frac{P_{jt+1}}{u'_j(C_{jt+1})} + \beta\frac{P_{it}}{u'_i(C_{it})} = 0, \tag{A.10}$$

which is (22).

The location choice maximizes the Lagrangian:

$$\ell_i(v,\epsilon) \in \arg\max_l \beta\omega\Xi_i[v_{ilt+1}(v,\epsilon),\epsilon) + \epsilon_{il}] + \frac{1}{R}S_{lt+1}(v_{il}(v,\epsilon)). \tag{A.11}$$

Substituting the expression for Ξ_i in (A.9) gives (23).

A.2.2 Proof of Proposition 2

Let Ξ_{jt} be the Lagrangian multiplier on the promise-keeping constraint (32). The first order conditions are

$$\Xi_{it} u_i'(C_{it}) = P_{it} \tag{A.12}$$

$$\frac{1}{R}\mu_{im}\partial_{v'_{im}} S_{mt+1}(v'_{im}) + \Xi_{it}\beta\mu_{im} + \frac{1}{R}\sum_k \frac{\partial \mu_{ik}}{\partial v'_{im}} S_{kt+1}(v'_k) = 0 \tag{A.13}$$

The envelope condition is

$$\partial_v S_{it}(v) = -\Xi_{it}. \tag{A.14}$$

Combining the three equations,

$$-\frac{1}{R}\mu_{im}\frac{P_{mt+1}}{u'_m(C_{mt+1})} + \frac{P_{it}}{u'_i(C_{it})}\beta\mu_{im} + \frac{1}{R}\sum_k \frac{\partial \mu_{ik}}{\partial v'_m} S_{kt+1}(v'_k) = 0 \tag{A.15}$$

Rewriting the above equation,

$$\mu_{im}\left[\beta R \frac{u'_m(C_{mt+1})/P_{mt+1}}{u'_i(C_{it})/P_{jt}} - 1\right] + \sum_k \frac{\partial \mu_{ik}}{\partial C_{mt+1}} \frac{S_{kt+1}(v'_k)}{P_{mt+1}} = 0, \tag{A.16}$$

as desired.

A.2.3 Proof of Proposition 3

The planning problem in a recursive form is

$$\mathcal{S}_t(\{v_{it}, L_{it}\}) = \max_{\{v_j, v_i^n, C_{it}, \mu_{ij}\}} \sum_i [w_{it}(1+\alpha_{it}) - P_{it}C_{it}]L_{it} \tag{A.17}$$

$$+ (1-\omega)\sum_i \Xi_i v_i^n L_{it} + \frac{1}{R}\mathcal{S}_{t+1}(\{v_{jt+1}, L_{jt+1}\}) \tag{A.18}$$

$$\text{s.t.} \qquad v_{it} = u_i(C_{it}) + \beta\omega\left[\sum_j \mu_{ij}v_{jt+1} - \psi(\{\mu_{ij}\}_j)\right] \tag{A.19}$$

$$v_{it}^n = \beta\left[\sum_j \mu_{ij}v_{jt+1} - \psi(\{\mu_{ij}\}_j)\right] \tag{A.20}$$

$$\mu_{ij} \in \arg\max_{\{\mu_{ij}\}_j} \sum_j \mu_{ij}v_{jt+1} - \psi(\{\mu_{ij}\}_j) \tag{A.21}$$

$$L_{jt+1} = \sum_k L_{kt}\mu_{kjt+1} \tag{A.22}$$

Let $\varkappa_{it}L_{it}$ be the Lagrangian multiplier on constraint (A.19). The first-order condition w.r.t. v_j is

$$\frac{1}{R}\left[\frac{\partial S_{t+1}}{\partial v_{jt+1}} + \sum_k \frac{\partial \mu_{ikt}}{\partial v_j}L_{it}\frac{\partial \mathcal{S}}{\partial L_{kt+1}}\right] + \beta\sum_i(\omega\varkappa_{it} + (1-\omega)\Xi_i)L_{it}\mu_{ij} = 0. \tag{A.23}$$

The first-order condition w.r.t. C_{jt} is

$$P_{jt} = \varkappa_{jt}u'(C_{jt}) \tag{A.24}$$

The envelope conditions are

$$\frac{\partial \mathcal{S}_t}{\partial v_{jt}} = -\varkappa_{jt}L_{jt} \tag{A.25}$$

$$\frac{\partial \mathcal{S}_t}{\partial L_{jt}} = w_{jt} - P_{jt}C_{jt} + (1-\omega)\Xi_j v_j^n + \frac{1}{R}\sum_k \mu_{jkt}\frac{\partial \mathcal{S}_{t+1}}{\partial L_{kt+1}} \tag{A.26}$$

Combining the above expressions,

$$\frac{1}{R}\left[-\frac{P_{jt+1}}{u'(C_{jt+1})}L_{jt+1} + \sum_k \frac{\partial \mu_{ikt}}{\partial v_j}L_{it}S_{kt+1}\right] + \beta\sum_i\left(\omega\frac{P_{it}}{u'(C_{it})} + (1-\omega)\Xi_i\right)L_{it}\mu_{ij} = 0, \tag{A.27}$$

which we can further rewrite as

$$\sum_i L_{it}\left[\mu_{ij}\left[\beta R\omega\frac{u'(C_{jt+1})/P_{jt+1}}{u'(C_{it})/P_{it}}+\beta R(1-\omega)\Xi_i\frac{u'(C_{jt+1})}{P_{jt+1}}-1\right]+\sum_k\frac{\partial\mu_{ikt}}{\partial C_{j+1}}\frac{S_{kt+1}}{P_{jt+1}}\right]=0. \tag{A.28}$$

A.2.4 Proof of Corollary 1

We divide both sides of the expression in Proposition 2 by $u_j(C_{jt+1})/P_{jt+1}$ to obtain

$$\mu_{ijt}\left[\beta R\frac{1}{u_i'(C_{it})/P_{it}}-u_j'(C_{jt+1})/P_{jt+1}\right]+\sum_k\frac{\partial\mu_{ik}}{\partial C_{jt+1}}u_j'(C_{jt+1})S_{kt+1}(v_{ik}')=0, \tag{A.29}$$

We further rewrite this as

$$\mu_{ijt}\left[\beta R\frac{1}{u_i'(C_{it})/P_{it}}-u_j'(C_{jt+1})/P_{jt+1}\right]+\sum_k\frac{\partial\mu_{ik}}{\partial u_{jt+1}}S_{kt+1}(v_{ik}')=0, \tag{A.30}$$

where $u_{jt+1}\equiv u_j(C_{jt+1})$. With a slight change in notation, we can equivalently express household's migration decisions as

$$V_{it}=\max_{\{\mu_{ik}\}_k}\sum_{\mu_{ik}}\mu_{ik}[u_k(C_{kt+1})+\beta\omega V_{kt+1}]-\psi(\{\mu_{ik}\}_k) \tag{A.31}$$

From this expression, uniformly increasing u_{jt+1} for all j would not affect the choice probability:

$$\sum_j\frac{\mu_{ik}}{\partial u_{jt+1}}=0 \tag{A.32}$$

for all i,k.

