NBER WORKING PAPER SERIES

CONTRASTING THE LOCAL AND NATIONAL DEMOGRAPHIC INCIDENCE OF LOCAL LABOR DEMAND SHOCKS

Richard K. Mansfield

Working Paper 32706
http://www.nber.org/papers/w32706

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2024

Contrasting the Local and National Demographic Incidence of Local Labor Demand Shocks
Richard K. Mansfield
NBER Working Paper No. 32706
July 2024
JEL No. J23,J61,R23,R58

ABSTRACT

This paper examines how spatial frictions that differ among heterogeneous workers and establishments shape the geographic and demographic incidence of alternative local labor demand shocks, with implications for the appropriate level of government at which to fund local economic initiatives. LEHD data featuring millions of job transitions facilitate estimation of a rich two-sided labor market assignment model. The model generates simulated forecasts of many alternative local demand shocks featuring different establishment compositions and local areas. Workers within 10 miles receive only 11.2% (6.6%) of nationwide welfare (employment) short-run gains, with at least 35.9% (62.0%) accruing to out-of-state workers, despite much larger per-worker impacts for the closest workers. Local incidence by demographic category is very sensitive to shock composition, but different shocks produce similar demographic incidence farther from the shock. Furthermore, the remaining heterogeneity in incidence at the state or national level can reverse patterns of heterogeneous demographic impacts at the local level. Overall, the results suggest that reduced-form approaches using distant locations as controls can produce accurate estimates of local shock impacts on local workers, but that the distribution of local impacts badly approximates shocks' statewide or national incidence.

Richard K. Mansfield
Economics Building Room 206C
University of Colorado Boulder
Boulder, CO 80309
and NBER
richard.mansfield@colorado.edu

An online appendix is available at http://www.nber.org/data-appendix/w32706

1 Introduction

Billions of dollars in local aid are spent each year by state, federal, and local agencies to support city-level or county-level economic development initiatives that seek to enhance labor market opportunities for workers who live or work within the local jurisdiction (Bartik (2004)). These often take the form of local infrastructure spending, tax breaks to lure firms to relocate, or discounted loans or subsidies aimed at startup companies. To determine which types of firms or projects to support, federal, state, and local policymakers must predict not only which types of workers from which locations the tax-supported firms would hire, but also whether the resulting ripple effects that operate through vacancy chains and pressure on local wages would primarily trickle down to lower-paid workers or out toward more distant locations. In particular, whether to fund such initiatives at the city, county, or state level depends critically on the shares of the initiative's employment and welfare incidence expected to redound to workers within city, county, and state borders, respectively.

While a large literature in economics seeks to evaluate the incidence of place-based labor demand policies and shocks, most reduced-form methods focus on quite local impacts. More distant towns, counties or states are either excluded from the sample or used as control groups, thereby ignoring the possibility that these more distant areas might collectively account for a sizeable share of shock incidence, even if no single area is strongly affected. Furthermore, due to their focus on policies or shocks occurring in one or a small number of locations, these studies generally feature samples that are too small or too geographically focused to allow comparison of shocks featuring different labor demand compositions on locations featuring different local labor supply compositions, or to examine differential demographic incidence among local and less local areas.

Motivated by this challenge, we adapt the marriage market assignment model of Choo and Siow (2006) to assess and forecast welfare incidence across location-by-demographic group categories from labor demand shocks featuring alternative target areas and establishment compositions. After fitting the model to tens of millions of job transitions and retentions, we perform a variety of simulations that demonstrate how labor market competition interacts with a shock's location and composition to shape its pattern of demographic and spatial incidence.

Two key features of Choo and Siow (2006)'s (hereafter CS) version of the assignment game make it particularly suitable for this analysis. First, it can accommodate multidimensional heterogeneity based on unordered categorical characteristics for agents on both sides of the matching market. This allows the model to feature arbitrary spatial links between workers and establishments in different geographic units (both large and small) that vary flexibly based on combinations of other worker and firm characteristics, such as past income, age, and industry.

Second, the key model parameters, mean relative joint surpluses among matched pairs of workers and positions belonging to observable types, can be mapped one-to-one into odds ratios or revealed comparative advantages that can be constructed from a single labor market matching of workers (with associated initial jobs) to positions. We show that these surplus difference-in-differences among possible job match partners, which we treat as policy-invariant composites of structural parameters, act as sufficient statistics for the job matching technology. Specifically, they allow changes in match outcomes and expected welfare for both workers and firms to be computed for any counterfactual change in labor supply and/or demand composition. Importantly, the sufficient statistics approach does not require the specification of a more fundamental structural model of utility, firm production, and moving or search costs. Thus, heterogeneity in observed matching patterns is not lost or mischaracterized in projecting the data onto a small number of interpretable structural parameters that reflect the authors' beliefs about the sources of comparative advantages.

We estimate the model using matched employer-employee data from the Longitudinal Employer-Household Dynamics (LEHD) database on 19 U.S. states that approved the use of their records. The data display three key properties that make it suitable for a rich assignment model: 1) they capture the (near) universe of job matches from the participating states, mitigating selection problems; 2) they include tens of millions of annual job matches, allowing precise estimates of the large number of parameters necessary to capture complex two-sided multidimensional sorting; and 3) workers' establishments are geocoded to the census tract level. These properties ensure that the data, when combined with the model, provide the necessary inputs to compute the shares of employment and welfare gains or losses from alternative local labor demand shocks that accrue to particular demographic groups located within particular jurisdictions both near and far from the shock.

The counterfactual simulations we consider involve establishment openings and closings that create or destroy 250 positions in particular U.S. census tracts featuring alternative combinations of establishment size, average pay, and supersector. While we model the diffusion of labor demand shocks via labor market competition much more richly than other structural models, we do not model the housing and product markets, though the estimated surplus parameters partly capture their impact through the way they affect worker flows. Thus, we recover "labor-related" welfare changes induced by these shocks that act as complementary inputs to local policy decisions alongside estimates of house and product price elasticities.[1] These simulations yield five primary findings.

First, across a wide variety of simulated shocks, we show that job stimuli generate very small per-worker impacts on employment probability and expected welfare for workers outside the targeted area. Averaging across simulations, we find that utility and employment gains (in parentheses) for initially local workers are 3.1 (2.8), 18 (19), and 2,850 (857) times as large as for workers in an adjacent tract, an adjacent PUMA[2], and a non-adjacent state, respectively, with expected utility gains (scaled in $ of annual earnings) of $322 for focal tract workers and just $0.11 for the most distant workers. Such rapid declines and tiny mean impacts for far away workers confirm that local labor markets are sufficiently isolated to allow accurate reduced-form estimates of treatment effects of local demand shocks on local workers when distant locations serve as control groups.

[1]For example, policymakers might wish to know whether a local initiative creating new high-skilled positions will create sufficient downstream earnings opportunities for low-income renters to offset any increases in rent.

[2]PUMAs or "public-use microdata areas" are mutually exclusive and exhaustive collections of contiguous counties and census tracts encompassing at least 100,000 residents. We focus on PUMAs instead of Commuting Zones because they better approximate the level at which decisions are made for many small, local economic initiatives. In particular, PUMAs are smaller and more consistent in their population, with 2,378 PUMAs nationally vs. 741 CZs), they do not cross state lines, and they are more likely to distinguish large suburbs of major cities from the city centers.

Second, despite very disproportionate per-capita gains for the most local workers, the cumulative share of welfare and particularly employment gains accruing to workers outside of the targeted local area can be quite large, since the local workforce makes up a very small share of the national labor market.[3] We find that only 9.9% and 5.8% of the job-related utility and net employment gains from such local stimuli accrue to workers initially or most recently working in the surrounding PUMA, while 35.9% and 62.0% of the utility and employment gains accrue to workers initially outside the state. This result, which most reduced-form approaches cannot capture, casts doubt on labor market-based justifications for funding local initiatives at the local level. Requiring newly-created jobs to be filled exclusively by local PUMA workers (or job seekers) does increase their share of net employment gains, but only to 17.6%. We also find that the within-PUMA share of employment gains is over twice as high for shocks targeting rural rather than urban areas (8.7% vs. 3.7%).

Third, we show that the degree to which local job creation improves local income inequality depends critically on which types of jobs are brought to town. The primary beneficiaries of local shocks vary widely with the establishment composition of the newly created jobs, suggesting opportunities for local officials to craft local development initiatives that target particular local sub-populations. For example, unemployed workers from the target tract reap the largest welfare gains ($620) from positions created at small, low paying firms in the other services sector and the smallest gains ($249) from positions at small, high paying professional & business services firms. By contrast, initially employed but low-paid workers benefit the least ($167) from positions created at large, high paying information firms and the most ($573) from positions at large, low-paying education & health firms, and the highest-paid workers benefit the least ($154) from positions at large, low paying information firms and the most ($641) from positions at large, high-paying education & health firms. Incorporating existing estimates of job multipliers only slightly alters these findings.

Fourth, despite the high degree of heterogeneity in local incidence, we show that different shocks become increasingly generic in their demographic incidence as one focuses on more distant workers: initially low-paid or unemployed workers enjoy only 1-2% higher shares of nationwide employment and welfare gains when job creation occurs at low-paying firms rather than high-paying firms. This occurs even though workers in the top initial earnings quartile take under 5% of newly created jobs at low-paying firms vs. over 26% at high paying firms. Furthermore, focal tract characteristics that predict relatively greater employment gains for local low-paid and unemployed workers from local job creation fail to predict any such employment redistribution at the national level.

Fifth, the remaining heterogeneity in incidence at the state or national level can reverse findings at the local level. For example, older initially unemployed workers generally enjoy a larger share of shock-induced local employment gains than their local workforce share, since their lower geographic mobility causes lower job-finding rates without the shock. However, at the national level, it is younger unemployed workers who reap disproportionate employment gains, as they are more willing to move to new job opportunities. Similarly, workers from the same industry as the new store/plant generally enjoy a much larger share of shock-induced local welfare gains than their local

[3] A single census tract generally only contains a few thousand workers, and a single PUMA contains around 100,000.

workforce share (since they are good fits for the new jobs), but they account for a nearly proportionate share of national gains. This is partly because most jobs vacated by those taking the new local jobs are in other industries, but also because they highly value their existing stable jobs, making them insensitive to distant opportunities. These findings suggests that reduced-form estimates of local treatment effect heterogeneity may be a particularly poor guide to shock incidence at more aggregate levels. Thus, at the state-level, officials may wish to prioritize job creation per dollar of funding over equity considerations when choosing local projects to fund.

Finally, we also perform a validation exercise in which the estimated model is used to forecast the realized reallocation around 421 census tracts that experienced openings or closings of more than 100 jobs within one year between 2003 and 2012. The model predicts these out-of-sample reallocations well and considerably better than relatively rich one-sided parametric models that fit firm or worker conditional choice probabilities with over 100 parameters. This exercise illustrates that the very large set of estimated parameters is not causing overfitting, but is instead necessary to capture the highly nonlinear and multidimensional nature of the U.S. job matching technology.

This paper builds primarily on three literatures. The first consists of evaluations of particular place-based policies or local economic shocks. Most papers in this branch use average wages or employment rates in the targeted location as the outcome of interest, seek to define a control group of alternative locations, and evaluate the policy or shock's impact using a treatment effect framework. This literature is vast, and is thoroughly discussed by survey articles such as Glaeser et al. (2008), Moretti (2010), Kline and Moretti (2013), and Neumark and Simpson (2015).[4]

Two recent papers in this vein are notable for incorporating spillovers to non-targeted locations driven by worker mobility. Sprung-Keyser et al. (2022) analyze mobility and earnings responses across workers' initial commuting zones of CZ-level wage variation induced by differential recovery of labor demand after the Great Recession. Like this paper, they show that substantial heterogeneity in geographic mobility by demographic group can affect incidence of local demand shocks. However, their framework does not allow initial mobility responses to change wage offers in other locations via vacancy chains and labor supply outflows. Like us, Hornbeck and Moretti (2024), analyzing incidence of decadal city-level manufacturing TFP growth, show that such wage offer changes can lead other locations to account for large shares of national earnings and employment gains. They also find that greater local gains for more educated workers are offset at the national level. We show that such reversals of incidence at more aggregate geographies are likely to be common, but that they are specific to the firm composition of the labor demand shock. Because they examine cross-sectional changes in city outcomes, they cannot distinguish mobility-induced compositional changes from actual longitudinal outcome changes among workers by initial location. More generally, both papers analyze long-run outcomes from differential CZ or MSA exposure to national shocks rather than short-run responses to small, hyper-local job creation.

[4]A particularly prominent example is Greenstone et al. (2010), who compare employment gains in counties making winning bids for "million-dollar" plants to control counties who made losing bids. Busso et al. (2013) is one of the few quasi-experimental papers to use their elasticity estimates to explicitly evaluate social welfare impact.

Our approach also complements a sub-literature on local job multipliers from increased product demand and agglomeration/congestion externalities created by a local job stimulus (e.g. Moretti (2010) or Bartik and Sotherland (2019)). Such papers generally do not assess which types of workers from which initial locations benefit from the net change in local job opportunities, while our assignment model takes as an input the new spatial distribution of positions (possibly reflecting multipliers) and evaluates the resulting skill and spatial incidence. We demonstrate this point by evaluating a shock combining 250 new manufacturing positions with 171 service jobs spread throughout the PUMA in accordance with the relevant multiplier estimate from Bartik and Sotherland (2019).

Second, the paper adds to a fast-growing literature on structural spatial equilibrium models designed to forecast the georaphic incidence of economic shocks. Several such models impose additional structure on the sources of match surpluses or incorporate additional markets, while our model offers a richer and more flexible labor market. For example, Schmutz and Sidibé (2019) use a search-and-matching model to separate the roles of search frictions and moving costs in determining the incidence of local shocks. Monte et al. (2018) and Caliendo et al. (2019) each estimate trade-theoretic models with labor, housing, and product market clearing and arbitrary spatial frictions in both labor and product markets. The former features joint choices of residential and work locations, and highlights the role of commuting in determining local shock incidence.[5] The latter shows how to evaluate counterfactual dynamic equilibrium paths for alternative structural shocks without estimating all the model primitives. Our paper focuses on short-run (one-year) predictions that are unlikely to be sensitive to longer-run housing and product market dynamics. We demonstrate robustness to unmodeled shock-induced changes in dynamic continuation values by simulating shocks that incorporate observed average surplus changes from actual establishment openings.

Each of these papers aggregates locations to at least the county level. Manning and Petrongolo (2017), by contrast, use a search-and-matching model to fit data on changes in vacancy stocks from local job search centers in Britain, and simulate the impact on the geographic distribution of unemployment of an exogenous increase in vacancies (new jobs) within particular census wards (analogous to census tracts). Like Marinescu and Rathelot (2018), they find evidence that labor markets are quite local, in the sense that moderate distance to vacancies substantially decreases the probability of an application. Nonetheless, like us, they find that ripple effects from overlapping markets cause very little of the employment gain to accrue to the targeted ward.

None of these papers feature any worker heterogeneity beyond initial location, and only Caliendo et al. (2019) (industry differences) features any observable firm heterogeneity besides location. Similarly, several spatial labor market models, such as Piyapromdee (2021) or Diamond (2016), feature imperfect substitution among observable worker types, but only differentiate firms by location. Because none of these models feature multidimensional two-sided sorting, the model featured in this paper is the only one equipped to evaluate differential incidence both across space and across skill/demographic groups from local labor demand shocks with alternative firm compositions. While Lindenlaub (2017) and Bonhomme et al. (2019) each estimate multidimensional labor

[5] Due to a lack of residential microdata, we do not consider whether new job matches involve residential mobility.

market sorting models, they do not incorporate geography or spatial frictions.

Indeed, Nimczik (2018), who characterizes labor markets as networks of firms disproportionately sharing worker flows, shows that the geographic and industrial scope of labor markets varies substantially across occupation and education categories. However, his stochastic block model defines distinct labor markets for each skill category. Thus, it is not designed to analyze the tradeoffs firms and workers make following local demand shocks between settling for skill mismatch and paying moving and search costs to overcome spatial mismatch. Fogel and Modenesi (2021) use a similar "revealed network" approach and do allow substitution across revealed worker types and markets, but focus on only Rio De Janeiro, precluding analysis of skill/spatial tradeoffs. More generally, this paper also builds on the reduced-form and descriptive literature capturing how worker mobility and the geographic extent of labor markets vary by worker and firm characteristics.[6]

Finally, this paper draws heavily from the theoretical literature on two-sided assignment games. Several early papers established properties of assignment equilibria (Koopmans and Beckmann (1957), Shapley and Shubik (1972), Roth and Sotomayor (1992), and Sattinger (1993)), with a more recent literature examining identification and estimation (Choo and Siow (2006), Chiappori and Salanié (2016), Menzel (2015), Galichon and Salanié (2022), and Mourifié and Siow (2021)). To our knowledge this is the first large-scale empirical application of a two-sided assignment model to the national labor market.[7] We make three contributions to this literature.