Using this property and summing (A.30) across j, we have

$$\beta R\frac{1}{u_i'(C_{it})/P_{it}}-\sum_j\mu_{ij}u_j'(C_{jt+1})/P_{jt+1}+\sum_k S_{kt+1}(v_{ik}')\underbrace{\sum_j\frac{\partial\mu_{ik}}{\partial u_{jt+1}}}_{=0}=0, \tag{A.33}$$

so that

$$\frac{1}{u_i'(C_{it})/P_{it}}=\frac{1}{\beta R}\mathbb{E}_{it}\left[\frac{u_j'(C_{jt+1})}{P_{jt+1}}\right], \tag{A.34}$$

as desired.

A.3 Decentralization

We present one example of implementation with no private savings, either because households do not have access to the credit market or because the government bans private savings. The underlying environment remains the same as described in Section 2. Here we focus on explaining the market structure.

The households supply labor in each location i at wage w_{it}. The price of final goods in each location is P_{it}. Let $\ell_t \in \{1, \ldots, J\}$ denote the location of living at time t, and let ℓ^t denote the history of location of living of any household. The government sends transfers $T_t(\ell^t)$ as a function of history of living locations.

The household problem in a recursive form is

$$v_{jt}(\ell^t) = \max_{C_{jt}, \{\mu_{jk}\}_k} u_{jt}(C_{jt}) + \beta\omega \left[\sum_k \mu_{jk} v_{kt+1}(\{\ell^t, k\}) - \psi(\{\mu_{jk}\}) \right] \tag{A.35}$$

$$\text{s.t.} \quad P_{jt} C_{jt} = w_{jt} + T_t(\ell^t). \tag{A.36}$$

Let $C_{jt}(\ell^t)$ and $\mu_{jkt}(\ell^t)$ denote the policy functions associated with the above problem.

The firm takes prices and population size $\{L_{kt}\}_k$ as given. The profit maximization problem of firm in location i is

$$\max_{\{l_{kjt}\}_k} P_{jt} f_{jt}(\{l_{kjt}\}, \{L_{kt}\}_k) - \sum_k w_{kt} l_{kj}. \tag{A.37}$$

Here, agglomeration/congestion forces are externalities that are not internalized by private agents. The government budget constraint is

$$\sum_{\ell^t} T_t(\ell^t) = 0. \tag{A.38}$$

Let $\Phi(\ell^t)$ denote the measure of households with history ℓ^t. The goods market clearing condition is

$$\int C_{jt}(\ell^t) d\Phi = f_{jt}(\{l_{kjt}\}, \{L_{kt}\}_k). \tag{A.39}$$

The factor market clearing condition is

$$\sum_j l_{kjt} = \int \mathbb{I}[\ell_t = k] d\Phi = L_{kt}. \tag{A.40}$$

The distribution evolves according to the following law of motion:

$$\Phi_{t+1}(\{\ell^t, k\}) = \mu_{\ell_t kt}(\ell^t)\Phi_t(\ell^t). \tag{A.41}$$

In the above decentralized equilibrium, appropriate choice of the transfer system $T_t(\ell^t)$ implements the constrained efficient allocation characterized in Proposition 2. To see this, first note that the continuation value in the constrained efficient allocation is only depends on location of living in the next period and the promised utility. Therefore, given the initial location that each household is born, the promised value is only a function of the history of living locations. Let $v(\ell^t)$ denote the promised value with a history of living location ℓ^t. Then, $C_{\ell_t t}(v(\ell^t))$ is the consumption of households currently in location ℓ_t with a history ℓ^t in the constrained efficient allocation.

Consider the following transfer system:

$$T_t(\ell^t) = P_{\ell_t t} C_{\ell_t t}(v(\ell^t)) - w_{\ell_t t}. \tag{A.42}$$

From the budget constraint, it is immediate to see that such transfer system implements the constrained efficient allocation as long as $\{P_{jt}, w_{jt}\}_j$ in the decentralized equilibrium coincide with those in the constrained efficient allocation. In fact, they do. To see why, the migration probabilities are identical given $\{P_{jt}, w_{jt}\}_j$. The optimality conditions of (A.39) is identical to (16) in the constrained efficient allocation. The market clearing conditions and the evolution of the distribution are identical in both economies by construction. Finally, the government budget (A.38) is satisfied in Warlas' law. Given that all the conditions in two economies coincide, the transfer scheme (A.42) implements the constrained efficient allocation as a decentralized equilibrium.

A.4 Capital Accumulation

We now introduce capital accumulation in the baseline model. Assume that in each location j, there is a capital stock denoted as K_{jt}. The production function of the final goods at location j is now given by

$$Y_{jt} = f_{jt}(\{l_{kjt}\}_k, \{k_{kjt}\}_k, \{L_{kt}\}_k), \tag{A.43}$$

where k_{kjt} denotes the use of capital stock from location k in location j. The capital stock in location j at time t depreciates at rate δ_{jt}, but the final goods in location j can be invested into the capital stock in the same location. The law of motion of capital stock in location j is

$$K_{jt+1} = K_{jt}(1-\delta_{jt}) + I_{jt}, \tag{A.44}$$

where I_{jt} is the investment. The investment incurs the adjustment cost of the form $\Psi_{jt}(I_{jt}, K_{jt})$ in the units of final goods in location j. The capital market clearing condition is

$$\sum_k k_{jkt} = K_{jt}. \tag{A.45}$$

The goods market clearing condition in location j is modified as

$$\int_0^1 C_{jt}(h)dh + I_{jt} + \Psi_{jt}(I_{jt}, K_{jt}) = f_{jt}\left(\{l_{kjt}\}_k, \{k_{kjt}\}_k, \{L_{kt}\}_k\right). \tag{A.46}$$

The rest of the environment remains unchanged.

In this environment, there is no change in the component planning problem (31) and (34). The only change comes from D_t in (14). Let r_{jt} be the Lagrangian multiplier on (A.45). The term D_t now includes the distribution of capital stock in each location as a state variable and is given by

$$D_t(\{K_{jt}\}) = \max_{\{l_{ij},k_{ij},L_i,K_{jt+1}\}} \sum_i P_{it} f_{it}\left(\{l_{ki}\}_k, \{k_{ki}\}_k, \{L_k\}_k\right) - \sum_i w_{it} \sum_j l_{ij} \tag{A.47}$$

$$- \sum_i \alpha_{it} w_{it} L_i - \sum_i r_{it} K_{it} - \sum_i r_{it} \sum_j k_{ijt} - \sum_i P_{it}\left[I_{it} + \Psi_{it}(I_{it}, K_{it})\right] \tag{A.48}$$

$$+ \frac{1}{R} D_{t+1}(\{K_{jt+1}\}). \tag{A.49}$$

$$\text{s.t.} \quad K_{jt+1} = K_{jt}(1-\delta_{jt}) + I_{jt}. \tag{A.50}$$

The first-order conditions with respect to l_{ij} and L_i remain essentially the same as in the main text:

$$P_{it} \frac{\partial f_{it}\left(\{l_{kit}\}_k, \{k_{kit}\}_k, \{L_{kt}\}_k\right)}{\partial l_{ki}} = w_{kt} \tag{A.51}$$

$$\sum_i P_{it} \frac{\partial f_{it}\left(\{l_{kit}\}_k, \{k_{kit}\}_k, \{L_{kt}\}_k\right)}{\partial L_k} = \alpha_{kt} w_{kt}. \tag{A.52}$$

The optimality condition of spatial allocation of capital services k_{ij} is given by

$$P_{it} \frac{\partial f_{it}\left(\{l_{kit}\}_k, \{k_{kit}\}_k, \{L_{kt}\}_k\right)}{\partial k_{ki}} = r_{kt}. \tag{A.53}$$

The first-order condition with respect to investment I_{jt} is

$$P_{jt}\left(1+\partial_I \Psi_{jt}(I_{jt}, K_{jt})\right) = \frac{\partial D_{t+1}(\{K_{jt+1}\})}{\partial K_{jt+1}}. \tag{A.54}$$

The envelope condition is

$$\frac{\partial D_t(\{K_{jt}\})}{\partial K_{jt}} = r_{jt} - P_{jt}\partial_K \Psi_{jt}(I_{jt}, K_{jt}) + (1-\delta)\frac{1}{R}\frac{\partial D_{t+1}(\{K_{jt+1}\})}{\partial K_{jt+1}}. \tag{A.55}$$

Let

$$q_{jt} \equiv \frac{\partial D_t(\{K_{jt}\})}{\partial K_{jt}} \tag{A.56}$$

be the "marginal q" of capital stock in location j at time t. Using this expression, we can equivalently write (A.54) and (A.55) as

$$P_{jt}\left(1+\partial_I \Psi_{jt}(I_{jt}, K_{jt})\right) = q_{jt+1}. \tag{A.57}$$

and

$$q_{jt} = r_{jt} - P_{jt}\partial_K \Psi_{jt}(I_{jt}, K_{jt}) + (1-\delta)\frac{1}{R}q_{jt+1}. \tag{A.58}$$

Therefore, optimal investment follows similarly to what is prescribed by the q-theory of investment.

A.5 Ex-Ante Heterogeneous Household Types

In the baseline model, we have assumed that households are ex-ante homogeneous. We now consider an extension of the baseline model to multiple ex-ante heterogeneous household types.

There are M heterogeneous household types (e.g., race, skills, or gender). Each household belongs to one of the types indexed by $\theta \in \{\theta_1, \ldots, \theta_M\}$, each of which has a mass ℓ^θ. We allow arbitrary heterogeneity across households with respect to θ including preferences, location preference shock distribution, and death probability. When a household of type θ dies, they are replaced by a newborn of the same type, so the mass of type θ remains fixed at ℓ^θ. Importantly, we assume the ex-ante types are observable to the planner.

The technology to produce the final goods consumed by the household θ in location j at time

t is

$$Y_{jt}^{\theta} = f_{jt}^{\theta}(\{l_{kjt}^{\tilde{\theta},\theta}\}_{j,\tilde{\theta}}, \{L_{kt}^{\tilde{\theta}}\}_{k,\tilde{\theta}}), \tag{A.59}$$

where $l_{kjt}^{\tilde{\theta},\theta}$ denotes the labor services of type $\tilde{\theta}$ shipped from location k to j used to produce final goods for consumption goods of type θ, and L_{it}^{θ} is the population size of households of type θ in location i at time t. Here, we allow for agglomeration/congestion forces to depend arbitrarily on the population size of different household types.

The planner's objective is to maximize the following social welfare function:

$$\mathcal{W}_0 = \sum_{t=0}^{\infty} \frac{1}{R^t} \sum_{i=1}^{J} \Lambda_i^{\theta} v_{it}^{\theta,n}(1-\omega^{\theta}) \int_0^1 \mathbb{I}[\ell_t^{\theta}(h) = i] dh, \tag{A.60}$$

where Λ_i^{θ} is the welfare weight attached to household of type θ born in location i.

Let φ_j^{θ} be the distribution over promised utility of households of type θ living in location j. The goods market clearing condition for consumption goods of type θ is

$$\int C_{jt}^{\theta}(v)^{\theta} d\phi_j^{\theta} = f_{jt}^{\theta}(\{l_{kjt}^{\tilde{\theta},\theta}\}_{j,\tilde{\theta}}, \{L_{kt}^{\tilde{\theta}}\}_{k,\tilde{\theta}}). \tag{A.61}$$

The labor market clearing condition for type θ is

$$\sum_{k,\tilde{\theta}} l_{jkt}^{\theta,\tilde{\theta}} = \int \mathbb{I}[\ell_t^{\theta}(v,\epsilon) = j] d\phi_j^{\theta}, \tag{A.62}$$

and the following equation dictates the agglomeration forces:

$$\int \mathbb{I}[\ell_t^{\theta}(v,\epsilon) = j] d\phi_j^{\theta} = L_{jt}^{\theta}. \tag{A.63}$$

The evolution of distribution is

$$\begin{aligned} \phi_j^{\theta\prime}(\mathbb{V}) = \omega^{\theta} \sum_i \mathbb{E}_{it}^{\theta} \left[\phi_i^{\theta}(v_{ij}^{\theta,-1}(\mathbb{V},\epsilon)) \mathbb{I}[\ell_i^{\theta}(v_{ij}^{\theta,-1}(\mathbb{V},\epsilon),\epsilon) = j] \right] \\ + (1-\omega^{\theta}) \int d\phi_i^{\theta} \mathbb{E}_{it}^{\theta} \left[\mathbb{I}[\ell_i^{n,\theta}(\epsilon) = j] \mathbb{I}[v_{ij}^{\theta,n\prime}(\epsilon) \in \mathbb{V}] \right]. \end{aligned} \tag{A.64}$$

Given all these environments, the value function D_t in (15) in the main text is now replaced by

$$D_t = \max_{\{l_{ij}^{\tilde{\theta},\theta}, L_i^\theta\}} \sum_{i,\theta} P_{it}^\theta f_{it}^\theta(\{l_{ki}^{\tilde{\theta},\theta}\}_{k,\tilde{\theta}}; \{L_k^{\tilde{\theta}}\}_{k,\tilde{\theta}}) - \sum_{i,\theta} w_{it}^\theta \sum_{j,\tilde{\theta}} l_{ij}^{\theta,\tilde{\theta}} - \sum_{i,\theta} \alpha_{it}^\theta w_{it}^\theta L_i^\theta + \frac{1}{R} D_{t+1}, \quad \text{(A.65)}$$

where P_{jt}^θ, w_{jt}^θ, and α_{jt}^θ are Lagrangian multipliers on (A.61), (A.62), and (A.63). The first-order optimality conditions are

$$P_{it}^\theta \frac{\partial f_{it}^\theta(\{l_{ki}^{\tilde{\theta},\theta}\}_{k,\tilde{\theta}}; \{L_k^{\tilde{\theta}}\}_{k,\tilde{\theta}})}{\partial l_{ki}^{\tilde{\theta},\theta}} = w_{kt}^\theta \quad \text{(A.66)}$$

$$\sum_{i,\theta} P_{it}^\theta \frac{\partial f_{it}^\theta(\{l_{ki}^{\tilde{\theta},\theta}\}_{k,\tilde{\theta}}; \{L_k^{\tilde{\theta}}\}_{k,\tilde{\theta}})}{\partial L_k^\theta} = \alpha_{kt}^\theta w_{kt}^\theta. \quad \text{(A.67)}$$