First, we consider implementation with millions of match observations and thousands of types on both the supply and demand side. We address the "granularity" problem of a somewhat sparse matching matrix highlighted by Dingel and Tintelnot (2020) by developing a smoothing procedure to aggregate matching patterns across "nearby" match types without removing the heterogeneity the model is designed to highlight. Second, we allow separate surplus values for job stayers relative to within-job-type movers, and show that this reveals asymmetry between the welfare losses and gains from negative and positive demand shocks. Third, because unfilled vacancy counts by detailed type are not available, we consider the limits to identification when the number of unmatched partners of each type is either unobserved or only observed on one side of the market.[8] We discuss conditions under which model predictions are invariant to ignoring unmatched partners, and show robustness of results to endogenizing the set of positions to be filled via a fixed point algorithm.

The rest of the paper proceeds as follows. Section 2 describes our two-sided assignment game. Section 3 establishes identification of a set of joint surplus parameters that fully determine the labor market incidence of counterfactual shocks. Section 4 describes the LEHD database and presents summary statistics that motivate the subsequent analysis. Section 5 describes the smoothing procedure and introduces the various labor demand shocks and the methods used to aggregate counterfactual job matchings into interpretable statistics that highlight variation in shock incidence. Section 6 presents the main findings and the model validation results, and Section 7 concludes.

[6] e.g. Malamud and Wozniak (2012), Cadena and Kovak (2016), Bayer et al. (2008).

[7] See Tervio (2008) and Chen (2017) for applications of the assignment game to the narrower market for CEOs.

[8] Existing identification results (e.g. CS and Menzel (2015)) rely on observing the number of singles on both sides.

2 The Two-Sided Assignment Model

We model the labor market as a static assignment game played by workers and establishments. We introduce several features and extensions necessary to adapt CS's marriage market model to a labor market setting. The exposition closely mirrors Galichon and Salanié (2022) (hereafter GS), which generalizes CS. Section 2.1 defines the matching game and describes how workers and positions and the job matches that determine the game's payoffs are aggregated to types and groups, respectively. Section 3.1 imposes further structure to facilitate the identification and estimation of the underlying group-level match surpluses that govern the frequencies of different kinds of job matches. Section 3.2 shows how to use the estimated surpluses to construct counterfactual simulations capturing the incidence of local labor supply and demand shocks of varying worker and position compositions.

2.1 Defining the Assignment Game

Suppose that in a given year there are I potential workers comprising the set $\mathcal{I}$ who participate in the labor market. Each worker i enters the market with an existing job match with a position $j(i)$ at establishment $m(j(i))$ taken from the set of possible positions $\mathcal{J}$. Let $m(j) = 0$ represent unemployment so that positing an initial "job" for each worker is without loss of generality. Each worker i also belongs to an observed worker type $l(i) \in \mathcal{L}$. In the empirical work, worker types (detailed further below) are defined by combinations of 1) an age category, 2) an age-adjusted prior earnings/unemployment category, 3) an indicator for whether the newly-created local jobs are in the worker's initial industry, and 4) the location of a worker's previous establishment.[9]

On the other side of the market there are K potential positions comprising set $\mathcal{K}$ at establishments that seek workers in the chosen year. The intersection of $\mathcal{K}$ and $\mathcal{J}$ may be quite large, so that many positions in $\mathcal{K}$ can potentially be "filled" by retaining an existing worker. We assume each establishment makes independent hiring decisions for each position so as to model positions' preferences over individual workers rather than establishments' preferences over collections of workers.[10] Each position $k \in \mathcal{K}$ belongs to a position type $f(k) \in \mathcal{F}$. Below, these types will consist of combinations of an employer's size category, average pay category, industry supersector, and location.[11]

Each potential job match (i, k) can be assigned to one of a set of mutually exclusive groups $g \in \mathcal{G}$ (with $G \equiv |\mathcal{G}|$). Let $g(i, k)$ denote match (i, k)'s group assignment, and let $z(i, j(i), k) \equiv z(i, k)$ denote any characteristics defining the match group that depend on both i and k. In the empirical work, the only z characteristic is a trichotomous indicator that equals one for continued

[9] Ideally, residential location would define the worker type and establishment location would define the position type. In the absence of data on workers' residential locations, initial (i.e. past) establishment locations are used as proxies.

[10] One justification for treating positions as independent is that there are nontrivial costs of coordinating multiple independent hires/retentions that outweigh the gains from better exploiting production complementarities. Roth and Sotomayor (1992) highlight the complications that arise when establishments have preferences over collections of workers.

[11] There is no inconsistency in using i's age-adjusted prior year earnings to proxy for worker skill and using $m(k)$'s prior year average pay to proxy for k's skill requirements, since earnings have been widely shown to contain persistent worker and firm components, and one can assume that a new hire develops the required skills by the end of the year.

employment at the same establishment, two for employer changes within the same supersector, and zero otherwise.[12] Then one can rewrite the mapping $g(i,k)$ as $g(l(i), f(k), z(i,k)) \equiv g(l,f,z)$. We use $l(g)$ to refer to group g's worker type and $f(g)$ to refer to its position type.

Worker i's payoff from accepting position k in the current year is denoted $U(i,k)$. The worker's potential earnings at the position in the chosen year, denoted w_{ik}, is assumed to be additively separable from other determinants of the worker's payoff, so that $U(i,k)$ has a money-metric form:[13]

$$U(i,k) = \pi^i_{ik} + w_{ik} \equiv \theta^l(g(i,k)) + \epsilon^i_{ik} + w_{ik} \tag{1}$$

$\pi^i_{ik} \equiv \theta^l(g(i,k)) + \epsilon^i_{ik}$ captures the combined value to worker i of a variety of payoff components. We show below that one need not specify any of the fundamental components or the functions governing their links to payoffs to construct counterfactual simulations capturing labor demand shock incidence. Any payoff function in which current worker earnings are additively separable will suffice. That said, careful thought about which determinants of the payoff are likely to be large and differential across alternative workers, positions, and job matches is necessary to guide the choice of characteristics used to assign workers and positions to types in section 4.2 below, as well as to evaluate the plausibility of assumptions laid out in section 3.2 that underlie the simulations.

Such components include worker i's valuation of various non-pecuniary amenities offered by position k (including the appeal of its location). In addition, though assignment games traditionally have been characterized and parameterized as "frictionless" models, the model's structure does not preclude either deterministic moving or training costs among worker-firm pairs or stochastic search frictions.[14] Thus, π^i_{ik} also captures any search, moving, or training costs paid by worker i to find, move to, or settle into position k from initial position j. While the model is not explicitly dynamic, in practice π^i_{ik} might also include the continuation value associated with starting the next year as a trained worker at position k, which might depend on productivity gains from firm-specific experience and the availability of other jobs in position k's local labor market.[15]

$\theta^l(g)$ captures the part of position k's value to worker i that is common to any worker of type $l(i)$ accepting a position of type $f(k)$ with match characteristics $z(i,k)$. For example, older workers may particularly value jobs in industries with less physically taxing tasks, low-paid workers may particularly value large firms with a well-defined promotion path, and all workers may value avoiding search and training costs by staying at their current establishment. ϵ^i_{ik} captures the part of k's value to i that is idiosyncratic to (i,k) conditional on (l,f,z). ϵ^i_{ik} might reflect, for example, the low psychic costs of a worker who is moving back to a familiar location.

[12] Mourifié and Siow (2021) use the same approach to distinguish marriage from cohabitation.

[13] Since we have data on annual earnings but not wages or hours, for simplicity we assume that the hours associated with a job match are fixed by contract and common across positions for a given worker.

[14] For example, Menzel (2015) shows that one can augment a deterministic assignment model with a probability that i and k meet that is independent of other payoff determinants, assign their joint surplus to $-\infty$ if the pair does not meet, and use these alternative payoffs to determine the stable matching.

[15] Mourifié (2019) shows that an augmented version of CS's static assignment model and Choo (2015)'s dynamic assignment model generate identical surplus estimates and matching functions, suggesting that static and dynamic models are likely to generate similar incidence predictions, particularly for short-run impacts from small shocks.

Let $V(i,k)$ denote the value to position k in establishment $m(k)$ of hiring (or retaining) worker i. The potential earnings paid by k to worker i in the chosen year is assumed to be additively separable from all other determinants of the position's payoff, so that $V(i,k)$ can be written as:

$$V(i,k) = \pi_{ik}^k - w_{ik} \equiv \theta^f(g) + \epsilon_{ik}^k - w_{ik} \tag{2}$$

Akin to π_{ik}^i, π_{ik}^k combines several payoff components that need not be fully specified. These components might include worker i's contribution to $m(k)$'s annual revenue, any recruiting, moving, and training costs borne by $m(k)$ in hiring worker i, and any continuation value from starting next year's market with i already in position k, such as the option to avoid further recruiting/training costs next year. As with π_{ik}^i, π_{ik}^k features a common group-level component $\theta^f(g)$ and an idiosyncratic component $\epsilon_{i,k}^k$. $\theta^f(g)$ might capture the possibility that larger firms may face smaller per-position costs of recruiting distant workers due to economies of scale, or that highly skilled workers may generate more revenue at high-paying firms whose output is particularly sensitive to worker skill. ϵ_{ik}^k might capture particular skills required by position k that worker i uniquely possesses.

We define the joint surplus from i and k's match as the combined worker and position valuations:

$$\pi_{ik} \equiv U(i,k) + V(i,k) = \pi_{ik}^i + \pi_{ik}^k \tag{3}$$

Since worker earnings are additively separable in both worker and position payoffs, the model exhibits transferable utility, mimicking Shapley and Shubik (1972)'s classic assignment game.

A matching or market-wide allocation in this labor market is an $I \times K$ matrix μ such that $\mu_{i,k}$ = 1 if worker i matches with position k, and 0 otherwise. As in GS, we focus on stable matchings, which require a division of joint surplus in each job match such that no currently unmatched worker-position pair can find any division of the joint surplus from their potential match that makes both strictly better off than under the proposed matching. Shapley and Shubik (1972) show that the set of stable matchings coincides with both the core of the game and the set of competitive equilibria from a decentralized market. Furthermore, they show that with transferable utility there will exist a unique matching (or, equivalently, competitive equilibrium allocation) of workers to positions as long as preferences are strict on both sides of the market. This equilibrium allocation/stable matching maximizes the aggregate surplus and solves a linear programming problem.[16]

Two key features of the equilibrium assignment should be noted. First, the stable assignment is fully determined by the joint surplus values $\{\pi_{ik}\}$ (See Appendix A1); no separate information on the worker and firm components π_{ik}^i and π_{ik}^k is needed. This implies that one need not impose additional assumptions to separately identify the amenity, productivity, and training/search costs components of the surplus in order to assess the incidence of local labor demand shocks.

Second, while market-clearing earnings amounts will in general be specific to worker-position

[16]Aggregate surplus is given by $\sum_{(i,k)\in\mathcal{I}\times\mathcal{K}} \mu_{i,k}\pi_{ik} + \sum_{i\in\mathcal{I}:\mu_{i,k}=0\ \forall\ k} \mu_{i,0}\pi_{i0} + \sum_{k\in\mathcal{K}:\mu_{i,k}=0\ \forall\ i} \mu_{0,k}\pi_{0k}$, where π_{i0} and π_{0k} denote i's payoff from unemployment and k's payoff from remaining vacant. Also, each position and worker is constrained to match with at most one counterpart: $\sum_i \mu_{i,k} \leq 1\ \forall\ k \in \mathcal{K}$ and $\sum_k \mu_{i,k} \leq 1\ \forall\ i \in \mathcal{I}$.

pairs (i,k), the market-clearing utilities r_i and profit contributions q_k (i.e. the game's payoffs) will be worker-specific and position-specific, respectively (they solve the dual version of the social planner's linear programming problem). We exploit this property below. Importantly, while the stable assignment μ is generally unique, the equilibrium payoffs and transfers are not: all r_i utility values can generally be shifted slightly up or down (with offsetting q_k shifts) without violating stability. The exact equilibrium payoffs/earnings depend on the market clearing mechanism.

While the model does not require a particular earnings-setting process, one candidate is a simultaneous ascending auction in which all positions bid on all workers. Workers set reservation utilities based on their values of remaining unemployed for a year. Each position bids utility values of a one year commitment U_{ik} (which include the value of starting the next year at k), and may only win the bidding for a single worker (or choose to remain vacant). The position k that bids the highest utility r_i retains worker i and pays annual earnings w_{ik} that, combined with the non-pecuniary component π^i_{ik}, equals the worker's promised valuation $U_{ik} = r_i$. The auction ends when no position wishes to change its bid for any worker. Some workers may remain unemployed and some positions may remain unfilled. Importantly, though positions start at different π^i_{ik} baselines, with transferable utility bid changes can always take the form of earnings increases. Thus, changes in equilibrium utilities r_i following demand shocks can be scaled in terms of annual earnings gains (though some utility gains are achieved by taking an earnings cut to get a position offering superior non-pecuniary values).

Since we wish to examine the incidence of local labor demand shocks with different position type composition across worker types rather than predict exact worker-position matches, we follow CS in analyzing group-level equilibria that are consistent with the underlying worker-position-level equilibria. To this end, we decompose π_{ik} into group-level and idiosyncratic components as follows:

$$\pi_{ik} = \theta_g + \sigma\epsilon_{ik} \tag{4}$$

where $\theta_g = \theta^l(g) + \theta^f(g)$ and $\epsilon_{ik} = \frac{\epsilon^i_{ik}+\epsilon^k_{(ik)}}{\sigma}$. σ is a scaling parameter that captures the relative importance of idiosyncratic surplus components compared to group-level components in determining the variation in match surpluses across potential pairs $(i,k) \in \mathcal{I} \times \mathcal{K}$. We show below that counterfactual job assignments do not depend on σ, but σ governs the marketwide scale of changes in utility bids r_i necessary to facilitate the reallocation that yields the stable assignment.

The goal is to use the observed matching μ to recover the set of group mean surplus values $\{\theta_g\}$.[17] We achieve identification by assuming that ϵ_{ik} draws are i.i.d across all alternative matches (i,k') and $(i',k) \in \mathcal{I} \times \mathcal{K}$ and follow a Type 1 extreme value distribution.[18] Sprung-Keyser et al. (2022) provide quasi-experimental support for a key property of i.i.d EV models: origin-specific

[17]GS show that one could impose further structure on the production, utility, search cost, and recruiting cost functions that comprise the joint surpluses and estimate the model via ML. We prefer to be agnostic about these structural functions, so we follow CS and leave $\{\theta_g\}$ unrestricted, even allowing job/industry staying premia to vary by (l,f) pair.

[18]Menzel (2015) shows that imposing i.i.d draws is the key assumption rather than the Type 1 EV distribution). GS and Chiappori et al. (2009) show how to allow certain forms of correlation in the idiosyncratic component across matches with shared characteristics. However, we use the standard i.i.d. assumption to ease our substantial computational burden.

changes in mobility rates to a location experiencing a labor demand shock increase monotonically with their baseline transition rates to the target location. Unlike in CS and GS, equation (4) allows the idiosyncratic component to be truly pair-specific: the combined surplus from two matches changes if the workers swap positions, even if they share a worker type and the positions share a position type. Gutierrez (2020) shows that allowing idiosyncratic components to be pair-specific ensures that the model does not suffer from the independence of irrelevant alternatives (IIA) problem, so that subdividing worker or position types in arbitrary ways does not change estimated joint surplus values or match probabilities. Sections 6.6 and 6.7 compare simulated forecasts and out-of-sample prediction accuracy between our model and the CS model. As discussed in Section 3.2 and Appendix A5, allowing such heterogeneity prevents the use of observed transfers to recover group-level worker and position subcomponents $\theta^l(g)$ and $\theta^f(g)$ defined above. Fortunately, this decomposition is not necessary to generate key measures of worker-level incidence.