The component planning problems are essentially the same as (31) except that now everything is indexed by θ:

$$S_{it}^\theta(v) = \max_{\{C_{it}^\theta, v_{ij}^{\theta\prime}, \mu_{ij}^\theta\}} w_{it}^\theta(1+\alpha_{it}^\theta) - P_{it}^\theta C_{it}^\theta + (1-\omega^\theta) S_{it}^{\theta,n} + \frac{1}{R}\omega^\theta \sum_j \mu_{ij}^\theta S_{jt+1}^\theta(v_{ij}^{\theta\prime}) \quad \text{(A.68)}$$

$$\text{s.t.} \quad v = u_i^\theta(C_{it}^\theta) + \beta^\theta \omega^\theta \left[\sum_j \mu_{ij}^\theta v_{ij}^{\theta\prime} - \psi_{it}^\theta(\{\mu_{ij}^\theta\}_j)\right] \quad \text{(A.69)}$$

$$\{\mu_{ij}^\theta\}_j \in \arg\max_{\{\tilde{\mu}_{ij}^\theta\}_j} \beta^\theta \omega^\theta \left[\sum_j \tilde{\mu}_{ij}^\theta v_{ij}^{\theta\prime} - \psi_{it}^\theta(\{\tilde{\mu}_{ij}\}_j)\right] \quad \text{(A.70)}$$

and the following replaces (34):

$$S_{it}^{\theta,n} = \max_{v_i^{\theta,n}, \{v_{ij}^{\theta,n}, \mu_{ij}^{\theta,n}\}} \Lambda_i^\theta v_i^{\theta,n} + \frac{1}{R} \sum_j \mu_{ij}^{\theta,n} S_{jt+1}^\theta(v_{ij}^{\theta,n}) \quad \text{(A.71)}$$

$$\text{s.t.} \quad v_i^{\theta,n} = \beta^\theta \sum_j \left[\mu_{ij} v_{ij}^{\theta,n\prime} - \psi_{it}^\theta(\{\mu_{ij}^{\theta,n}\}_j)\right] \quad \text{(A.72)}$$

$$\{\mu_{ij}^{\theta,n}\}_j \in \arg\max_{\{\tilde{\mu}_{ij}^{\theta,n}\}} \beta^\theta \sum_j \left[\tilde{\mu}_{ij}^{\theta,n} v_{ij}^{\theta,n\prime} - \psi_{it}^\theta(\{\tilde{\mu}_{ij}^{\theta,n}\}_j)\right]. \quad \text{(A.73)}$$

The following formula is analogue of Proposition 2:

$$\mu_{ijt}^\theta \left[\beta^\theta R \frac{u_j^{\theta\prime}(C_{jt+1}^\theta)/P_{jt+1}^\theta}{u_i^{\theta\prime}(C_{it}^\theta)/P_{it}^\theta} - 1\right] + \underbrace{\sum_k \frac{\partial \mu_{ik}^\theta}{\partial C_{jt+1}^\theta} \frac{S_{kt+1}^\theta(v_{ik}^{\theta\prime})}{P_{jt+1}^\theta}}_{\equiv \xi_{ijt}^\theta} = 0. \quad \text{(A.74)}$$

A.6 Lagged Agglomeration/Congestion Forces

In the baseline model, we assumed that agglomeration/congestion forces only depend on the contemporaneous population size distribution. We now allow for the agglomeration/congestion forces to depend on the lagged population size distribution.

The only modification is that now the production function takes the following form:

$$Y_{jt} = f_{jt}\left(\{l_{kjt}\}_k, \{L^t_{kt}\}_k, \{L^t_{kt-1}\}_k, \dots, \{L^t_{kt-T^L}\}_k\right), \tag{A.75}$$

which replaces (3) in the main text. Here L^t_{kt-s} denotes the population size in location k at time $t-s$ that enters as agglomeration/congestion forces for production at time t. We allow for the population size of $t-1, \dots, t-T^L$ with $T^L \geq 1$ enters the production function. The lagged population sizes are defined as

$$L^t_{jt-s} = \int d\phi_{jt-s} \quad \text{for } s = 1, \dots, T^L. \tag{A.76}$$

Let $\alpha^t_{jt-s} w_{jt}$ be the Lagrangian multiplier on the above equation.

Given all these modification, the value function D_t in (15) in the main text is now replaced by

$$D_t = \max_{\{l_{ij}, L_i\}} \sum_i P_{it} f_{jt}\Big(\{l_{kjt}\}_k, \{L^t_{kt}\}_k, \{L^t_{kt-1}\}_k, \dots, \{L^t_{kt-T}\}_k\Big) - \sum_i w_{it} \sum_j l_{ij} \tag{A.77}$$

$$- \sum_{s=0}^{T^L} \sum_i \alpha^t_{it-s} w_{it} L^t_{it-s} + \frac{1}{R} D_{t+1}. \tag{A.78}$$

The first-order optimality conditions are

$$P_{it} \frac{\partial f_{it}}{\partial l_{ki}} = w_{kt} \tag{A.79}$$

$$\sum_i P_{it} \frac{\partial f_{it}}{\partial L_{kt-s}} = w_{kt-s} \alpha^t_{kt-s}. \tag{A.80}$$

The value function $S_{it}(v)$ now takes into the fact that increasing current population size of a location changes the technology in the future:

$$S_{it}(v) = \max_{\{C_{it}, v'_{ij}, \mu_{ij}\}} w_{it} \left[(1+\alpha_{it}) + \textstyle\sum_{s=1}^{T^L} \frac{1}{R^s} \alpha_{it}^{t+s}\right] - P_{it} C_{it} + (1-\omega) S_{it}^n + \frac{1}{R} \omega \sum_j \mu_{ij} S_{jt+1}(v'_{ij})$$

$$\text{s.t.} \qquad v = u_i(C_{it}) + \beta\omega \left[\sum_j \mu_{ij} v'_{ij} - \psi_{it}(\{\mu_{ij}\}_j)\right]$$

$$\{\mu_{ij}\}_j \in \arg\max_{\{\tilde{\mu}_{ij}\}_j} \beta\omega \left[\sum_j \tilde{\mu}_{ij} v'_{ij} - \psi_{it}(\{\tilde{\mu}_{ij}\}_j)\right].$$

The value function for newborns $S_{it}^n(v)$ remains unchanged and is given by (34).

B Quantitative Appendix

B.1 Details on Calibration

The aggregate trade flow from location i to j, denoted as $x_{ij} \equiv w_i l_{ij}$, is given by

$$x_{ij} = \frac{(w_i/\mathcal{A}_{ij}(L_i))^{1-\sigma}}{\sum_k (w_k/\mathcal{A}_{kj}(L_k))^{1-\sigma}} \int P_j C_j(a) d\varphi_j. \tag{B.1}$$

Taking the ratio of x_{ij} to x_{jj}, we have

$$\frac{x_{ij}}{x_{jj}} = \frac{(w_i/\mathcal{A}_{ij}(L_i))^{1-\sigma}}{(w_j/\mathcal{A}_{jj}(L_j))^{1-\sigma}}, \tag{B.2}$$

which we can rewrite as

$$\frac{1}{\mathcal{A}_{ij}(L_i)} = \left(\frac{x_{ij}}{x_{jj}}\right)^{\frac{1}{1-\sigma}} \frac{w_j}{w_i} \frac{1}{\mathcal{A}_{jj}(L_j)} \tag{B.3}$$

The price index of location j is

$$P_j = \left[\sum_i (w_i/\mathcal{A}_{ij}(L_i))^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{B.4}$$

$$= \left[\sum_i \left(\left(\frac{x_{ij}}{x_{jj}}\right)^{\frac{1}{1-\sigma}} \frac{w_j}{w_i} \frac{1}{\mathcal{A}_{jj}(L_j)} w_i\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{B.5}$$

$$= \frac{w_j}{\mathcal{A}_{jj}(L_j)} \frac{1}{(x_{jj})^{\frac{1}{1-\sigma}}} \left[\sum_i x_{ij}\right]^{\frac{1}{1-\sigma}} \tag{B.6}$$

$$= \frac{w_j}{A_{jj} L_j^{\alpha}} \frac{1}{(x_{jj})^{\frac{1}{1-\sigma}}} \left[\sum_i x_{ij}\right]^{\frac{1}{1-\sigma}}, \tag{B.7}$$

where we used (B.9) in the second line. As a result, conditional on the choice of (σ, α), we can infer A_{jj} given the data on trade flows $\{x_{ij}\}$, price index $\{P_j\}$, population size, L_j, and output per capita, $w_j = \sum_j x_{ij}/L_i$:

$$A_{jj} = \frac{w_j}{P_j L_j^{\alpha}} \frac{1}{(x_{jj})^{\frac{1}{1-\sigma}}} \left[\sum_i x_{ij}\right]^{\frac{1}{1-\sigma}}. \tag{B.8}$$

With $\{A_{jj}\}$ in hand, we also infer all $\{A_{ij}\}$ using (B.9):

$$A_{ij} = \frac{1}{L_i^\alpha} \left(\frac{x_{jj}}{x_{ij}} \right)^{\frac{1}{1-\sigma}} \frac{w_i}{w_j} A_{jj} L_j^\alpha. \tag{B.9}$$

We choose the remaining parameter values, $\{\chi_{ij}\}$, $\underline{a}$, β, by repeatedly solving the model to exactly match (i) migration flows in the data, (ii) the steady-state real interest rate of 2%, and (iii) the marginal propensity to consume of 0.3. We normalize $\chi_{ii} = 1$ for all i, since what matters for the migration decision is χ_{ij}/χ_{ii}. In calibrating the migration cost, we use the following updating rule. Given the guess of $\{\chi_{ij}^o ld\}$, we can solve the model to obtain the aggregate migration flows from region i to j:

$$\mu_{ij}^{model} \equiv \int \mu_{ij}(a) d\varphi_i. \tag{B.10}$$

Given the data on migration probabilies in the data, μ_{ij}^{data}, we update χ_{ij} as follows

$$\chi_{ij}^{new} = \xi \frac{\mu_{ij}^{data}}{\mu_{ij}^{model}} \chi_{ij}^{old} + (1-\xi)\chi_{ij}^{old}, \tag{B.11}$$

where $\xi \in (0, 1]$ is the degree of updating. For β and $\underline{a}$, we update using the bisection method.

B.2 Computational Algorithms

We describe the computational algorithm to solve the satus quo economy and the constrained efficient allocation. For both cases, we describe the algorithm for the steady state. The algorithm for the transitions are similar with everything indexed by time t.

B.2.1 Computational Algorithm for Status Quo Economy

In the steady state, households solve

$$v_j(a_t) = \max_{C_j, \{\mu_{jk}\}_k, a' \geq \underline{a}} u_j(C_j) + \beta\omega \left[\sum_k \mu_{jk} v_k(a') - \psi(\{\mu_{jk}\}) \right] \tag{B.12}$$

$$\text{s.t.} \quad P_j C_j + a' = (1+r)a_t + w_j + T_j. \tag{B.13}$$

The first-order condition is

$$u_j'(C_j)/P_j \geq \beta\omega \left[\sum_j \mu_{jk} \partial_a v_k(a') \right] \tag{B.14}$$

with equality whenever $a' > \underline{a}$.

For the inner problem, where we solve the Bellman equation givne prices, we proceed as follows. The algorithm extends the endogenous gridpoint method by Carroll (2006) to incorporate dynamic discrete choices. Let $\mathbb{A} \equiv [a_1, \ldots, a_I]$ denotes the gridpoints in assets.

1. For each grid point in $a' \in \mathbb{A}$, guess $\{v_k(a')\}$.

2. Given $\{v_k(a')\}$, one can compute migration probabilities conditional on saving a' by solving

$$\{\mu_{jk}^{EGM}(a')\}_k \in \arg\max_{\{\mu_{jk}\}} \sum_k \mu_{jk} v_k(a') - \psi(\{\mu_{jk}\}) \tag{B.15}$$

for each j and $a' \in \mathbb{A}$.

3. Assuming the first-order condition holds with equality, invert the consumption using

$$C_j^{EGM}(a') = u_j'^{-1}\left(P_j \beta \omega \left[\sum_j \mu_{jk}^{EGM}(a') \partial_a v_k(a')\right]\right) \tag{B.16}$$

for each j and $a' \in \mathbb{A}$. Then, we are able to obtain the current asset level that is consistent with next period saving a' and non-binding borrowing constraint.

$$a_j^{EGM}(a') = \frac{1}{1+r}\left(a' + P_j C_j^{EGM}(a') - w_j - T_j\right). \tag{B.17}$$

4. For a such that $a \leq a_j^{EGM}(\underline{a})$, the borrowing constraint is binding. Therefore, we recover the saving policy functions as follows.

$$a_j'(a) = \begin{cases} a_j^{EGM,-1}(a) & \text{if } a > a_j^{EGM}(\underline{a}) \\ \underline{a} & \text{if } a \leq a_j^{EGM}(\underline{a}) \end{cases}, \tag{B.18}$$

where $a_j^{EGM,-1}$ denote the inverse function of $a_j^{EGM}(a)$.

5. The migration policies are givne by

$$\mu_{jk}(a) = \mu_{jk}^{EGM}(a_j'(a)). \tag{B.19}$$

and the consumption function is

$$C_j(a) = \frac{1}{P_j}\left((1+r)a + w_j - a_j'(a)\right). \tag{B.20}$$

Now we can update the value function as

$$v_j^{new}(a) = u_j(C_j(a)) + \beta\omega \left[\sum_k \mu_{jk}(a) v_k(a'_j(a)) - \psi(\{\mu_{jk}(a)\}) \right] \tag{B.21}$$

If $|v_j^{new}(a) - v_j(a)| < tol$, we are done. Otherwise, go back to 2 with $v_j(a) = v_j^{new}(a)$.