3 Identification

3.1 Identification of the Set of Group-Level Match Surpluses $\{\theta_g\}$

Shapley and Shubik (1972) show that a necessary condition for a matching μ to be sustainable as a competitive equilibrium is that there exists a set of worker payoffs $\{r_i\}$ such that $\mu_{ik} = 1$ implies that $i \in \arg\max_{i \in \mathcal{I}} \pi_{ik} - r_i$ for any potential match $(i, k) \in \mathcal{I} \times \mathcal{K}$. Combining this result with the i.i.d. Type 1 EV assumption for ϵ_{ik} yields a standard logit expression for the probability that worker i maximizes k's payoff (Appendix A1). We wish to aggregate this logit formula to the group level.

Define $n(l)$ as the share of workers assigned to type l, define C_l as the mean of $e^{-\frac{r_i}{\sigma}}$ among type l workers, and define $\overline{S}_{g|l,f}$ as the mean among type f positions of the share of type l workers whose hire/retention would be assigned to group g. This is the share from the same firm if $z(g) = 1$, the same industry share if $z(g) = 2$, and the share from other industries if $z(g) = 0$. With two additional assumptions, Appendix A1 derives a tractable expression for the conditional probability $P(g|f)$ that a position of type f wishes to hire a type l worker whose job match would be assigned to group g:

$$P(g|f) = \frac{e^{\frac{\theta_g}{\sigma}} \overline{S}_{g|l,f} n(l) C_l}{\sum_{l' \in \mathcal{L}} \sum_{g' \in (l,f)} e^{\frac{\theta_{g'}}{\sigma}} \overline{S}_{g'|l',f} n(l') C_{l'}} \tag{5}$$

This expression depends only on the group g and the types l and f rather than individual workers i and positions k.[19] Appendix A1 presents and proves this result formally as Proposition A1. Intuitively, the first assumption imposes that the utility payoffs required in equilibrium by workers from the same initial earnings, age, and industry categories and local area must not differ systematically across initial establishments. This becomes a better approximation as worker types are defined by

[19] Note that in contrast to CS, the probability that a type l worker is chosen depends on the share of workers of type l in the population, $n(l)$. This difference stems from allowing the idiosyncratic surplus component to be pair-specific. Menzel (2015) derives a similar formula in his nontransferable utility assignment model.

more categories and finer geography, so that workers of the same type become close substitutes for one another. The second assumption imposes that establishments of the same position type feature roughly the same number and worker type distribution of incumbent workers. This approximation improves as position types are defined by narrower establishment location, industry, average pay, and particularly size categories. Importantly, these additional assumptions are only necessary to isolate the surplus from hiring a within-firm incumbent relative to a worker from another firm in the same census tract. Violations (discussed in Appendix A1) lead to slight over or understatements of deviations among (l, f) type combinations from the average surplus premium for job staying.

Next, let $\hat{\mu}$ denote an observed matching. Since each job match can be assigned to a unique group g, one can easily aggregate the individual-level matching into an empirical group-level distribution. Let $\hat{P}_g$ denote the fraction of observed matches that are assigned to group g, let $\hat{n}(l)$ denote the fraction of matches featuring type l workers and $\hat{h}(f)$ denote the fraction featuring type f positions.[20] One can then estimate the conditional choice probability $P(g|f)$ by calculating the observed fraction of type f positions that were filled via group g matches: $\hat{P}(g|f) = \frac{\hat{P}_g}{\hat{h}(f)}$. As the number of observed matches gets large, each member of the set of empirical CCPs $\{\hat{P}(g|f)\}$ should converge to the corresponding expression in (5). The average shares $\{\overline{S}_{g|l,f}\}$ can also be estimated using averages of the incumbent indicator $1(m(j(i)) = m(k))$ and same supersector indicator $1(s(j(i)) = s(k))$ across all possible matches (i, k) sharing type pairs $(l(i), f(k))$.

One may now assess the amount of information contained in the observed empirical choice probabilities $\{\hat{P}(g|f)\}$ about the mean match surplus values $\{\theta_g\}$. First, using (5), we can derive an expression for the log odds between two CCPs involving an (arbitrarily chosen) position type f_1 and two (arbitrarily chosen) match groups g_1 and g_2 for which $f(g_1) = f(g_2) = f_1$:

$$\ln(\frac{\hat{P}_{g_1|f_1}}{\hat{P}_{g_2|f_1}}) = (\frac{\theta_{g_1} - \theta_{g_2}}{\sigma}) + \ln(\frac{\overline{S}_{g_1|l(g_1),f_1}}{\overline{S}_{g_2|l(g_2),f_1}}) + \ln(\frac{n(l(g_1))}{n(l(g_2))}) + \ln(\frac{C_{l(g_1)}}{C_{l(g_2)}}) \tag{6}$$

Since the worker type shares $n(l(g_1))$ and $n(l(g_2))$ and shares of potential firm or industry stayers $\overline{S}_{g_1|l(g_1),f_1}$ and $\overline{S}_{g_2|l(g_2),f_1}$ are either directly estimable or observed (depending on whether a sample or the full population is available), to establish identification one can treat their terms as known and bring them to the left hand side. Even these adjusted log odds still conflate the relative mean (re-scaled) surplus values from match groups g_1 and g_2, $(\frac{\theta_{g_1}-\theta_{g_2}}{\sigma})$, with the log ratio of mean exponentiated worker re-scaled utilities between the two worker types, $\ln(\frac{C_{l(g_1)}}{C_{l(g_2)}})$.

However, consider two more groups g_3 and g_4 for which $f(g_3) = f(g_4) = f_2$, $l(g_3) = l(g_1)$, and $l(g_4) = l(g_2)$.The groups g_1 to g_4 can be chosen to be the two ways to match two positions with two workers. Dividing (6) by its analogue using g_3 and g_4 conditional on f_2 and rearranging yields:

$$\ln(\frac{\hat{P}_{g_1|f_1}/(\overline{S}_{g_1|l(g_1),f_1}n(l(g_1)))}{\hat{P}_{g_2|f_1}/(\overline{S}_{g_2|l(g_2),f_1}n(l(g_2)))} / \frac{\hat{P}_{g_3|f_2}/(\overline{S}_{g_3|l(g_3),f_2}n(l(g_3)))}{\hat{P}_{g_4|f_2}/(\overline{S}_{g_4|l(g_4),f_2}n(l(g_4)))}) = \frac{(\theta_{g_1} - \theta_{g_2}) - (\theta_{g_3} - \theta_{g_4})}{\sigma} \tag{7}$$

[20] Because we do not observe unfilled vacancies, in the empirical work we augment $\mathcal{K}$ to include a sufficient number of unemployment "positions" to ensure that each match will have both a worker and a "position".

Thus, the adjusted log odds ratio identifies the expected gain in scaled joint surplus from swapping partners in any two job matches. Note that differencing and conditioning, respectively, necessarily remove any information about the mean payoffs or welfare of worker types and position types. However, the identified set of surplus difference-in-differences $\Theta^{D-in-D} \equiv \{\frac{(\theta_g-\theta_{g'})-(\theta_{g''}-\theta_{g'''})}{\sigma} \,\forall\, (g, g', g'', g''') : l(g) = l(g''), l(g') = l(g'''), f(g) = f(g'), f(g'') = f(g''')\}$ preserves the critical information about the relative efficiency of alternative matchings present in the observed group frequencies.

For example, if one observes that high-paying firms tend to hire other firms' high-paid workers and low-paying firms tend to hire other firms' low-paid workers but not vice versa, there must be greater combined surpluses from the first two kinds of matches relative to their alternatives. Whether due to production complementarities, greater tastes for workplace amenities by low-paid workers, or reduced moving costs if the two sets of workers and firms are spatially segregated, we show that evaluating the incidence of counterfactual labor demand shocks does not require identifying the source of this comparative advantage as long as it is not meaningfully affected by the shock.

3.2 Counterfactual Simulations

We now show that identification of Θ^{D-in-D} is sufficient to generate the unique counterfactual aggregated assignment $P^{CF}(g)$ and the shares of utility and profit gains or losses by worker and position type following arbitrary changes in the distributions of these types. If multiple matchings are observed, σ can also be (roughly) estimated and utility and profit gains can be scaled in dollars.

We characterize the set of workers to be reallocated via the worker type distribution, $n^{CF}(l)$, where "CF" indicates that this distribution could be counterfactual (e.g. capturing a proposed influx of refugees). Similarly, we use $h^{CF}(f)$ to capture the set of counterfactual positions to be filled, and $\{\theta_g^{CF}\}$ to denote the relevant group mean surplus values (i.e. the prevailing matching technology). $n^{CF}(l)$, $h^{CF}(f)$, and $\{\theta_g^{CF}\}$ are all inputs that are either observed or chosen by the researcher.

As a motivating example, suppose a local development board has forecasted the number and location of new manufacturing positions that a plant opening would generate, and has data on past job match patterns. The board may wish to predict the change in job-related utility and the employment rate induced by the plant opening among existing local workers/job seekers in the chosen and surrounding neighborhoods (and perhaps the profits of local and less local firms).

We assume that the counterfactual assignment also satisfies the assumptions of Proposition A1 above. We also assume that the set of position type averages of the shares of potential job and industry stayers among each worker type, $\{\overline{S}_{g|l,f}^{CF}\}$, is known, and treat it as an input. When the counterfactual type shares $n^{CF}(l)$ and $h^{CF}(f)$ are set equal to those from some observed year y, these shares can be directly computed. Then the counterfactual CCP $P^{CF}(g|f)$ can be expressed as (5) with $(\theta_g^{CF}, n^{CF}(l), h^{CF}(f), \overline{S}_{g|l(g),f(g)}^{CF}, C_l^{CF})$ replacing $(\theta_g, n(l), h(f), \overline{S}_{g|l(g),f(g)}, C_l)$. The worker type-specific mean exponentiated (and rescaled) utility values $\mathbf{C}^{CF} \equiv \{C_1^{CF} \ldots C_L^{CF}\}$ are ex ante unknown equilibrium objects affected by the counterfactual changes reflected in $(\theta_g^{CF}, n^{CF}(l), h^{CF}(f))$. Thus, each counterfactual CCP must initially be treated as a function of the set $\mathbf{C}^{CF}$.

GS and Decker et al. (2013) each show that a unique probability distribution over match groups $P^{CF}(g)$ satisfies the aggregate analogues to the stability and feasibility conditions. However, these papers as well as CS assume when proving identification that one observes the total number of agents of each type, including unmatched partners, on both sides of the market. While counts of unemployed workers by type can be accurately constructed, the LEHD data contain no information about unfilled vacancies.[21] Because each submatching of a stable matching must also be stable, observing only filled positions does not threaten identification of the remaining elements of Θ^{D-in-D}; the estimated relative surpluses would not change if data were augmented with vacancies.

In principal, though, unfilled positions may put upward pressure on wages that alters the division of surplus between workers and positions, even if they do not affect the final job assignment. However, many unfilled vacancies may not be the second-best option for any worker, or may only be slightly more appealing than a third-best position that settles for another worker, so that a large share of unfilled vacancies ignored by $h^{CF}(f)$ may negligibly affect the division of surplus. A related concern is that firms might wish to alter how many positions they choose to fill when wages change in the wake of labor demand shocks. However, for relatively small and localized shocks, firms may display a highly inelastic extensive margin response if the costs of adjusting the number of positions at establishments (and perhaps changing workers' tasks) are large relative to the shock-induced changes in the minimized cost of an efficiency unit of labor. In this case establishments may only adjust the composition of workers they choose to fill a fixed set of positions.

While our baseline estimates maintain an assumption of perfect inelasticity on the extensive margin, for larger shocks that make this untenable, one can use existing wage elasticity and multiplier estimates to incorporate the endogenous response into $h^{CF}(f)$ and re-interpret $h^{CF}(f)$ as a post-adjustment distribution. As a robustness check in Section 6.6, we explicitly endogenize the extensive margin in this way by iterating between assignment model equilibria and calibrated extensive margin responses to changes in a position's expected profitability until a fixed point is found.

Treating the set of positions that will be filled as exogenous (at least within an iteration) simplifies the choice of variation used to identify relative surplus values. One need not isolate labor supply shocks in order to identify extensive margin labor demand elasticities by type. Instead, surplus diff-in-diffs Θ^{D-in-D} (along with σ) are essentially determining equilibrium elasticities of substitution for each position type among different worker types. Elasticities of substitution are fully determined by *relative* prices, so they should be insensitive to the source of relative cost changes for different worker types: upward (downward) shifts in the number of local (distant) workers seeking positions and downward (upward) shifts in the number of local (distant) positions tending to prefer local (distant) workers are all valid sources of variation in relative prices of workers from different initial locations. So there is no inconsistency in using the full set of year-to-year job flows, implicitly driven by a mix of many small and large local supply and demand shocks, to recover Θ^{D-in-D}.

Requiring all positions in h_f^{CF} to fill also simplifies the computation of counterfactual equilib-

[21] Constructing vacancy counts for our position types from publicly available vacancy data is also not straightforward.

ria. With an unknown number of unmatched partners on each side, GS show that one must solve $L+F$ non-linear equations that combine the feasibility and stability conditions for the mean equilibrium payoffs of all worker and firm types ($\{C_l^{CF}\}$ and $\{C_f^{CF}\}$). By contrast, when the "supply" of positions by type is assumed known, each can be set equal to worker "demand" for such positions to create F market clearing conditions that determine $\{C_f^{CF}\}$.[22] Equivalently, if a dummy "position" type is added with mass equal to the share of workers who will end up unmatched, then the augmented demand (including "demand" from unemployment) for each worker type l will equal the supply $n^{CF}(l)$, facilitating market clearing on the worker side.[23] Since relative payoffs among worker types fully determine the equilibrium assignment (so one can normalize $C_1^{CF}=0$), and the worker type distribution $n^{CF}(*)$ must sum to one, we obtain $L-1$ market clearing conditions:

$$\begin{aligned}
&\sum_{f\in\mathcal{F}} h^{CF}(f)(\sum_{g:l(g)=2} P^{CF}(g|f,\mathbf{C^{CF}})) = n^{CF}(2)\\
&\vdots\\
&\sum_{f\in\mathcal{F}} h^{CF}(f)(\sum_{g:l(g)=L} P^{CF}(g|f,\mathbf{C^{CF}})) = n^{CF}(L)
\end{aligned} \tag{8}$$

Given a solution to (8), one can then construct the counterfactual probability for any match group via $P^{CF}(g)=\sum_f h^{CF}(f)P^{CF}(g|f,\mathbf{C^{CF}})$. Since the solution also satisfies the stability and feasibility conditions, it must be the unique aggregate counterfactual stable assignment.

Because only $\min\{L,F\}$ equations must be solved, this approach provides considerable computational savings when the number of types is much larger on one side of the market. Below we present results that average over 300 counterfactual allocations featuring around 5,000 worker and 10,000 position types that would be prohibitive to compute with unmatched agents on both sides.

3.3 Interpreting the Counterfactual Simulations

We generally use data from the 2012-2013 set of job matches (i.e. $y=2012$) to form our simulation inputs, so that $\Theta^{CF}=\Theta^{2012}$, $n^{CF}(*)=n^{2012}(*)$, and $h^{CF}(*)$ will equal $h^{2012}(*)$ plus a shock consisting of positions added to or subtracted from a chosen type f. We wish to interpret the difference between the resulting counterfactual reallocation and the observed 2012-2013 reallocation as the one-year impact that such a shock would have caused in that economy. However, a few additional assumptions and clarifications are needed to justify and elaborate on this interpretation.

First, constructing the market-clearing conditions (8) requires a full set of group joint surpluses $\Theta^{2012}\equiv\{\theta_g^{2012}\ \forall\ g\in\mathcal{G}\}$, but the identification argument in section 3.1 suggests that only the set of diff-in-diffs $\Theta^{D-in-D,2012}$ is identified. In Appendix A2, we prove the following proposition:

[22] Koopmans and Beckmann (1957) point out that when unmatched agents only exist on one side of the market, the dual problem payoffs need only be recovered on one side of the market in order to construct the stable assignment.

[23] These dummy unemployment positions represent a computational mechanism for incorporating workers' payoffs from unemployment, $\{\pi_{i0}^i\}$, akin to "balancing" an unbalanced assignment problem (Hillier and Lieberman (2010)). A formal proof of equivalence is proposed as Proposition A2 and proved in Appendix A3.

Proposition 1:

Define the set $\Theta^{D-in-D} \equiv \{\frac{(\theta_g-\theta_{g'})-(\theta_{g''}-\theta_{g'''})}{\sigma} \; \forall \, (g, g', g'', g''') : l(g) = l(g''), l(g') = l(g'''), f(g) = f(g'), f(g'') = f(g''')\}$. *Given knowledge of* Θ^{D-in-D}, *a set* $\tilde{\Theta} = \{\tilde{\theta}_g \; \forall \; g \in \mathcal{G}\}$ *can be constructed such that the unique group level assignment* $P^{CF}(g)$ *that satisfies the market-clearing conditions (8) using* $\theta_g^{CF} = \tilde{\theta}_g \; \forall \; g$ *and arbitrary marginal PMFs for worker and position types* $n^{CF}(*)$ *and* $h^{CF}(*)$ *will also satisfy the corresponding market-clearing conditions using* $\theta_g^{CF} = \theta_g \; \forall \; g \in \mathcal{G}$ *and the same PMFs* $n^{CF}(*)$ *and* $g^{CF}(*)$. *Furthermore, denote by* $\tilde{\mathbf{C}}^{\mathbf{CF}} \equiv \{\tilde{C}_1^{CF}, \ldots, \tilde{C}_L^{CF}\}$ *and* $\mathbf{C}^{\mathbf{CF}} \equiv \{C_1^{CF}, \ldots, C_L^{CF}\}$ *the utility vectors that clear the market using* $\theta_g^{CF} = \tilde{\theta}_g$ *and using* $\theta_g^{CF} = \theta_g$, *respectively. Then* $\tilde{\mathbf{C}}^{\mathbf{CF}}$ *will satisfy* $\tilde{C}_l^{CF} = C_l^{CF} e^{\frac{-\Delta_l}{\sigma}} \; \forall \; l \in \mathcal{L}$ *for some set of worker type-specific constants* $\{\Delta_l : l \in [1, L]\}$ *that is invariant to the choices of* $n^{CF}(*)$ *and* $h^{CF}(*)$.

Essentially, the proposition states that the identified set of surplus difference-in-differences Θ^{D-in-D} contains sufficient information to generate the unique counterfactual group-level assignment $P^{CF}(g)$ associated with the complete set of surpluses Θ. Furthermore, the utility premia $\tilde{\mathbf{C}}^{\mathbf{CF}}$ that clear the market using the artificially completed surpluses $\tilde{\Theta}$ will always differ from the "true" premia $\mathbf{C}^{\mathbf{CF}}$ that clear the counterfactual market under Θ by the same l-type-specific constants regardless of the compositions of supply $n^{CF}(l)$ and demand $h^{CF}(f)$ that define the counterfactual.

The "bias" terms $\{\Delta_l\}$ in Prop. 1 imply that relative levels of baseline utility among worker types are not identified. However, because Δ_l values are constant across counterfactuals with different $n^{CF}(l)$ and $h^{CF}(f)$ distributions, relative *changes* $[(\ln(C_l^{CF1}) - \ln(C_l^{CF2})) - (\ln(C_{l'}^{CF1}) - \ln(C_{l'}^{CF2}))] \approx (\frac{(\overline{r}_l^{CF1} - \overline{r}_l^{CF2}) - (\overline{r}_{l'}^{CF1} - \overline{r}_{l'}^{CF2})}{\sigma})$ in mean rescaled utilities across worker types among two counterfactuals are identified.[24] Below, we pair counterfactuals that feature targeted local demand shocks with otherwise identical counterfactuals that do not. We assume that the small, very local stimuli and plant closings we consider do not alter utility for the least affected (usually quite distant) worker type, so that utility changes $\frac{\overline{r}_l^{CF1} - \overline{r}_l^{CF2}}{\sigma}$ for other types can be identified, as can each worker type's share of total welfare gains or losses from the shock. The model's symmetry between workers and positions implies that mean changes in profits and shares of profit gains or losses by position type also are identified. Thus, given data on a single matching, the model can produce a fairly complete account of job-related welfare incidence from labor supply and demand shocks.

Second, besides these normalizations, in order for the predicted allocation and welfare gains to accurately reflect what would have happened had the simulated shocks occurred, one must also assume that the joint surpluses diff-in-diffs $\Theta^{D-in-D,CF}$ and marginal type distributions $n^{CF}(*)$ and $h^{CF}(*)$ that act as simulation inputs are exogenous to (i.e. unaffected by) the shock itself. Any reallocation and welfare changes are assumed to be driven exclusively by the changes in transfers across worker types required to eliminate shock-induced imbalances between supply and demand.

Exogeneity of $h^{CF}(*)$ imposes that the shock does not cause further changes in firms' location

[24] This insight mirrors that of Caliendo et al. (2019). The approximation requires limited variation in utility values among workers of the same type, so that $\ln(C_l) \equiv \ln(\frac{1}{|l|} \sum_{i:l(i)=l} e^{\frac{r_i}{\sigma}}) \approx \ln(\frac{1}{|l|} \sum_{i:l(i)=l} e^{\frac{\overline{r}_l}{\sigma}}) = \frac{\overline{r}_l}{\sigma}$.

and size decisions. To highlight heterogeneity in incidence by firm size, average pay, and industry, we generally consider simple "apples-to-apples" comparisons where each shock adds or subtracts a common number of jobs to a single position type. However, in addition to endogenizing firm responses to shock-induced wage changes (discussed above), we consider a second robustness check that incorporates product market spillovers by adding extra service positions in locations near the original "exogenous" shock, guided by job multiplier estimates from Bartik and Sotherland (2019). Other agglomeration and congestion forces could be similarly built into the simulated shock.

There are also plausible mechanisms by which the joint surpluses Θ^{2012} might respond to the shock, particularly for large shocks representing a "big push" (Kline and Moretti (2014)).[25] However, for reasonably small local shocks, the most obvious endogenous surplus changes are likely to be minuscule relative to the size of existing surplus variation in worker types' relative productivities, amenity valuations, and moving costs across firm types, so that such exogeneity violations generate minimal bias. Note also that only changes in surplus diff-in-diffs Θ^{D-in-D} affect the counterfactual assignment, so that the components of endogenous changes to productivities, amenities, or continuation values among position types that are common to all workers do not affect the shock's worker incidence.[26] Nonetheless, to assess sensitivity to unmodeled changes in local continuation values, we consider simulations that build into the shock the average joint surplus changes among groups within the surrounding PUMA from a sample of observed establishment openings.

Another caveat relates to shock duration. We focus on forecasting reallocations and welfare changes that occur within one year of the shocks and we assume that job matches with shock-generated positions create the same surplus as those with existing positions of the targeted position type. Implicitly, this requires that the new positions have the same expected duration as other positions of their type.[27] As is, the model is designed to show that the incidence of very local shocks may spread quite widely across space and demographic groups even over a short period, despite movers' strong tendencies to take nearby jobs, consistent with large short-run mobility frictions.

A final, important caveat relates to the absence of a housing market in the model (and residential choices in the data). Standard models of spatial equilibrium in urban economics (e.g. Roback (1982) or Kline and Moretti (2013)) emphasize that if housing supply is inelastic and workers are mobile, increases in housing and rent prices may offset a substantial share of job-related utility gains to local workers if they are also nearby renters. Sprung-Keyser et al. (2022)'s estimates suggest that increases in rent and other local non-tradeables offset between 30% and 50% of earnings gains in target commuting zones from broader local labor demand shocks, so that the majority of the job-related utility gains we identify are likely to remain after accounting for other price changes.

[25] A new establishment might increase the demand for other local firms' intermediate goods, raising their value of workers. Alternatively, if search/recruiting/moving costs increase with distance, then jobs at nearby establishments might now have greater continuation value because future job searches will begin in a local area featuring greater labor demand.

[26] In Appendix A2, such surplus changes only affect Δ_f^2, which shifts the position type's profit but does not enter into equilibrium utilities for worker types $\{C_l^{CF}\}$. This partly motivates the focus on incidence among workers, for whom differential agglomeration effects among firms across shock compositions may be less important.

[27] One could allow separate surpluses for short- and long-lived jobs. Fully analyzing shocks of varying duration requires a fully dynamic assignment model akin to Choo (2015) that specifies worker expectations and serial correlation in ϵ_{ik}^y.

These considerations suggest that local low-paid workers would be justified in resisting local initiatives focused on bringing "good" jobs to town if they are likely to generate an employment-related incidence that is either geographically dispersed or concentrated among higher-paid workers.

Furthermore, Hornbeck and Moretti (2024) and Sprung-Keyser et al. (2022) show that house prices increase less in places with relatively elastic housing supply (e.g. rural areas, areas with weak zoning laws). Similarly, in areas with low commuting costs, adjustment to the small, localized shocks we consider may occur primarily via changing commuting patterns rather than residential moves, with diluted house price impacts across the variety of locations from which workers commute.[28] Since changes in commuting costs from work location changes are implicitly captured in the model as a component of joint surplus θ_g, shock-induced commuting changes will generally be reflected in our welfare estimates.[29] Thus, job-related welfare gains may closely approximate total welfare gains in these cases. While a complete welfare analysis requires explicitly incorporating housing and product markets, this paper's goal is to highlight the roles of differential geographic scopes of local labor markets for different types of workers and firms and the skill vs. spatial mismatch tradeoff in determining the incidence of alternative local labor demand interventions.

3.4 Identifying σ

The share of welfare gains or losses for workers (or firms) can be recovered without estimating σ. However, since payoffs are additive in worker earnings, knowledge of σ allows the estimated utility gains $\frac{\bar{r}_l^{CF1}-\bar{r}_l^{CF2}}{\sigma}$ and profit gains $\frac{\bar{q}_f^{CF1}-\bar{q}_f^{CF2}}{\sigma}$ to be scaled in dollars, making it easy to gauge the economic importance of welfare changes from various shocks. Conditional on Θ, σ sets the elasticity of matching choices with respect to relative wages or required utility bids, so that it governs the magnitude of utility reallocations across worker types due to changes in labor demand.[30]

As Galichon et al. (2017) have noted, identification of σ requires combining information from multiple matchings, so we estimate σ using observed matchings between 2003-2004 and 2012-2013. Because the procedure (described fully in Appendix A4) requires additional strong assumptions, estimates of σ are likely to be quite rough, though they are fairly consistent across years.[31] We use

[28]Recall that our incidence assessments classify workers by work location rather than residential location.

[29]Differential willingness to pay for locational amenities will be reflected in the relative propensities for different worker types to move to positions at particular locations, which are captured by the odds ratios used to identify Θ^{D-in-D}.

[30]Intuitively, when position type C disproportionately chooses type A workers over type B workers compared to position type D, it could be because $\theta_{AC}-\theta_{AD} \gg \theta_{BC}-\theta_{BD}$ and σ is substantial, or because $\theta_{AC}-\theta_{AD}$ is marginally larger than $\theta_{BC}-\theta_{BD}$ but σ is tiny. When the former is true, large changes in required utility bids are necessary to engender sufficient substitution across worker types to overcome strong comparative advantages from matching certain types of workers and positions. If the latter is true, small utility changes suffice to clear the market after a shock.

[31]Essentially, differences in worker types' observed mean earnings changes between years $y-1$ and y are regressed on model-generated log differences in predicted scaled utility values $\ln(C_l^{CF,y})-\ln(C_{l'}^{CF,y}) \approx (\bar{r}_l^{CF,y}-\bar{r}_{l'}^{CF,y})/\sigma^y$. These predicted values are constructed by computing counterfactual equilibria in which worker and position type distributions evolve as they actually did but surpluses are fixed at their 2003-2004 values. The coefficient on $(\bar{r}_l^{CF,y}-\bar{r}_{l'}^{CF,y})/\sigma^y$ will approximately equal σ^y under the assumptions that a) the evolution in the utility premia enjoyed by different worker types was due primarily to changes in supply and demand composition rather than relative changes in the moving costs, recruiting costs, tastes, and productivities that compose the surpluses Θ, and b) mean utility gains for each worker type in the chosen year generally consisted of earnings increases rather than increases in amenities or continuation values.

the mean of $\hat{\sigma}^y$ across all years, $\overline{\sigma} = 18,420$, to assign dollar values to all utility changes.

As noted by GS, in the CS model observed earnings also can be used to separate each mean joint surplus θ_g into worker and position components θ_g^l and θ_g^f. In Appendix A5 we show that clean identification of θ_g^l and θ_g^f breaks down without the particular structure CS place on the unobserved match quality component ϵ_{ik} unless further strong assumptions are imposed. We do not pursue this approach because we have shown this decomposition is not needed to recover the dollar-valued welfare incidence across worker and position types of alternative local labor demand shocks.

4 Data

We construct a dataset of workers' pairs of primary jobs in consecutive years using the Longitudinal Employer-Household Dynamics (LEHD) database. The core of the LEHD consists of state-level records containing quarterly job earnings and unique worker and firm IDs (state EINs) for nearly all jobs in the state.[32] The worker IDs are then linked across states, and the data are augmented with establishment assignments, establishment characteristics (notably location and industry) from an extract of the ES-202/QCEW report, and worker demographics from the Social Security Administration (including age, race and sex but not occupation nor education for most workers).[33]

4.1 Sample Selection

Our sample consists of all LEHD records from the 19 U.S. states that opted to provide data to our project.[34] A person is included in the initial sample if he/she is ever observed as employed in one of the sample states in any of the years for which data was provided, and for disclosure avoidance reasons a 50% random sample of all initial sample members is taken to form the final sample. The 2014 LEHD snapshot includes a file that indicates whether a worker was employed in some U.S. state in each quarter, even among states not providing records to a particular project, as long as the state provided data to the Census Bureau. Thus, job transitions into and out of the 19 observed states can be distinguished from transitions to and from nonemployment. While the estimation of σ and the model validation exercise use all the data after 2002 (when the last sample state begins reporting data), the model simulations use surplus parameters estimated from 2012-2013 data.[35] Preliminary work suggested that the shock incidence forecasts were quite insensitive to the years chosen.[36]

[32] The database does not include farm jobs, self-employed workers, or federal employees.

[33] A worker's establishment must be imputed for multi-establishment firms, and is fixed within a spell at the firm. However, the LEHD's unit-to-worker imputation procedure assigns establishments with probabilities that depend on the distance between that establishment and the worker's residence, so any mistakes will likely misattribute the worker's job to another nearby establishment, limiting scope for significant measurement error. We use the LEHD's Successor-Predecessor file to reclassify as retentions any spurious job transitions stemming from changes to a firm's structure that do not alter a worker's location. See Abowd et al. (2009) and Vilhuber et al. (2018) for further details about the LEHD.

[34] By agreement with the Census Disclosure Avoidance Review staff, the identities of the states cannot be revealed, but they include large, medium, and small states, and are spread throughout the U.S., albeit unevenly.

[35] An advantage of less recent data is that the vast majority of positions were in-person rather than remote, so that job transition distances and the surplus values they imply should be relevant for job stimuli that create in-person positions.

[36] This was true despite the decreasing job-to-job mobility over this time period documented by Hyatt et al. (2016).

To form job change/retention observations, we select each worker's highest earnings job in each year among those lasting at least one full quarter and then append the next year's primary job.[37] Workers are considered nonemployed in a given year if they did not earn above $2,000 at any job in any full quarter in any observed state and are not reported as employed in an out-of-sample state.

To try to isolate workers who are in the labor force, each worker's presence in the sample begins and ends with his/her first and last years of observed employment. We also drop workers with ages below 20 or over 70. This limits the influence of "nonemployment" spells consisting of full-time education or retirement followed by part-time work, so that parameters related to unemployment are identified by prime-aged workers who were unemployed or temporarily out of the labor force.

Since most results presented below rely on parameters estimated using 2012-2013 matches and sample coverage ends in 2015Q1, excluding nonemployment spells without an observed resumption of employment may cause a slight undercount of E-to-U and U-to-U transitions, since a small share of unemployed workers in 2013 likely remained in the labor force but did not find jobs by 2015Q1. We address this by using the American Community Survey, which distinguishes unemployment from labor force exit, to construct estimated counts of E-to-U and U-to-U transitions by combination of initial U.S. state, destination state, 5-year age bin, and initial earnings category (for E-to-U only). These aggregated match groups are coarser than those in the model, so we use the LEHD's E-to-U and U-to-U transitions only to distribute the ACS group counts across the model's finer groups. We supplement the ACS data with BLS national unemployment counts by age group to align the scale of the labor force with standard measures. Appendix A6 details these imputation procedures.