The outer problem iterates over prices $\{r, \{w_{j,}, P_j\}_j\}$. We divide the outer problem in two layers. In the inner layer, we iterate over r to clear the bond market. In the outer layer, we iterate $\{w_j\}$ to clear the final goods market for each location.

1. Guess $\{w_j, L_j\}$, where we take location 1's wage as numeraire, $w_1 = 1$, and L_j is the population size of location j.

2. Given $\{w_i, L_i\}$, compute the price indices in each location:

$$P_j = \left[\sum_i \left(\frac{w_i}{A_{ij} L_i^{\alpha}} \right)^{1-\sigma} \right]^{1/(1-\sigma)} \tag{B.22}$$

3. Given $\{w_j\}$, iterate over r or β until the bond market clears, $\sum_j \int a d\varphi_j \approx 0$. We use bisection to update r or β. We iterate over r when we solve for the counterfactual. We iterate over β when we calibrate β to match the target interest rate r.

4. We then update wages $\{w_j\}$ and population size $\{L_j\}$ as follows. Given the implied distribution φ_j and consumption policy functions $C_j(a)$ from the guess $\{w_j\}$, we compute

$$w_i^{new} = \xi^w \left[\frac{1}{L_i} \sum_j \frac{(1/(A_{ij} L_i^{\alpha}))^{1-\sigma}}{\sum (w_l/(A_{lj} L_l^{\alpha}))^{1-\sigma}} \int P_j C_j(a) d\varphi_j \right]^{\frac{1}{\sigma}} + (1 - \xi^w) w_j \tag{B.23}$$

$$L_i^{new} = \xi^L \int d\varphi_i + (1 - \xi^L) L_i \tag{B.24}$$

where $\xi \in (0,1]$ is the degree of updating. If $|w_i^{new} - w_i| < tol$ and $|L_i^{new} - L_i| < tol$ for all i, we are done. Otherwise, set $w_i := w_i^{new}$ and $L_i := L_i^{new}$ and go back to step 1.

Practically, the above algorithm finds the equilibrium prices orders of magnitude faster than more conventional algorithms such as Newton's method.

B.2.2 Computational Algorithm for Constrained Efficient Allocation

Throughout, we impose the parametric functional form (56) that we use in the quantitative exercise. We first derive the optimality conditions under the function form of (56), which will be useful for our computation. We then explain how we can efficiently solve the constrained efficient allocation on the computer.

The problem in the steady state is

$$S_i(v) = \max_{\{v_{ij}, C_{it}, \mu_{ij}\}} w_i (1+\alpha_i) - P_i C_i + (1-\omega) S_i^n + \frac{1}{R}\omega \sum_j \mu_{ij} S_j(v_{ij}) \tag{B.25}$$

$$\text{s.t.} \quad v = u_i(C_i) + \beta\omega \left[\sum_j \mu_{ij} v'_{ij} - \psi_i(\{\mu_{ij}\}_j)\right] \tag{B.26}$$

$$\{\mu_{ij}\}_j \in \arg\max_{\{\tilde{\mu}_{ij}\}_j} \sum_j \tilde{\mu}_{ij} v'_{ij} - \psi_i\{\tilde{\mu}_{ij}\}_j) \tag{B.27}$$

and the value of the newborn in the steady state is

$$S_i^n = \max_{v_i^n, \{v_{ij}^n, \mu_{ij}^n\}} \Lambda_i v_i^n + \frac{1}{R} \sum_j \mu_{ij}^n S_j(v_{ij}^n) \tag{B.28}$$

$$\text{s.t.} \qquad v_i^n = \beta \sum_j \left[\mu_{ij} v_{ij}^{n\prime} - \psi_i(\{\mu_{ij}^n\}_j)\right] \tag{B.29}$$

$$\{\mu_{ij}^n\}_j \in \arg\max_{\{\tilde{\mu}_{ij}^n\}} \beta \sum_j \left[\tilde{\mu}_{ij}^n v_{ij}^{n\prime} - \psi_i(\{\tilde{\mu}_{ij}^n\}_j)\right], \tag{B.30}$$

A challenge in numerically solving the above problem is the dimensionality of the control variables. We need to optimize over continuation value for each location. However, we show below that with our functional form assumption (56), the problem essentially collapses to a one-dimensional optimization problem.

We first describe the problem for the incumbent generation (B.25). We solve for C_j to rewrite the problem as

$$S_i(v) = \max_{\{v_{ij}, C_{it}, \mu_{ij}\}} w_i (1+\alpha_i) - P_i C_i + (1-\omega)\Lambda_i S_i^n + \frac{1}{R}\omega \sum_j \mu_{ij} S_j(v_{ij}) \tag{B.31}$$

$$\text{s.t.} \quad C_i = u_i^{-1}\left(v - \beta\omega\left[\sum_j \mu_{ij} v'_{ij} - \psi_i(\{\mu_{ij}\}_j)\right]\right) \tag{B.32}$$

$$\{\mu_{ij}\}_j \in \arg\max_{\{\tilde{\mu}_{ij}\}_j} \sum_j \tilde{\mu}_{ij} v'_{ij} - \psi_i(\{\tilde{\mu}_{ij}\}_j) \tag{B.33}$$

The first-order conditions with respect to v'_{im} are

$$\frac{\omega}{R}\mu_{im}\partial_v S_m(v'_{im}) + \frac{P_j}{u'_i(C_i)}\beta\omega\mu_{im} + \frac{\omega}{R}\sum_k \frac{\partial \mu_{ik}}{\partial v'_m} S_k(v'_k) = 0, \tag{B.34}$$

where we have omitted the dependence for brevity. Under the logit specification (56), we have

$$\frac{\partial \mu_{jk}}{\partial v'_m} = \begin{cases} \theta \mu_{jk} \left(1 - \mu_{jk}\right) & \text{for } k = m \\ -\theta \mu_{jk} \mu_{jm} & \text{for } k \neq m \end{cases} \tag{B.35}$$

Therefore, the FOC simplifies to

$$\partial_v S_m(v'_{im}) + \beta R \frac{P_i}{u'_i(C_i)} + \theta S_m(v'_{im}) - \theta \sum_k \mu_{ik} S_k(v'_{ik}) = 0. \tag{B.36}$$

From this expression, the policy functions must satisfy:

$$\begin{aligned} \partial_v S_m(v'_{im}(v)) + \theta S_m(v'_{im}(v)) &= \partial_v S_n(v'_{in}(v)) + \theta S_m(v'_{in}(n)) && \text{(B.37)} \\ &= \theta \sum_k \mu_{ik} S_k(v'_{ik}(v)) - \beta R \frac{P_i}{u'_i(C_i(v))} && \text{(B.38)} \\ &\equiv M_i(v) && \text{(B.39)} \end{aligned}$$

for all m and n. This observation leads to a substantial simplification. Instead of optimizing over $\{v_{im}(v)\}_m$, we instead optimize over a one-dimensional object $M_i(v)$. Given the guess of $M_i(v)$, we can immediately obtain $v_{im}(v)$ by solving

$$\partial_v S_m(v'_{im}(v)) + \theta S_m(v'_{im}(v)) = M_i(v) \tag{B.40}$$

for each m. Once we obtain $\{v_{im}(v)\}_m$, we can obtain $\{\mu_{im}(v)\}_m$ using (B.33). With $\{v_{im}(v)\}_m$ and $\{\mu_{im}(v)\}_m$ in hand, consumption is residually determined from the promise-keeping constraint (B.32):

$$C_i(v) = u_i^{-1} \left(v - \beta \omega \left[\sum_j \mu_{ij}(v) v'_{ij}(v) - \psi_i(\{\mu_{ij}(v)\}_j) \right] \right). \tag{B.41}$$

Given all the steps for a given guess of $M_i(v)$, we can search for the optimal $M_i(v)$ using the standard one-dimensional optimization routine such as Brent method. In practice, we can obtain further speed gain with the endogenous grid point method by Carroll (2006). Below, we describe the algorithm for value function iteration that relies on the endogenous grid point method.