A drawback of the LEHD sample is that the establishment and pay of employed workers are only observed among the 19 sample states. Rather than exclude out-of-sample workers, which would cause us to overstate the geographic concentration of shock incidence, we aggregate all out-of-state employment into a single out-of-sample "state" and "tract". As with flows to unemployment, we use aggregate ACS counts to set the scale of flows between in- and out-of-sample states, and then use the LEHD to impute the joint distribution of worker and position characteristics among flows into and out of each in-sample census tract (see Appendix A6). Because incidence forecasts may be especially sensitive to observing worker flows to and from states adjacent to the focal state, our simulations generally sample target tracts only from 10 states in the west/southwest/great plains area where coverage is nearly complete and almost all adjacent states are observed.[38]

4.2 Assigning Workers and Positions to Types and Job Matches to Groups

For each pair of years $(y-1, y)$ we assign each observation to a worker type $l(i)$, a position type $f(k)$, and a match group $g(i,k)$. Workers' type assignments are based on the combination of their $y-1$ primary establishments' locations (discussed in Section 5.1), the age-adjusted earnings quartile

[37] A job is observed in a full quarter if it features positive earnings in the preceding and following quarter as well.

[38] Using ACS 1-year residential mobility data and weighting states by their census tract count, we estimate that for the 10 states supplying target tracts, about 47% of year-to-year worker inflows from other states and about 92% of total job-to-job changes ending in one of these 10 states (including within-state flows) originate in one of 19 in-sample states.

associated with their $y-1$ earnings at this establishment, their age category ($\leq$ 30, 31-50, or $>$ 50), and whether their $y-1$ industry supersector matches that of the simulated job creation or destruction.[39] For workers who were not employed in $y-1$, the location of their most recent establishment is used (or, for new entrants, the location is imputed using ACS/LEHD data) and the earnings quartile is replaced by a separate category for unemployment. Workers' year y positions are assigned to position types based on the combination of their establishment $m(k)$'s location, supersector, employment size (below/above the worker-weighted median) and average worker earnings (below median, quartile 3, or quartile 4). These characteristics were chosen because they are consistently observable and likely to be key determinants of productivity complementarities, recruiting, search and moving costs, and the other components of the match surpluses among heterogeneous positions and workers. Match groups $g(i,k) \equiv g(l(i), f(k), z(i,k))$ are based on the worker's type $l(i)$, the position's type $f(k)$, and a trichotomous indicator for whether the match keeps the worker at his/her $y-1$ firm ($z(i,k)=1$), his/her $y-1$ industry but not firm ($z(i,k)=2$), or neither ($z(i,k)=0$).

4.3 Summary Statistics

Figure 1a (Col. 1 of Table A3) presents the distribution of distance between the locations of origin and destination establishments for workers who changed primary jobs $(m(j) \neq m(k))$ between 2012 and 2013. 3.2% of job switchers took new jobs within the same census tract, while another 5.7%, 6.1%, and 12.2% moved to jobs one, two, or 3+ tracts away within the same PUMA. 54.7% found jobs in another PUMA within the same state, while 18.1% changed states. The sizable share of workers accepting new jobs very near their previous jobs is prima facie evidence that either search/moving costs are large or preferences for particular locations are strong, so that conditions in workers' local labor markets may still hold outsized importance for their job-related welfare.

Row 1 of Table 1 Panel A shows that 15.6% of sample observations involve job-to-job transitions, with 8.3% changing supersector. A full 69.5% of workers keep the same primary job, while 9.3%, 2.8% and 2.8% make U-to-E, E-to-U, and U-to-U transitions, respectively. In total, the 2012-2013 estimation sample for the set $\{\Theta^{D-in-D}\}$ contains 24.2 million observations.

Examining other rows of Panel A, we see that 77.1% of workers who were unemployed in 2012 found jobs in 2013. $U-E$ rates vary sharply by age, however: 86.5% of age $\leq$ 30 workers (including many new entrants) find jobs, while only 68.7% and 60.8% of initially unemployed workers aged 31-50 and over 50 find jobs. Among those employed in 2012, younger workers ($\leq$30) were also far less likely to stay at their establishment (66.3%) than those aged 31-50 (81.4%) or over 50 (87.9%). Similarly, workers in the lowest age-adjusted earnings quartile in 2012 were far less likely than the highest paid workers to stay at their job (70.1% vs. 84.3%) and far more likely to become unemployed (5.6% vs. 1.6%) or take another job (24.2% vs. 14.1%). Given a job change, the highest paid were also more likely to stay within the same industry (53.2% vs. 41.3% for the lowest

[39] Earnings quartile cutoffs are defined using the distribution of primary job annual earnings among all same-aged workers in year $y-1$, and are based on prorating earnings from full quarters. The age and proration adjustments allow the quartile to better capture full-time pay relative to peers rather than experience or share of the year he/she worked.

quartile), but were the most likely to leave their original PUMA (78.3% vs. 69.7%) and their state (24.4% vs. 15.7%), suggesting that the geographic scope of labor markets varies across earnings categories. These differences motivate using age, earnings, and industry to define worker types.

Panel B of Table 1 shows that the highest paying quartile of firms retain a much greater share of their workers (80.3%) than those with below-median pay (68.3%), but hire distant workers more often when filling a vacancy: 22.1% of their new hires had been working out of state and 22.2% had been working in the same PUMA, compared to 16.5% and 29.0% for those with below-median pay. Firms above median size are more likely than small firms to retain workers (78.0% vs. 70.0%), but less likely to hire from within the same PUMA (21.9% vs. 30.8%). Industries (Panel C) also vary widely in their job retention rates (from 62.1% for leisure & hospitality to 82.9% for manufacturing) and shares of hires from unemployment (from 6.1% for finance to 17.9% for leisure & hospitality).

The heterogeneity in job staying rates in particular reveals important differences in joint surpluses across match groups that shape the demand shock incidence analyzed below. To see this, note that on average workers who remain employed in the same tract are 136.7 times more likely to be firm stayers than firm switchers, even though a random worker is on average only 1/20th as likely to be a given firm's incumbent as an incumbent at a different firm of the same type in the same tract (since the sample mean of $\overline{S}_{z=1}$ is near .05). Thus, job retentions occur nearly 2,714 times more than random matching would predict conditional on worker and firm type, implying a relative surplus value of $\frac{\theta_{stay}-\theta_{move}}{\sigma} = \log(2714) = 7.9$. However, retentions among workers under 30 occur only around 1,600 times as often as under random matching, versus 5,712 for workers over 50.

While these statistics illustrate the data patterns driving the variation in joint surpluses, motivate the choices of types, and illustrate the need to consider shocks featuring different establishment compositions, they do not condition on any other firm, location, or worker characteristics. Comparing incidence across counterfactual shocks that hold all but one characteristic fixed will be more informative about how the scope of labor markets differs across types of workers and firms.

5 Estimation

5.1 Collapsing the Type Space for Distant Geographic Areas

Since match groups g are defined by several other worker and position characteristics in addition to worker and position locations, treating all 28,000 census tracts in our 19 states as separate locations would generate trillions of match groups. Given the particular interest in the incidence of alternative shocks among locations near the shock, we combine initial types (and thus groups) that share the same worker and position characteristics and are geographically close to each other but far from the shock. Specifically, beyond a five tract radius around the targeted tract, a type's location is defined only by its PUMA. Beyond the targeted state, a type's location is defined only by its state.

Coarsening the type space for distant locations dramatically reduces the number of groups and the sparsity of the empirical group distribution $\hat{P}(g)$. While many workers move between nearby tracts, very few move between most distant tract pairs, so relative surpluses for groups whose tracts are in different states would be weakly identified without such coarsening. This approach still uses in each simulation all observed job matches and all locations in the 19 state sample as well as the out-of-sample "state", and each local labor market remains nested within a single national market.

Even after combining types, there are relatively few observed matches per group g, particularly for groups local to the shock, so that Dingel and Tintelnot (2020)'s concerns about overfitting with granular data remain relevant. Thus, following Hotz and Miller (1993) and Arcidiacono and Miller (2011), we smooth $\hat{P}(g)$ prior to estimation by replacing each element's value with a kernel-density weighted average of $\hat{P}(g)$ among groups featuring "similar" worker and position characteristics.

Because excessive smoothing erodes the signal in the data about the degree of heterogeneity in joint surpluses from matches with different worker and position characteristics and locations, we create a customized smoothing procedure, detailed in Appendix A7. It is based on the idea that the hiring establishment's location is critical in determining the origin locations from which hires create the most surplus (i.e. least moving/search cost), while non-location attributes (size, avg. pay, and industry) primarily determine the surplus-maximizing worker earnings/age/industry category. Table A1 repeats the summary statistics from Table 1 for the smoothed sample. The smoothed and raw transition rates generally differ by .001 or .002 and almost never by more than one percent, providing reassurance that the procedure is preserving the essential variation in matching patterns.

The type aggregation and smoothing procedures imply that type and group spaces vary by target tract.[40] FSRDC disclosure rules also bar the release of results for specific substate locations. Thus, for each shock type we report averages of incidence measures across 300 simulations that randomly choose a target tract from the 10 state southwest/west/great plains subsample.[41] Averaging results across targeted tracts requires redefining match groups post-simulation. Worker and position type locations are replaced with bins of distance to the targeted tract, and we report incidence estimates for various distance rings around the shock.[42] We mostly focus on distance bins defined by tract, PUMA, and state pathlengths, since the number of workers contained within circles defined by the same pathlength is more consistent across urban and rural areas than circles with miles-based radii.

5.2 Defining the Local Labor Demand Shocks

Baseline simulated shocks either add or remove 250 jobs from the stock of positions to be filled in a chosen census tract and remove or add 250 national unemployment "positions".[43] This represents about a 10% change in labor demand for an average tract with around 2,500 jobs. For each chosen

[40] Because a worker's type is partly determined by whether their initial supersector matches that of the plant or store opening, changes in the target supersector also change workers' assigned types.

[41] A census tract is only eligible to be a target tract in the simulations if it features at least 250 jobs, so that the parameters governing local firms' and workers' choices are well-identified. The same set of 300 randomly chosen target tracts is used for each shock specification to facilitate fair comparisons among alternative specifications.

[42] Spatial links between adjacent and nearby tracts are not restricted during simulations, so the model does not impose a priori assumptions about the role of distance beyond the initial aggregation of distant tracts to PUMAs and states.

[43] We experimented with "plant relocations" that move jobs to a new location from a distant state. These shocks had nearly identical employment and welfare incidence to their stimulus analogues among workers within the receiving state.

tract, we first simulate 32 “stimulus packages” featuring new establishments with different combinations of the non-location attributes that define a position type: establishment size, average pay, and industry supersector. Table A2 details each shock’s composition. We then consider packages that require the new positions to be filled only by workers from the surrounding PUMA, reflecting stipulations included in some economic development contracts between cities and incoming firms.[44] Comparing these “restricted” specifications to their unrestricted counterparts illustrates the value of such provisions to cities or states. Next, to examine asymmetry between positive and negative shocks and sensitivity to shock scale, we consider several pairs of analogous positive and negative shocks of various magnitudes involving either large high-paying manufacturing firms (“plant openings” and “plant closings”) or large low-paying retail firms (“store openings” and “store closings”). Finally, we run several simulations that evaluate sensitivity of results to key model assumptions.

5.3 Inference

Given that we observe the universe and not a sample of job matches within the available states, it is unclear how to define the relevant population for the purposes of inference. Furthermore, since we estimate nearly a million surplus parameters $\theta_g \in \Theta$, and each counterfactual incidence statistic depends on the full set Θ, any confidence intervals should provide information about the precision of incidence forecasts as opposed to specific parameters. Rather than characterizing sampling error in isolation, we rely on the model validation results presented in section 6.7 to assess the combined contribution of sampling error and misspecification to out-of-sample forecast accuracy.[45]

6 Results

6.1 How Local Are Labor Markets? Aggregated Incidence by Distance to Focal Tract

We focus first on characterizing the geographic scope of labor markets for a “typical” local stimulus by averaging the predicted changes in assignments across the 32 baseline stimuli. This effectively integrates over the joint distribution of establishment industries, sizes, and average pay levels. We primarily discuss figures, but provide accompanying tables in parentheses.

Figure 2a (Table A3, col. 3) displays the mean probability of taking one of the 250 new stimulus jobs among individuals initially or most recently working at different distances from the focal tract. The figure highlights a sense in which U.S. labor markets are quite local: the probability of obtaining a stimulus job is three times higher for a worker from the target tract (.0054) than for one in an adjacent tract (.0017), over 8 times higher than for a worker 2 tracts away (.0007), and almost 20 times higher than for one initially 3 or more tracts away within the same PUMA (.0003). Additional

[44] For example, Empowerment Zones only subsidize wages for employees that are local residents (Busso et al. (2013)).

[45] The first few results tables do provide standard errors that only reflect the sampling error stemming from averaging over a 300 sample of target tracts rather than all available tracts (around 28,000). These standard errors are tiny, suggesting little value to additional simulations per specification. As a result, subsequent tables do not report standard errors.

distance from the focal tract continues to matter at greater distances: a target tract worker is 35 times more likely to obtain a stimulus job than one from an adjacent PUMA, 68 and 233 times more likely than a worker two PUMAs away or 3 or more PUMAs away within the same state, respectively, and 4,279 and 26,566 times more likely than a random worker one state or two or more states away.

However, the target tract contains only 0.002% of the workforce at risk of obtaining these jobs, while other within-PUMA tracts contain 0.146%, other PUMAs within the state contain 6.05%, and other states contain 93.8% (Figure 1b, Table A3 col. 2). Thus, one obtains a very different impression of incidence by swapping the conditioned term and calculating $P(\text{distance from target }|\text{new job})$, the share of stimulus jobs obtained by workers from each distance bin. Figure 2b (Table A3, col. 4) shows that 3.4% of new jobs go to workers from the target tract, another 22.7% go to other workers in the PUMA, 52.8% go to workers in different PUMAs within the state, and 21.1% go to out-of-state workers. Thus, workers far from the target area take a very large share of the new jobs.

Analyzing which workers take the new stimulus jobs may not be very informative about the true incidence of the shock. This is because many workers who take the new jobs would have obtained other similar jobs in the absence of the stimulus, and other workers now obtain these jobs, and so on, creating ripple effects through vacancy chains that determine the true employment and welfare incidence. This is where a flexible equilibrium model provides additional insight.

Figure 2c (Table A3, col. 5) reports the change in the probability of any employment, relative to a no-stimulus counterfactual, by distance from the traget tract. The change in employment rate is quite locally concentrated, but less so than the probability of landing a stimulus job. The stimulus increases target tract workers' employment rate by 0.09%. This is 2.8, 6.2, and 12.6 times greater than for workers 1, 2, or 3+ tracts away within the same PUMA, 19, 29, and 55 times greater than for workers 1, 2, or 3+ PUMAs away within the state, and 339 and 857 times greater than for workers one state and 2+ states away, respectively. The odds of net employment gains for workers 2+ states away relative to focal tract workers are 31 times higher than for obtaining a stimulus job.

Figure 2d (Table A3, col. 6) displays the share of the 250 job increase in national employment that accrues to workers from each distance bin. Only 0.55% of the net employment gain redounds to target tract workers, with 5.3% of the gains going to workers in other tracts within the PUMA, 32.2% to workers in other PUMAs within the state, and 62.0% to workers from out of state.

Figure 2e (Table A3, col. 7) provides the average utility impact, scaled in \$ of 2023 annual earnings, by distance bin from the target tract for the "typical" stimulus package. Recall that we report utility gains relative to the worker type estimated to gain the least.[46] Focal tract workers receive an estimated \$322 increase in money metric utility, while workers 1, 2, and 3+ tracts away receive expected gains of \$105, \$51, and \$26 respectively. Workers initially 1, 2, and 3+ PUMAs away within the state receive \$17, \$11, and \$7, while workers one state away, 2+ states away, and out-of-sample receive \$0.81, \$0.11, and \$0.12. Figure 2f (Table A3, col. 8) plots the share of total utility gains (relative to the normalized type) that accrue to workers in each distance bin. Only 0.9%

[46]The normalized type varies with stimulus composition, but is generally young, unemployed workers in a distant state.

of worker welfare gains accrue to focal tract workers, with 9.0% going to those from other within-PUMA tracts, 54.1% to those from other within-state PUMAs, and 35.9% to out-of-state workers. Thus, welfare gains are considerably more geographically concentrated than employment gains.

Figure 2 (Table A4) displays the incidence measures using miles-based bins. The story is the same: only 6.6% of employment gains and 11.2% of welfare gains accrue to workers within 10 miles of the target tract even though they fill 27.9% of stimulus jobs. 74.2% of employment gains and 54.4% of welfare gains accrue to workers more than 250 miles away or in out-of-sample states.