The newborn's problem (B.28) can be solved similarly, or is even simpler. The first-order condition with respect to $v_{im}^{n\prime}$ is

$$\partial_v S_m(v_{im}^{n'}) + \beta R \Lambda_i + \theta S_m(v_{im}^{n'}) - \theta \sum_k \mu_{ik}^n S_k(v_{ik}^{n'}) = 0, \tag{B.42}$$

which is analogous to (B.36). Therefore, it must be that

$$\partial_v S_m(v_{im}^{n'}) + \theta S_m(v_{im}^{n'}) = M_i^n \tag{B.43}$$

for some M_i^n. For a given guess of M_i^n, we can find the continuation value $v_{im}^{n'}$ that is consistent with M_i^n for all m by inverting (B.43). Given $v_{im}^{n'}$, we can find the migration probabilities using the incentive compatibility constraint (B.30).

Algorithm. We first describe the algorithm for solving the Bellman equation for given vector of $\{w_i, P_i\}_i$. Note that with out functional form assumption (54), α_i is exogenously fixed at α. The outer loop updates $\{w_i, P_i\}$, which we describe later. We let $\mathbb{V} \equiv [v_1, \ldots, v_{N_V}]$ denote the grid point of the promised utility. We let $\mathbb{M}$ denote the grid points for $M_{it}(v) \in \mathbb{M} \equiv [M_1, M_2, \ldots, M_{N_M}]$.

1. Guess the value function $S_{it}(v)$.
2. For each $i = 1, \ldots, J$,
   (a) For each $m \in \mathbb{M}$
      i. Compute $v_{im}^{EGM\prime}(M)$ that is consistent with (B.40) with $M_{it}(v) = M$:

$$\partial_v S_m(v_{im}^{EGM\prime}(M)) + \theta S_m(v_{im}^{EGM\prime}(M)) = M, \tag{B.44}$$

      for each $m = 1, \ldots, J$.
      ii. Using $\{v_{im}^{EGM\prime}(M)\}_m$ obtained from the previous step, compute $\{\mu_{im}^{EGM}(M)\}_m$ using (B.33) associated with $v'_{im} = v_{im}^{EGM\prime}(M)$.
      iii. Find $C_i(M)$ that is consistent with the optimality conditions (B.38) and (B.39):

$$C_i^{EGM}(M) = u_i^{\prime -1}\left(\frac{\beta R P_i}{[\theta \sum_k \mu_{ik}^{EGM}(M) S_k(v_{ik}^{EGM\prime}(M)) - M]}\right). \tag{B.45}$$

      iv. Now we can find the value of today's promised utility v that is consistent with

M using the promise keeping constraint (B.32):

$$v_i^{EGM}(M) = u_i(C_i^{EGM}(M)) + \beta\omega \left[\sum_j \mu_{ij}^{EGM}(M) v_{ij}^{EGM\prime}(M) - \psi_i(\{\mu_{ij}^{EGM}(M)\}_j) \right] \tag{B.46}$$

(b) Now we invert the mapping of $v_i^{EGM}(M)$ to obtain the optimal M for each $v \in \mathbb{V}$: $M_i(v) \equiv v_i^{EGM,-1}(M)$. With $M_i(v)$ for each $v \in \mathbb{V}$ in hand, we can compute all the associate policy functions from the previous step.

(c) For newborn's problem (B.28), we simply optimize over M_i^n and finds associated continuation values $\{v_{im}^{n\prime}\}_m$ using (B.43) to maximize the right hand side of (B.28) to obtain S_{it}^n.

(d) Now we can update the value function:

$$S_i^{new}(v) = w_i\,(1+\alpha_i) - P_i C_i(v) + (1-\omega)\Lambda_i S_i^n + \frac{1}{R}\omega \sum_j \mu_{ij}(v) S_j(v_{ij}'(v)) \tag{B.47}$$

3. If $|S_{it}(v)^{new} - S_{it}(v)| < tol$ for all i and $v \in \mathbb{V}$, the value function has converged. If not, update the value function, $S_{it}(v) := S_{it}(v)^{new}$ and go back to step 2.

The outer loop updates Lagrangian multipliers $\{w_i, P_i\}$. We proceed as follows.

1. Guess $\{w_i, L_i\}$, where L_i is the population size of location i.

2. Given $\{w_i, L_i\}$, compute the implied $\{P_j\}$ with the CES price index:

$$P_j = \left[\sum_i \left(\frac{w_i}{A_{ij}\, L_i^{\alpha}} \right)^{1-\sigma} \right]^{1/(1-\sigma)} \tag{B.48}$$

3. With $\{w_i, P_i\}$ in hand, solve the Bellman equation using the algorithm described above. This gives us the consumption policy function $\{C_j(v)\}$ and the steady-state distribution associated with the policy function, $\{\phi_j\}_j$.

4. We then update wages $\{w_j\}$ and population size $\{L_j\}$ as follows. Given the distribution ϕ_j

and consumption policy functions $C_j(a)$ from the guess $\{w_j\}$, we compute

$$w_i^{new} = \xi^w \left[\frac{1}{L_i} \sum_j \frac{(1/(A_{ij} L_i^{\alpha}))^{1-\sigma}}{\sum (w_l/(A_{lj} L_l^{\alpha}))^{1-\sigma}} \int P_j C_j(v) d\phi_j \right]^{\frac{1}{\sigma}} + (1 - \xi^w) w_j \tag{B.49}$$

$$L_i^{new} = \xi^L \int d\varphi_i + (1 - \xi^L) L_i \tag{B.50}$$

where $\xi \in (0, 1]$ is the degree of updating. If $|w_i^{new} - w_i| < tol$ and $|L_i^{new} - L_i| < tol$ for all i, we are done. Otherwise, set $w_i := w_i^{new}$ and $L_i := L_i^{new}$ and go back to step 2.

B.3 Aggregate Shock

For notational simplicity, we drop x superscript and let t denote the time elapsed since the arrival of the aggregate shock instead. Note that with our assumption of $p \to 0$, the economy is in the deterministic steady state before the arrival of the shock. We let $t = 0$ to denote the deterministic steady state.

The first-order condition with respect to v'_{j1} is

$$p \frac{1}{R_t^1} \frac{dS_{j1}(v_j^{1\prime})}{dv'_{j1}} + p \frac{1}{R_t^1} \sum_k \frac{\partial \mu_{jk}^1}{\partial v_j^{1\prime}} S_{kt+1}(v_k^{1\prime}) + p \frac{P_{jt}}{u'(C_{jt})} \mu_{ji}^1 = 0. \tag{B.51}$$

Given the path of Lagrangian multipliers $\{P_j^x, w_j^x, \alpha_j^x\}_{x=1}^{\infty}$, which we denote in vector format (stacking both j and x), $\{\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}\}$, the first-order condition described above, together with the incentive compatibility and promise-keeping constraints, determines the sequence of policy functions following the shock: $\{v_{ij}^{x\prime}(v), \mu_{ij}^x(v), C_j^x(v)\}_{x=1}^{\infty}$. These policy functions, in turn, determine the evolution of distribution, $\{\phi_j^x\}_{x=1}^{\infty}$. Notice that even when $p \to 0$, the optimality condition (B.51) is well defined, as p cancels out.

Likewise, given the sequence of Lagrangian multipliers, $\{\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}\}$, the following optimality conditions pin down $\{l_{ij}^x, L_j^x\}_{x=1}^{\infty}$:

$$P_i^x \frac{\partial f_i^x}{\partial l_{ki}^x} = w_k^x \tag{B.52}$$

$$\sum_i P_i^x \frac{\partial f_i^x}{\partial L_k^x} = w_k^x \tag{B.53}$$

for all i, k

The sequences of Lagrangian multipliers, $\{\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}\}$, are determined to satisfy the following three conditions:

$$0 = \int C_i^x(v) d\phi_i - f_i^x(\{l_{ki}^x\}_k, \{L_k^x\}_k) \equiv \mathcal{F}^{x,C}(\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}) \tag{B.54}$$

$$0 = \sum_j l_{ij}^x - \int d\phi_i^x \equiv \mathcal{F}^{x,L}(\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}) \tag{B.55}$$

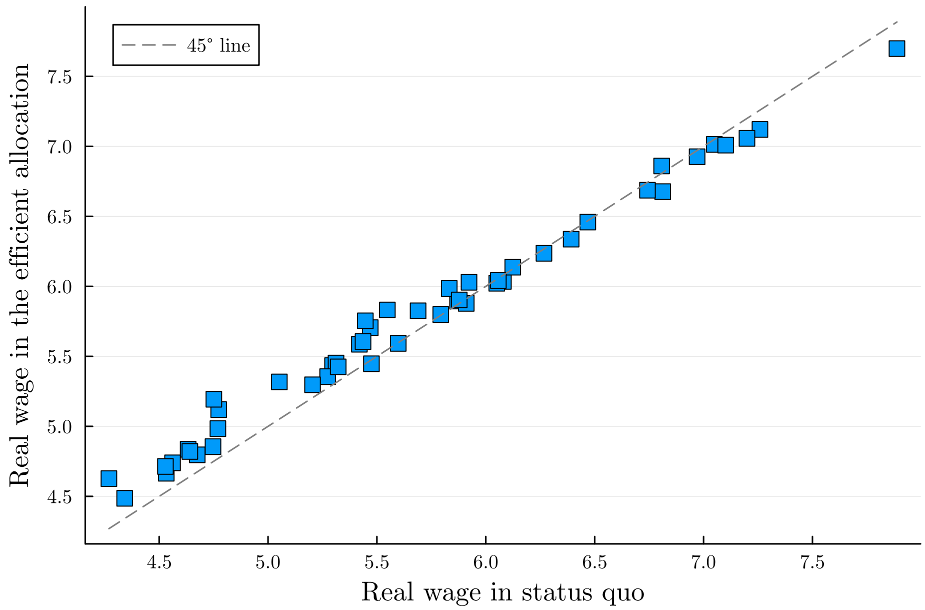
$$0 = L_i^x - \int d\phi_i^x \equiv \mathcal{F}^{x,A}(\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}). \tag{B.56}$$

As described earlier, since policy functions and distributions are functions of the sequence of Lagrangian multipliers $\{\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}\}$, these conditions can be expressed solely as a function of $\{\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}\}$, which we denoted as $\mathcal{F}^{x,C}, \mathcal{F}^{x,L}, \mathcal{F}^{x,A}$ above.

Our approach is to obtain first-order solutions to $\{\boldsymbol{P}, \boldsymbol{w}, \boldsymbol{\alpha}\}$ in sequence space by solving the fixed points described in (B.54)-(B.56). We obtain the sequence space Jacobian (Jacobian of the $\mathcal{F}^{x,C}, \mathcal{F}^{x,L}, \mathcal{F}^{x,A}$) using the methodology in Auclert et al. (2021).

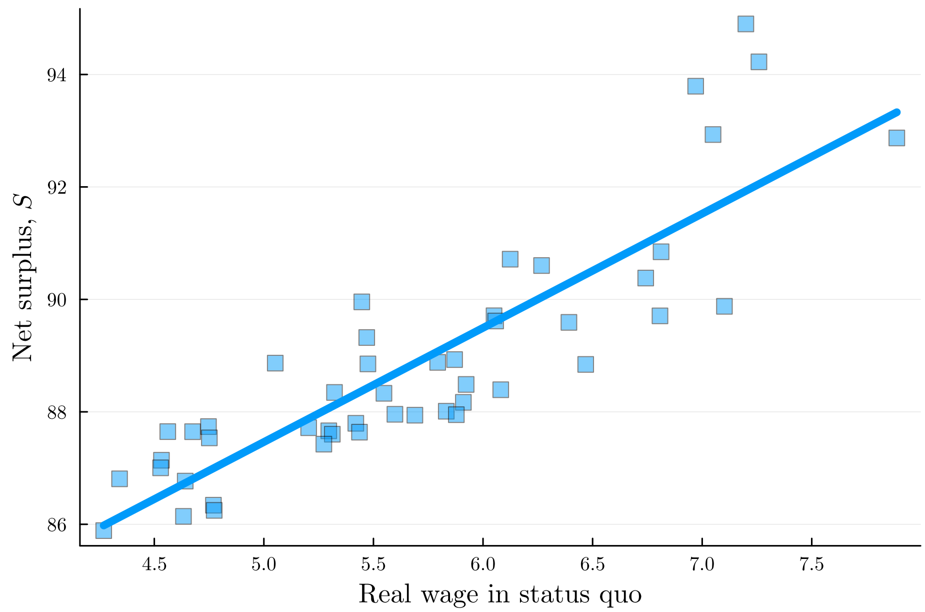
C Additional Figures

Figure C.1: Real Wage: Status Quo vs. Planner



Note: The figure compares the real wage w_j/P_j in the efficient allocation (y axis) and in the status-quo economy (x axis). Each square dot corresponds to a US state. The dashed grey line is a 45 degree line.

Figure C.2: Net Surplus S in the Planner's Solution and Real Wage Status Quo



Note: The figure compares the population weighted average of the net surplus S_j in the planner's solution (y axis) and real wage in the status quo economy (x axis). Each square dot corresponds to a US state. The solid blue line is the best linear fit.