Figure A2 (Table A5) illustrates the impact on incidence of requiring stimulus positions to only hire workers from the surrounding PUMA. The employment rate for target tract workers rises by 0.5% instead of 0.06%, and increases by 3-4 times more than the unrestricted stimulus for workers from other within-PUMA tracts. Overall, the within-PUMA share of net employment gains increases from 5.1% to 17.5%. The hiring restrictions increase the expected utility gains by over seven-fold ($296 to $2076) for focal tract workers, with 3-5 fold increases in gains for other within-PUMA workers, depending on distance. The share of utility gains accruing to the local PUMA increases from 9.9% to 29.1%. Thus, local development initiatives such as empowerment zones that add stipulations restricting hiring or wage subsidies to only local workers likely cause a much more locally concentrated labor market incidence, even though additional downstream hiring caused by initially employed workers vacating jobs to take the new positions remains unrestricted.

6.2 Heterogeneity in Local Incidence by Worker and Firm Characteristics

The first columns of Panels A and B of Table 2 display the expected increases in employment rate and utility among focal tract workers from various subpopulations defined by initial earnings, age, or same/different industry category, while Figures 3a and 3b plot the shares of local employment and welfare gains accruing to each local subpopulation against its baseline local employment share. Figure 3a shows that the 9.6% of local workers from the same industry as the newly-opened establishment account for 23.7% of local employment gains, partly because their industry knowledge (reflected in substantial surplus premia for same-industry moves) allow them to claim a large share of the new jobs (27.4%). Initially unemployed local workers also enjoy a quite disproportionate 49.2% of local employment gains despite representing 12.1% of the local workforce, while shares of employment gains among initially employed local workers decline with initial earnings quartile. This reflects the lower unemployment risk faced by higher paid workers in the absence of the shock.

Overall, young workers also account for a disproportionate share of local employment gains (39.4% vs. 31.3%), in part because they are often new entrants who are actively searching for jobs. However, further disaggregation reveals additional heterogeneity: among initially unemployed workers, younger workers actually receive less disproportionate gains than older ones (Figure A3a), who are much less likely to find a job otherwise. But this is offset by more disproportionate employment gains for younger compared to older employed workers within the same (age-adjusted) earnings quartile due to higher baseline rates of transition to unemployment.

Local welfare gains are more evenly distributed across initial earnings and age groups (Figure A3b), but here higher paid workers receive slightly larger shares of gains than their workforce shares. And the same-industry share of local welfare gains is even more outsized than for employment gains, with 9.6% of workers enjoying 34.8% of gains. In both cases, their low baseline unemployment risk suggests that most welfare gains take the form of raises and job changes.

The other columns of Table 2 show that the local employment and welfare gains accruing to each subpopulation vary substantially with the industry of the newly created jobs. For example, young local workers benefit the most from leisure & hospitality positions ($442) and the least from professional and business services (PBS) positions ($279), while workers over 50 benefit most from education/health positions ($464) and least from information sector positions ($217). Manufacturing and PBS positions both show substantial gradients in local earnings gains by initial earnings quartile that are absent in government and education/health, with manufacturing producing the third lowest gains for 1st quartile workers and the second highest gains for 4th quartile workers.

More generally, information sector stimuli yield smaller local utility gains due to the sector's greater propensity to hire distant workers, while education/health, which tends to hire locally and offer jobs at various skill and experience levels, generates large gains for all local subpopulations.

Table 3 presents employment and utility gains for focal tract workers by firm size and pay category combinations. As expected, creating positions at the highest paying quartile of firms (regardless of size) generates much larger gains for local high-paid workers ($\sim$ $445) and smaller gains for low-paid ($\sim$ $270) and unemployed ($\sim$ $350) workers. Firms with below median average pay show the opposite pattern, with payoffs of $\sim$ $240, $\sim$ $370, and $\sim$ $525 for the same three groups. Firm size seems to be most important for unemployed local workers, who reap larger utility gains from jobs created at smaller firms, suggesting that local officials may help such workers more by encouraging startups than luring one large establishment to open or relocate.

In addition, substantial further heterogeneity in local incidence exists at the three-dimensional sector/size/pay cell level. Figure A4 plots welfare gains by initial earnings status among focal tract workers for all 32 stimulus compositions. The range of predicted gains is huge. Welfare gains for unemployed workers range from $249 (large, high-paying PBS firms) to $620 (small, low-paying other services). For 1st quartile workers, they range from $167 (large, high-paying information) to $573 (large, low-paying educ./health). For the 4th quartile, they range from $154 (large, low-paying information) to $641 (large, high-paying educ./health). For small precinct councilors concerned with very local incidence, these large differences in the scale and skill intensity of utility incidence may justify tailoring the design of economic development packages to target certain subpopulations, and would be obscured by an analysis that ignored worker heterogeneity or used coarser geography.

6.3 Heterogeneity in National Incidence by Worker and Firm Characteristics

Tables 4 and 5 display the cumulative shares of subpopulation-specific employment and welfare gains, respectively, accruing to workers closer than or within each distance bin. The roughly similar distributions of cumulative shares indicate that per-worker gains decline rapidly with distance for all groups. However, there are subtle but consequential differences in rates of decay. For unemployed workers, 7.0% of their nationwide employment gains accrue to those within the target tract's PUMA, with 44.6% accruing to those in the same state. These values are 5.4% and 33.9% for 1st quartile workers and only 3.6% and 30.3% for the top earnings quartile. Within-state shares of nationwide employment gains are also larger for younger workers (40.3%) than for mid-career (37.0%) and older workers (35.3%), while workers from the shock's industry have twice as large a share of gains within the focal tract as those from different industries. The patterns are very similar for welfare gains, with unemployed, low-paid, younger, and same industry workers all displaying much more locally concentrated gains than their high-paid, older, different industry counterparts.

Such heterogeneity in the geographic scope of labor markets suggests that the substantial variation in local incidence across subpopulations and shock compositions need not translate to the state or national levels. To this end, column 1 of Table A6 reports the national shares of net employment gains by subpopulation, while Figure 4a graphs these shares against their national workforce shares. Like its local counterpart Figure 3a, Figure 4a shows that younger, lower-paid, and particularly unemployed workers enjoy disproportionate shares of employment gains from job stimuli. However, in contrast to Figure 3a, Figure 4a shows that workers already employed in the shock's industry reap a *smaller* share of national employment gains than their workforce share (6.6% vs. 9.6%).

This counterintuitive result reflects a couple of factors. First, the set of positions vacated by workers taking stimulus jobs better approximate the U.S. establishment distribution than the original shock, and each successive ripple of shock-induced reallocation yields an increasingly generic composition of vacated positions. This implies that farther from the site of the shock, workers from the targeted industry have an increasingly small advantage in securing vacated positions. Furthermore, due to the large surplus premium from staying at one's job, most employed workers are not inclined to seek other jobs, making them inelastic to potential job opportunities relative to the unemployed population. Indeed, the only industry whose workers receive a higher national employment gain share than population share when receiving a job stimulus is leisure & hospitality, which has the lowest baseline job-staying rate but the highest industry-staying rate in Table 1.

Figure 4b shows the corresponding national shares of welfare gains by subpopulation. Again, while the slightly disproportionate national shares for higher-paid and younger workers match the local results, same-industry workers only enjoy a slightly larger welfare gain share than their national workforce share (11.5% vs. 9.6%), a major departure from their large local share (34.8%). Disaggregating further to unemployed$\times$age combinations (Figure A5b) reveals a second local vs. national discrepancy. Local mid-career and older unemployed workers disproportionately benefit from local job creation, while at the national level their utility gain share is smaller than their workforce share, reflecting their relative immobility. These results suggest that reduced-form estimates of heterogeneous local effects can be a misleading guide to heterogeneity in incidence at the state or national level, again showing the need for an equilibrium model with two-sided heterogeneity.

The increasingly generic composition of vacated positions with greater distance also implies that which sector receives the shock barely affects the magnitude of the shock's impact nor its geographic, age, or initial earnings incidence beyond the surrounding PUMA. In particular, the share of net employment (utility) gains accruing to within-PUMA workers is between 5.3% and 6.6% (9.1% and 11.5%) regardless of the shock's supersector (Table A7). And shocks to all supersectors feature shares of national employment and utility gains accruing to each earnings and age category that are nearly always within 1% of the category's overall average. This contrasts starkly with the high sensitivity of very local incidence to shock composition. It suggests that county-level and particularly state- and federal-level policymakers may safely ignore differences in demographic and geographic incidence when deciding between local initiatives featuring different sectors.

Changing the firm size/pay composition also barely shifts geographic, and more surprisingly, earnings and age incidence beyond nearby tracts (Tables A6 and A8). Stimuli with low-paying rather than high-paying firms only yield 1-2% higher national shares of employment and welfare gains for low-paid or unemployed workers, compared to 8% higher local shares for such workers (Table 3). Thus, the local incidence understates the degree to which employment and welfare gains from shocks biased toward high-paid workers eventually "trickle down" to unemployed workers.

6.4 Local and National Incidence of Plant and Store Closings

The first row of Table 6 compares the average change in focal tract workers' employment rate (col. 1-2) and expected welfare (col. 5-6), respectively, for both "plant openings" and "plant closings" that create or destroy 250 positions at large, high paying manufacturing firms. The estimates average across 200 focal tracts randomly selected from the subset with at least 500 positions of this position type at baseline, so that such tracts are realistic targets for plant openings and closings. Due to higher baseline job counts, the plant opening represents a smaller percentage change for these tracts, and so only raises the employment rate and welfare gain among focal tract workers by 0.03% and $150. However, a dramatic asymmetry is instantly apparent in the table: the same-sized plant closing lowers these workers' employment rate by 0.59% and their annual earnings-scaled welfare by a whopping $5,624. Focal tract workers account for 0.4% and 1.7% of national employment and welfare gains for the plant openings and 8.6% and 35.7% of losses for the plant closings.

What causes this asymmetry? Plant openings or expansions require new hires, and because local hiring still requires hefty search and training costs, it only yields somewhat larger surplus than hiring more distant workers, so labor demand for locals only increases modestly. The plant opening and job stimuli simulations capture this by imposing that new positions cannot be filled by "job stayers" (groups with $z(g) = 1$).[47] By contrast, plant closings remove a previously large source of joint surplus from worker retention, since recruiting and moving costs had already been paid and workers had acquired firm-specific skill. The high retention rates in all industries in Table 1 reflect the generally large surpluses from preserving matches. Thus, as in the mass layoff literature, with

[47] In a companion paper, Carballo and Mansfield (2023), we show that the asymmetry disappears when we equalize surpluses for retention and replacement by a worker of the same type by setting $z(i, k) = 0$ for all job matches.

far inferior outside options, laid-off workers suffer large welfare losses. This asymmetry illustrates the value of distinguishing retention from replacement by a similar worker and using job-level microdata rather than aggregate job match counts by worker and firm type combinations.

Figure 5a plots each subpopulation's share of all within-tract employment losses against its local workforce share among all 200 plant closing simulations, while both shares are reported in columns 1 and 3 of Table 7, Panel A. In contrast to plant openings, initially unemployed workers account for just 1.6% of local net employment losses, as their unemployment rate only rises by 0.08% (Table 6). This is primarily because none directly lost jobs, but also because they were less likely to be employed even without the shock. Since the shock targeted high paying firms, the share of lost local employment increases in workers' pay quartile from 10.7% for the lowest-paid to 33.4% for the highest-paid. Similarly, because mid-career workers are over-represented in manufacturing (Table 1), they account for a slightly disproportionate share of local employment loss (45.8% compared to 42.4% of workers). A whopping 88.2% of local employment loss is borne by the 8.3% of workers initially in manufacturing. While a high share is expected for the directly affected population, it also suggests that relatively few non-manufacturing local workers were outcompeted for other jobs by displaced manufacturing workers.

Figure 5b shows that local manufacturing workers also suffer nearly all (95.5%) of local welfare losses, with a focal tract manufacturing worker losing the equivalent of $18,699 in earnings (Table 6). Local workers' welfare loss shares also increase more steeply with initial earnings than employment loss shares. Thus, for high-paid and manufacturing workers, losses are relatively more likely to consist of lost income, search costs, or lower amenities than lost employment. Local welfare loss shares also exceed those for employment for age 31-50 workers (49.1% vs. 45.8%), whose high initial retention rates suggest they give up especially large job-staying surpluses.

Figure A6 displays the change in employment rate and welfare by distance bin for both plant openings and closings. While the dramatic asymmetry in focal tract impacts dominates the comparison, beyond the focal tract the gains and losses from plant openings and closings exhibit similar magnitudes and decay rates, leading to very similar spatial patterns of incidence shares.

Table A9 displays separate estimates of cumulative shares of employment and welfare incidence within various worker subpopulations by distance. Differences in spatial decay rates are even more striking for closings than openings, in part due to the much greater local losses from closings. Only 10.6% and 14.4% of the shock-induced welfare losses incurred by unemployed and the lowest paid workers, respectively, accrue to those within 10 miles of the target tract, compared to 59.3% and 45.4% for the two highest paid quartiles, with corresponding differences in the geographic concentration of employment loss. Similarly, only 27.1% of welfare losses for workers under age 30 are incurred by those within 10 miles, compared to 44.3% and 49.6% for those 31-50 and over 50, and 82.0% of manufacturing workers' losses occur within 10 miles, compared to 8.8% for non-manufacturing. Part of this heterogeneity is because the shock becomes increasingly generic in both sectoral and skill demand composition as it ripples outward, so subgroups with greater local per-worker impact naturally have more locally concentrated distributions of aggregate incidence.

However, distant workers who are higher-paid, older, and already in manufacturing also tend to have considerably higher rates of job staying, suggesting that their job matches are creating large surpluses that generally insulate them from the shock.[48] By contrast, young and/or low-paid workers that frequently need or wish to switch jobs are harmed more by the reduction in their opportunities.

These stark differences in decay rates cause even stronger contrasts between subgroups' national and local shares of employment and welfare losses than for plant openings. As depicted in Figure 6a, low-paid workers and younger workers actually experience larger shares of national employment losses than higher-paid and older workers, even though high-paid and older workers are more likely to be initially employed at the closing plants. Indeed, among out-of-state workers, the bottom pay quartile is ten times more likely to endure shock-induced employment loss than the top quartile. Essentially, the high-paid and experienced workers outcompete low-paid and inexperienced workers for now scarcer positions, so that the employment incidence passes down the skill and experience ladder. Similarly, initially unemployed workers account for only 1.6% of local employment losses vs. 36.3% of national losses, as they tend to be the labor force's marginal workers. While national shares of welfare losses do increase with initial pay quartile (Figure 6b), the highest two quartiles' shares are much smaller nationally (39.1% and 25.9%) than locally (50.4% and 32.3%).

Most notably, workers from manufacturing bear only 13.5% and 45.4% of national employment and welfare loss versus 88.2% and 95.5% of within-tract losses. As discussed, this massive discrepancy partly reflects manufacturing's high baseline job staying rate, but it also reflects its tendency to hire non-manufacturing workers when turnover does occur: only 20% of their new hires in the sample come from other manufacturing firms, compared to around 30% in other supersectors.

Figures 7-9 (Table 7) reinforce this intuition by comparing plant closings among large high-paying manufacturing firms with "store closings" among large low-paying retail/wholesale firms. For focal tract workers (Fig. 7 and 8), the store closing creates a much larger per-worker employment rate decrease (1.1%) and share of local employment losses (39.6%) for the lowest-paid quartile than the highest-paid quartile (0.3% and 12.7%), since low-paid workers are both more targeted and less able to compete for other jobs. Local welfare losses from store closings are only slightly larger for the bottom two earnings quartiles ($3,920 and $4,313) than the top two ($2,663 and $3,431), since low-paid workers' greater exposure is partly offset by smaller baseline retention rates, so that more would have left jobs even without the shock. The low retention rate in retail/wholesale relative to manufacturing also explains why the overall average welfare loss among tract workers is smaller for the store closing ($3,134) than the plant closing ($5,624), since it suggests that retaining retail/wholesale jobs is less valuable to workers or firms (or both) than manufacturing jobs.

Since the retail shock also becomes generic with distance, the national shares of employment and welfare losses accruing to various worker subgroups differ far less across shock compositions than the local incidence would suggest (Figure 9). As with plant closings, much smaller shares of employment and welfare losses stay within the target industry at the national than local level

[48] These two forces outweigh the greater spatial mobility of high-paid workers conditional on switching jobs (Table 1).

(20.2% and 36.8% vs. 91.5% and 94.0%), and the national share of net employment loss borne by unemployed workers dwarfs the local share (39.1% vs. 1.2%). The gap in welfare shares for low vs. high paid workers is also attenuated (though not eliminated) nationally, and young workers' share of national welfare loss exceeds their workforce share, in contrast to the local level.

While quantifying the employment and utility incidence of negative labor demand shocks is important for allocating relief funds, policymakers and local communities also care about flows of workers away from sites suffering negative shocks. Thus, Figures A7a and A7b (Tab. A10) display the change in focal tract workers' probability of ending up employed in each distance bin. The probability of continuing to work in this tract only decreases by 4.5% and 3.6% for plant and store closings, even though the closings generally reduce total tract employment by around 10%. This is both because local workers retain a disproportionate share of remaining jobs relative to would-be job movers from afar, but also because a large minority of locals would have taken jobs elsewhere even without the shock. Indeed, less displacement for the store closing reflects retail's higher baseline turnover rate. An extra 0.7% of local workers become unemployed due to both closings, while an extra 0.8% (0.5%) move to other tracts in the PUMA after plant (store) closings, an extra 1.6 (1.9)% move to other within-state PUMAs, and an extra 1.6% (0.6%) find jobs out of state.

Figure A8 (Tab. A10) presents destination distributions by subpopulation. High earners are much more likely than low earners to find distant jobs following plant closings, with 87.7% of those induced to switch locations finding work in a different PUMA and 46.6% changing states, compared to 72.6% and 23.1% for the lowest earners (Panel A), reflecting their respective baseline tendencies to make such moves from Table 1. For the store closing (Panel B), which targets low earners, we see a large increase in their flows to unemployment, nearby tracts, and other PUMAs, but small flows out of state. This reflects low earners' less integrated labor markets, but also the fact that other opportunities in retail tend to be closer than in manufacturing. Although store closings caused more displacement for younger workers due to their greater presence in retail, a smaller share of their displacement leads to unemployment compared to older workers, who are less able or willing to move to more distant jobs. We also see a small additional outflow by local unemployed workers who would have found local jobs without the shock, illustrating the need to examine equilibrium reallocation rather than just the destinations of the initially laid-off workers.

Finally, Figure A9 (Table A11) shows how geographic incidence evolves as the shock size is scaled from 125 to 250 to 500 positions. For both plant openings and closings, the changes in employment rate and expected welfare scale nearly linearly with shock magnitude. Closings do exhibit a slight convexity in local employment rate changes with scale, as the focal tract workers' share of employment losses rises from 7.5% to 8.6% to 9.9% for the three shocks. For smaller closings, local workers disproportionately retain the remaining jobs at the expense of distant workers who would have been hired in the shock's absence but whose matches with the focal tract create less surplus. As shock size grows, the local workers become the marginally employed workers. By contrast, the local share of welfare gains is slightly concave in shock size, since larger shocks cause enough of an exodus to meaningfully affect labor supply to more distant areas.

6.5 Heterogeneity in Incidence by Focal Tract Characteristics

Heterogeneity in geographic incidence also stems from the choice of focal tract. Among the 300 tracts receiving shocks, Figure A10 (Tables A12 and A13) provides the mean employment and welfare incidence within the top and bottom quintiles of population density, # of jobs within 5 miles, rent for an average two-bedroom apartment and poverty rates.

Both welfare and employment gains are more geographically concentrated for tracts with lower population density. The expected utility gain for workers within the focal tract or 1, 2, or 3+ tracts away within the PUMA are all several times larger for the most rural relative to the most urban focal tracts ($805 vs. $216, $239 vs. $23, $90 vs. $16, and $37 vs. $14, respectively). The differences in welfare gains among nearby workers are even larger for tracts featuring few vs. many jobs within 5 miles (e.g. $878 vs. $132 for focal tract workers). These differences partly stem from the fact that 250 new jobs is a larger per-worker shock to low density areas, which tend to have fewer workers in the focal and surrounding tracts. However, substantial density-based differences also exist in within-PUMA shares of welfare and employment gains, so that larger per-worker gains in low density areas more than offset smaller labor force shares: the average share of welfare (employment) gains enjoyed by within-PUMA workers is 15.2% (8.8%) among the 60 most rural tracts versus 5.4% (3.7%) for the 60 most urban tracts (and 9.9% (5.7%) among all selected tracts). Combining the nearly linear relationship between shock size and average impact with the urban/rural differences in local concentration of incidence, the results suggest that targeting several rural areas with small development initiatives might generate larger local employment and welfare gains per job created than a single large plant opening in a dense urban area (barring large job multiplier differences).

Comparisons for tracts with low vs. high average two-bedroom rent closely mirror the rural/urban results. Since low rent may indicate a high housing supply elasticity, the job-related welfare gains may better approximate total welfare gains for such tracts. High-poverty tracts exhibit larger local welfare gains and within-PUMA shares of gains, suggesting that targeting local initiatives at poorer areas may yield greater local labor market benefits than for a typical tract.

Since residential sorting leads to high correlations among many tract characteristics, Table A14 reports the results of a set of regressions that relate various measures of shock incidence to a broader set of focal tract characteristics, where each has been standardized to have zero mean and unit s.d. to ease coefficient comparability. To improve power, the sample here consists of 3,200 plant opening simulations with 250 new jobs at large, high-paying manufacturing firms but different focal tracts.

Columns 2-5 confirm that the unconditional relationships from Tables A12 and A13 survive as partial correlations: one s.d. increases in two-bedroom rent and population density still predict lower average welfare gains ($21 and $3, respectively) and shares of total welfare gains (2.8% and 1.2%) for target PUMA residents even conditional on other tract characteristics. Similarly, lower median household income, higher poverty rates, and particularly fewer jobs within 5 miles ($14) all predict larger within-PUMA welfare gains, with the latter more strongly predicting local incidence than focal tract job density. A one s.d. (3.2 pp) increase in the PUMA's share of manufacturing workers only predicts small increases in the within-PUMA share of welfare gains (0.68%) and especially employment gains (0.04%), consistent with shocks becoming generic within quite a narrow range.

Columns 6-9 focus more narrowly on employment and welfare gains for low-paid within-PUMA workers, and show similar patterns, but with larger coefficient magnitudes for employment and smaller for welfare, consistent with earlier initial earnings incidence results. However, column 10, which examines employment gain shares among all U.S. low-paid workers, reveals another local vs. national discrepancy: tract characteristics that predict greater employment gains for local low-paid workers tend to predict smaller gains for low-paid workers nationwide.

Thus, reduced-form estimates of larger local treatment effects for low-paid workers that rely on classifying distant areas as "untreated" could cause incorrect inferences about which focal areas would best alleviate poverty, since larger gains for the local poor in certain local areas captured (and slightly overstated!) by such regressions would be outweighed by smaller expected gains among many less proximate workers. One possible explanation is that these characteristics may predict higher search costs that cause firms to hire local low-paid workers rather than more distant low-paid or even high-paid workers (since the jobs they vacate may be taken by their lower-paid neighbors).

To test the importance of mismatch between the skills of local workers and those required by the new jobs, Table A15 mimics Table A14 but replaces "plant openings" with large low-paying retail "store openings". Evidence of a role for mismatch is fairly mild: focusing on incidence for low-paid local workers (col. 6-9), the employment coefficients on poverty rate and median income increase and decrease by about 20% from Table A14, respectively, while impacts on welfare gains and shares are inconsistent. Changing the shock's firm composition also minimally affects how focal tract characteristics predict low-paid workers' share of national employment and welfare gains.

Finally, focusing on contrasts among observed tract characteristics masks additional unexplained heterogeneity in incidence among alternative focal tracts. For each shock specification, the within-PUMA shares of employment gains range from below 2% to above 10% and the within-state shares (partly driven by state size) range from below 15% to above 55%, though these ranges may partly reflect sampling error. Shares of welfare gains display even greater variation: within-PUMA shares range from 2% to over 20% and within-state shares range from 41% to 83%.

6.6 Robustness Checks

Table A16 examines sensitivity to alternative model assumptions of the baseline geographic incidence predictions from a standard 250 job "plant opening" (col. 1). Columns 2 and 6 display employment and earnings results from a model featuring job multipliers. Specifically, we adopt Bartik and Sotherland (2019)'s estimate that each new high-tech manufacturing job (presumably at large, high paying firms) generates an extra 0.71 jobs after one year. While this estimate captures the net effect of all spillover sources, we assume that increased product demand for local services is the dominant source. Thus, we add 250*.71 = 171 additional retail/wholesale and leisure/hospitality jobs, distributed across within-PUMA tracts in proportion to their workers' shares of expected earnings gains from the baseline results. The augmented shock increases average employment and welfare gains within the PUMA by only slightly more than the 171% multiplier, albeit with modest shifts in the shares of employment and welfare gains toward surrounding tracts and away from the target tract. These results indicate that explicitly introducing job multipliers rather than treating the simulated shocks as implicitly post-multiplier would not alter the paper's key findings.

Columns 3 and 7 display results from a specification that allows firms to endogenously update their desired number of positions in response to shock-induced changes in labor costs. We assume a constant elasticity of demanded positions with respect to changes in each position type's expected per-position payoff ($\overline{q}_f$), and assign a value of -0.197 based on the mean short-run employment elasticity estimate from Lichter et al. (2015)'s meta-analysis of the minimum wage literature. We then iterate between 1) computing equilibrium assignments and payoffs given a vector $h^{CF}(f)$ of position counts by type and 2) updating $h^{CF}(f)$ for each type by applying the elasticity to $\%\Delta\overline{q}_f$. We include a fixed cost of adjusting the stock of positions equal to 1% of average earnings to prevent fractional worker adjustments by a large share of firms. This process converges to a fixed point in which the final vector $h^{CF}(f)$ aligns with firms' optimal position count given their expected payoffs from filling a position. Across 300 simulations with different focal tracts, the mean adjustment reduces the shock size from 250 to 246 positions, with a standard deviation of 7. This adjustment slightly decreases the magnitude of employment and welfare gains, but, as with job multipliers, it barely changes the shares of gains by distance bin. Thus, explicitly incorporating endogenous responses in desired employment does not substantively alter our main findings.

Columns 4 and 8 display results from a specification that adopts the Choo-Siow structure of unobserved surplus components, which includes both worker $\times$ position type and worker type $\times$ position components ($\epsilon^1_{if(k)} + \epsilon^2_{l(i),k}$) rather than a single worker $\times$ position component (ϵ_{ik}). This approach assumes perfect rather than zero correlation in individuals' preferences for positions within firm types and vice versa. The geographic distribution of employment rate changes and gain shares among workers are surprisingly similar to their baseline counterparts, reflecting very similar worker reallocation following shocks. The CS specification generates slightly smaller employment and slightly larger welfare gains for local tract workers, with a slightly slower rate of decay with distance. This results in 5.2% (10.0%) of employment (welfare) gains accruing to workers within 10 miles and 21.2% (41.3%) accruing to workers within 250 miles, compared to 5.6% (10.8%) and 23.5% (43.6%) for the baseline specification. Thus, the model's incidence predictions seem quite insensitive to assumptions about within-type correlation in surplus components.

Columns 5 and 9 examine sensitivity to allowing the plant opening to change relative joint surplus values among job matches featuring within-PUMA worker and firm types, perhaps due to differential unmodeled adjustments in perceived expectations about local labor market dynamics. We estimate typical surplus changes by finding the median realized surplus change per job created for each such group g among our model validation sample of actual establishment openings (described in the next section), and re-scale to match a 250 job opening. These changes are then built into the simulated shock. This specification produces 25% larger average within-PUMA welfare gains, suggesting that large, high-paying manufacturing openings may benefit nearby workers somewhat more than pre-estimated surpluses would predict, perhaps due to anticipated openings by upstream suppliers. However, the same exercise for large high-paying retail or PBS openings produces welfare changes that are only 5% smaller and 1% larger, respectively, than their baseline counterparts, perhaps because such establishments are more likely to compete with existing within-PUMA businesses. This suggests that the limited changes in welfare incidence from incorporating dynamic considerations may be less relevant outside manufacturing, at least for short-run analyses.

Table A17 assesses sensitivity of model predictions to restricting surplus heterogeneity in various ways. Here we focus on local welfare changes across initial earnings and industry categories, where worker and firm heterogeneity was shown to matter most. In column 2, we equalize joint surplus values across all categories of firm industry, size, and average pay, so that location is the only firm characteristic. Because this model ignores complementarity between high-skilled workers and high-paying firms, it mistakenly predicts that local welfare gains will be larger for unemployed and low-paid workers whose low baseline retention rate suggests they are more open to alternative job opportunities, even when the shock features high-paying firms. Analogously, column 3 removes surplus variation among categories of all worker characteristics except initial location. This eliminates variation in local incidence by earnings categories except to the extent that initial earnings predicts welfare-relevant tract characteristics. Column 4 removes the surplus premium from moving/hiring within the same industry, conditional on switching firms. This halves the shock-induced welfare gain for same-industry workers, thus understating the concentration of local welfare gains. Finally, column 5 removes the surplus premium from job staying/retention. In this case, new jobs immediately create the same surplus as existing jobs, essentially ignoring any within-tract recruiting, search, and training costs. This produces enormous local welfare gains that mimic the losses from plant closings. These results show that the full extent of two-sided heterogeneity in the baseline model is needed to generate the disparities in local welfare gains presented above.

6.7 Model Validation

The estimated surplus parameters $\hat{\Theta}^{D-in-D}$ that underlie the simulations are identified from millions of quotidian job transitions driven by small firm expansions/contractions, labor force turnover, and preference or skill changes over the life cycle that cause considerable offsetting churn in the U.S. labor market. Thus, one might wonder whether parameters governing ordinary worker flows can capture the response to sizable, locally focused positive or negative shocks. To address this concern, we perform a model validation exercise in which surplus parameters estimated on pre-shock worker flows are used to forecast worker reallocation after actual observed local demand shocks. We evaluate model fit using the index of dissimilarity between the predicted and actual match group distributions $P(g)$ among affected workers, defined as those initially or most recently working in the target PUMA. We average this index across 421 shocks defined by tract-years that feature 1) a single opening or closing establishment with at least 100 workers; 2) a net change in total tract employment in the same direction of at least 100 workers and 10% of its pre-shock employment; 3) no offsetting contemporaneous shocks to the PUMA's other tracts; and 4) no qualifying shocks to the same tract in other years. Appendix A8 offers further detail, while Table 8 reports the results.

To summarize, on average the model would need to reallocate 35.1% of job matches of workers originating in the target PUMA to other groups g to perfectly match the true within-PUMA distribution. However, most "incorrect" predictions involve either slight differences in destination tract within the same distance bin or slightly mismatched firm size/avg. pay/sector cells.

When the group space is collapsed post-simulation so that worker and position locations are defined by 14 distance bins from the target tract, the share of job matches that must be reallocated across groups falls to 11.1% (row 2), and collapsing non-location position characteristics (and retaining all worker characteristic categories) pushes the necessary reallocation rate to 2.3% (row 3). This is despite the fact that $P(g)$ still contains 1,500 groups with only 155 restrictions imposed by $n(l)$ and $h(f)$. The model also fits well the worker and position type distributions among workers who either enter or exit unemployment after the shock (row 4), particularly when locations are aggregated to distance bins (row 5), where only 0.95% of within-PUMA workers' job matches require reassignment to match the actual allocation. This suggests that the counterfactual forecasts of employment incidence among demographic/distance bin combinations are likely to be accurate.

Furthermore, the assignment model vastly outperforms a one-sided parametric conditional logit model fit to the same pre-shock CCPs $P(g|f)$. Thus, with many million observed job matches, it appears that the risk of overfitting from using a highly saturated, just-identified model is far outweighed by the inability of a more parsimonious parametric model (still featuring $\sim$ 200 parameters!) to capture the rich multidimensional matching patterns contained in the data. The two-sided model also outperforms (though by much less) other one-sided nonparametric forecasts that hold fixed the full set of either raw or smoothed CCPs (so $P(g)^{y,CF} = h^y(f)P^{y-1}(g|f)$). This suggests that requiring market clearing does have additional predictive value, even for smallish shocks. The baseline model also outperforms the Choo-Siow model, which assumes perfect rather than zero correlation in workers' preferences for positions within position types, particularly for more aggregated predictions. The baseline model also generates much more accurate predictions than the alternatives from Table A17 that restrict surplus heterogeneity across worker types, firm types, or mover/stayer status. Taken together, the model predicts pretty well the reallocation of workers across job types and particularly employment statuses that follows substantial local labor demand shocks.

7 Conclusion

This paper models the U.S. labor market as a large-scale assignment game with transferable utility, and uses the model estimates to simulate the employment and welfare incidence across locations and worker demographic categories of a variety of local labor demand shocks representing different local development initiatives and establishment openings or closings.

We find that U.S. labor markets are quite local, in that the per-worker employment and welfare gains from a locally targeted labor demand shock are substantially larger for workers in the focal and adjacent census tracts than even for workers several tracts away. Nonetheless, because these very local workers are a tiny share of the U.S. labor force competing for positions, we also find that, regardless of establishment composition, around 62% (36%) of the employment (welfare) gain from a large establishment opening redounds to workers initially working out of state, with only around 6% (11%) going to existing workers within 10 miles of the focal tract.

We also document a high degree of heterogeneity in incidence by initial earnings, age, and initial industry among very local workers across demand shocks with different establishment composition and/or different focal tract attributes, suggesting that the type of establishment and community targeted by a local development policy has major implications for the groups of workers most likely to benefit. That said, as these alternative shocks ripple across space through a chain of job transitions, their incidence across worker subgroups becomes increasingly similar, so that the overall demographic and spatial composition of worker welfare gains slightly farther from the site is extremely similar across different types of shocks and target areas. Thus, state-level funders of local projects who internalize these ripple effects can safely devolve the selection of local projects to local leaders.

These findings demonstrate both the value and the limitations of reduced-form research analyzing place-based policies. The simulation results suggest that per-person employment and welfare impacts of local labor demand shocks become quite small at greater distances, so that research designs treating distant but similar locations as control groups may be valid for estimating treatment effects on local populations. However, the results also indicate that the distribution of local impacts need not resemble the distribution of state-level or national impacts. In fact, some worker subgroups that receive disproportionate shares of local impacts are comparatively insulated nationally.

We also find that negative shocks produce a much greater concentration of employment and welfare losses than the corresponding gains from equally-sized positive shocks. This is because many local workers would have been working anyway without a positive shock, but have jobs at risk from negative local shocks, and removing the option to keep one's job generates large welfare losses, presumably due to both job switching costs and the loss of firm-specific experience.

Methodologically, we show that one can still produce forecasts of welfare incidence on both sides of the market from changes in either side's composition even when singles are either not observed or observed on only one side. By basing simulations on millions of composite joint surplus parameters rather than a much smaller set of fundamental utility or production function parameters, the sufficient statistics approach used here can fully exploit the massive scale of the LEHD data to capture multidimensional heterogeneity on both sides of a two-sided market without placing unjustified structure on the job matching technology. Given appropriate matching data, the approach here could also be easily adapted to the student-college or patient-doctor contexts, among others.

References

Abowd, John M, Bryce E Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L McKinney, Marc Roemer, and Simon Woodcock, "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators," in "Producer Dynamics: New Evidence from Micro Data," University of Chicago Press, 2009, pp. 149–230.

Arcidiacono, Peter and Robert A Miller, "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Econometrica*, 2011, *79* (6), 1823–1867.

Bartik, Timothy J., "Economic Development," in J. Richard Aronson and Eli Schwartz, eds., *Management Policies in Local Government Finance*, 2004.

Bartik, Timothy J and Nathan Sotherland, "Local Job Multipliers in the United States: Variation with Local Characteristics and with High-Tech Shocks," 2019.

Bayer, Patrick, Stephen L Ross, and Giorgio Topa, "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 2008, *116* (6), 1150–1196.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 2019, *87* (3), 699–739.

Busso, Matias, Jesse Gregory, and Patrick Kline, "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *The American Economic Review*, 2013, *103* (2), 897–947.

Cadena, Brian C and Brian K Kovak, "Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession," *American Economic Journal: Applied Economics*, 2016, *8* (1), 257–90.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 2019, *87* (3), 741–835.

Carballo, Jeronimo and Richard Mansfield, "What Drives the Labor Market Incidence of Trade Shocks?: An Equilibrium Matching Analysis of China's WTO Accession," 2023. Working Paper.

Chen, Liang, "Compensation, Moral Hazard, and Talent Misallocation in the Market for CEOs," 2017.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *The Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162.

Chiappori, Pierre-André and Bernard Salanié, "The Econometrics of Matching Models," *Journal of Economic Literature*, September 2016, *54* (3), 832–61.

Chiappori, Pierre-Andre, Murat Iyigun, and Yoram Weiss, "Investment in Schooling and the Marriage Market," *The American Economic Review*, 2009, *99* (5), 1689–1713.

Choo, Eugene, "Dynamic Marriage Matching: An Empirical Framework," *Econometrica*, 2015, *83* (4), 1373–1423.

— **and Aloysius Siow**, "Who Marries Whom and Why," *Journal of Political Economy*, 2006, *114* (1), 175–201.

Decker, Colin, Elliott H Lieb, Robert J McCann, and Benjamin K Stephens, "Unique Equilibria and Substitution Effects in a Stochastic Model of the Marriage Market," *Journal of Economic Theory*, 2013, *148* (2), 778–792.

Diamond, Rebecca, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 2016, *106* (3), 479–524.

Dingel, Jonathan I and Felix Tintelnot, "Spatial Economics for Granular Settings," Technical Report, National Bureau of Economic Research 2020.

Fogel, Jamie and Bernardo Modenesi, "What Is a Labor Market? Classifying Workers and Jobs Using Network Theory," 2021.

Galichon, Alfred and Bernard Salanié, "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," *The Review of Economic Studies*, 2022, *89* (5), 2600–2629.

_ , _ et al., "The Econometrics and Some Properties of Separable Matching Models," *American Economic Review*, 2017, *107* (5), 251–255.

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

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

Gutierrez, Federico H, "A Simple Solution to the Problem of Independence of Irrelevant Alternatives in Choo and Siow Marriage Market Model," *Economics Letters*, 2020, *186*, 108785.

Hillier, Frederick S and Gerald J Lieberman, *Introduction to Operations Research - Ninth Edition*, New York, USA: McGraw-Hill Higher Education, 2010.

Hornbeck, Richard and Enrico Moretti, "Estimating Who Benefits from Productivity Growth: Local and Distant Effects of City Productivity Growth on Wages, Rents, and Inequality," *Review of Economics and Statistics*, 2024, *106* (3), 587–607.

Hotz, V Joseph and Robert A Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 1993, *60* (3), 497–529.

Hyatt, Henry R, Erika McEntarfer, Ken Ueda, and Alexandria Zhang, "Interstate Migration and Employer-to-Employer Transitions in the US: New Evidence from Administrative Records Data," *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-44*, 2016.

Kline, Patrick and Enrico Moretti, "Place Based policies with Unemployment," *American Economic Review*, 2013, *103* (3), 238–43.

_ and _ , "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 2014, *129* (1), 275–331.

Koopmans, Tjalling C and Martin Beckmann, "Assignment Problems and the Location of Economic Activities," *Econometrica*, 1957, pp. 53–76.

Lichter, Andreas, Andreas Peichl, and Sebastian Siegloch, "The Own-Wage Elasticity of Labor Demand: A Meta-Regression Analysis," *European Economic Review*, 2015, *80*, 94–119.

Lindenlaub, Ilse, "Sorting Multidimensional Types: Theory and Application," *The Review of Economic Studies*, 2017, *84* (2), 718–789.

Malamud, Ofer and Abigail Wozniak, "The Impact of College on Migration Evidence from the Vietnam Generation," *Journal of Human Resources*, 2012, *47* (4), 913–950.

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

Marinescu, Ioana and Roland Rathelot, "Mismatch Unemployment and the Geography of Job Search," *American Economic Journal: Macroeconomics*, 2018, *10* (3), 42–70.

Menzel, Konrad, "Large Matching Markets as Two-Sided Demand Systems," *Econometrica*, 2015, *83* (3), 897–941.

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 2018, *108* (12), 3855–90.

Moretti, Enrico, "Local Multipliers," *The American Economic Review*, 2010, pp. 373–377.

Mourifié, Ismael, "A Marriage Matching Function with Flexible Spillover and Substitution Patterns," *Economic Theory*, 2019, *67* (2), 421–461.

_ **and Aloysius Siow**, "The Cobb-Douglas Marriage Matching Function: Marriage Matching with Peer and Scale Effects," *Journal of Labor Economics*, 2021, *39* (S1), S239–S274.

Neumark, David and Helen Simpson, "Place-Based Policies," in "Handbook of Regional and Urban Economics," Vol. 5, Elsevier, 2015, pp. 1197–1287.

Nimczik, Jan Sebastian, "Job Mobility Networks and Endogenous Labor Markets," 2018.

Piyapromdee, Suphanit, "The Impact of Immigration on Wages, Internal Migration, and Welfare," *The Review of Economic Studies*, 2021, *88* (1), 406–453.

Roback, Jennifer, "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 1982, *90* (6), 1257–1278.

Roth, Alvin E and Marilda Sotomayor, "Two-Sided Matching," *Handbook of Game Theory with Economic Applications*, 1992, *1*, 485–541.

Sattinger, Michael, "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 1993, *31* (2), 831–880.

Schmutz, Benoît and Modibo Sidibé, "Frictional Labour Mobility," *The Review of Economic Studies*, 2019, *86* (4), 1779–1826.

Shapley, LS and Martin Shubik, *Game Theory in Economics: Chapter 3: the" Rules of the Games"*, Rand, 1972.

Sprung-Keyser, Ben, Nathaniel Hendren, Sonya Porter et al., *The Radius of Economic Opportunity: Evidence from Migration and Local Labor Markets*, US Census Bureau, Center for Economic Studies, 2022.

Tervio, Marko, "The Difference that CEOs Make: An Assignment Model Approach," *American Economic Review*, 2008, *98* (3), 642–68.

Vilhuber, Lars et al., "LEHD Infrastructure S2014 Files in the FSRDC," *US Census Bureau, Center for Economic Studies Discussion Papers, CES*, 2018.

Table 1: Summary Statistics Describing Heterogeneity in the Spacial Scope of Labor Markets by Worker and Establishment Characteristics

Panel A: By Worker Earnings or Age Category

		Share of All Transitions						Share of Job to Job Transitions					
Worker Subpop.	% of Pop.	Unemp. to Unemp.	Unemp. to Emp.	Emp. to Unemp.	Stay at Same Job	Same Ind.	Diff. Ind.	Same PUMA	New PUMA, Same State	New State	$<$ 10 Miles	10-250 Miles	$>$250 Miles
All		0.028	0.093	0.028	0.695	0.073	0.083	0.277	0.576	0.148	0.303	0.517	0.180
Unemployed	0.120	0.229	0.771					0.288	0.618	0.095	0.315	0.552	0.132
1st Earn. Q.	0.220			0.056	0.701	0.100	0.142	0.303	0.540	0.157	0.321	0.497	0.182
2nd Earn. Q.	0.220			0.032	0.790	0.082	0.097	0.287	0.558	0.155	0.309	0.514	0.176
3rd Earn. Q.	0.220			0.022	0.829	0.076	0.074	0.258	0.562	0.180	0.289	0.505	0.206
4th Earn. Q.	0.220			0.016	0.843	0.075	0.066	0.217	0.540	0.244	0.264	0.447	0.289
Age $<$ 30	0.310	0.028	0.181	0.042	0.525	0.093	0.131	0.266	0.568	0.165	0.292	0.511	0.197
Age 31-50	0.426	0.028	0.061	0.023	0.742	0.073	0.072	0.265	0.553	0.182	0.299	0.490	0.211
Age $>$50	0.265	0.026	0.041	0.018	0.820	0.049	0.045	0.278	0.554	0.168	0.304	0.497	0.199

Panel B: By Destination Establishment Pay Quartile and Size Quartile

		Share of All Transitions						Share of Job to Job Transitions					
Estab. Subpop.	% of Pop.	Unemp. to Unemp.	Unemp. to Emp.	Emp. to Unemp.	Stay at Same Job	Same Ind.	Diff. Ind.	Same PUMA	New PUMA, Same State	New State	$<$ 10 Miles	10-250 Miles	$>$250 Miles
FE Quartiles 1 & 2	0.519		0.141		0.680	0.083	0.095	0.290	0.545	0.165	0.301	0.507	0.192
FE Quartile 3	0.241		0.059		0.791	0.070	0.080	0.269	0.556	0.175	0.296	0.505	0.199
FE Quartile 4	0.240		0.045		0.801	0.073	0.081	0.222	0.558	0.221	0.288	0.448	0.264
FS $<$ Median	0.514		0.117		0.699	0.086	0.099	0.308	0.505	0.187	0.332	0.472	0.197
FS $>$ Median	0.486		0.079		0.775	0.069	0.077	0.219	0.610	0.172	0.252	0.523	0.224

Panel C: By Destination Establishment Industry

		Share of All Transitions						Share of Job to Job Transitions					
Estab. Industry	% of Pop.	Unemp. to Unemp.	Unemp. to Emp.	Emp. to Unemp.	Stay at Same Job	Same Ind.	Diff. Ind.	Same PUMA	New PUMA, Same State	New State	$<$ 10 Miles	10-250 Miles	$>$250 Miles
Nat. Resources	0.018		0.132		0.686	0.077	0.105	0.386	0.391	0.224	0.192	0.561	0.248
Construction	0.049		0.113		0.687	0.093	0.107	0.242	0.535	0.223	0.247	0.531	0.222
Manufacturing	0.089		0.054		0.826	0.036	0.084	0.339	0.490	0.172	0.296	0.518	0.187
Wholesale/Retail	0.204		0.107		0.732	0.078	0.083	0.234	0.570	0.196	0.251	0.522	0.228
Information	0.023		0.070		0.750	0.060	0.120	0.226	0.585	0.190	0.320	0.434	0.246
Financial Activities	0.059		0.062		0.758	0.075	0.105	0.237	0.601	0.162	0.297	0.493	0.211
Prof. Bus. Services	0.143		0.118		0.662	0.091	0.129	0.228	0.584	0.189	0.281	0.478	0.242
Ed. & Health	0.224		0.070		0.792	0.081	0.058	0.308	0.537	0.155	0.344	0.487	0.169
Leis. & Hosp.	0.113		0.182		0.616	0.117	0.086	0.298	0.525	0.177	0.336	0.468	0.196
Oth. Serv.	0.031		0.121		0.714	0.042	0.122	0.301	0.531	0.168	0.353	0.458	0.190
Government	0.047		0.036		0.880	0.024	0.060	0.344	0.544	0.112	0.319	0.520	0.162

Notes: "Unemployed": Workers who were unemployed in the prior year. "Earn. Q.": Workers in the chosen quartile of the distribution of annualized earnings based on pro-rating earnings in full quarters. "FE Quartile": Firms (SEINs) in the chosen quartile of the (worker-weighted) firm distribution of per-worker annual earnings. "FS $<$($>$) Median": Firms below (above) the median of the worker-weighted firm employment distribution. *: For initially unemployed workers, the share of unemployment-to-employment transitions by distance category is reported in place of share of job-to-job transitions. The locations of initially unemployed workers are assumed to be the location of their most recent employer if previously observed working, otherwise they are imputed from the conditional distribution among job-to-job transitions of origin locations given the destination employer location.

"Nat. Resources": Natural Resources. "Wholesale/Retail": Wholesale/Retail Trade and Transportation. "Prof. Bus. Services": Professional & Business Services. "Ed. & Health": Education and Healthcare. "Leis. & Hosp.": Leisure and Hospitality. "Oth. Serv.": Other Services (includes repair, laundry, security, personal services).

Table 2: Expected Employment and Welfare Gains From New Stimulus Positions Among Workers in Different Subpopulations Initially Employed in the Focal Tract by Industry Supersector of the Stimulus Package (Averaged Across Firm Size/Firm Average Earnings Combinations)

Panel A: Change in P(Employed)

Worker		Industry							
Category	Avg.	Info.	Manu.	R/W Trd.	Prof. Bus.	Ed./Hlth	Lei/Hosp.	Gov.	Oth. Serv.
All	0.0009	0.0008	0.0009	0.0008	0.0007	0.0012	0.0010	0.0009	0.0009
Unemployment	0.0034	0.0031	0.0034	0.0031	0.0026	0.0041	0.0034	0.0037	0.0036
1st Earn Q.	0.0009	0.0007	0.0008	0.0008	0.0007	0.0012	0.0010	0.0008	0.0009
2nd Earn Q.	0.0005	0.0004	0.0005	0.0004	0.0004	0.0007	0.0005	0.0005	0.0004
3rd Earn Q.	0.0004	0.0003	0.0004	0.0003	0.0004	0.0005	0.0003	0.0003	0.0003
4th Earn Q.	0.0003	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002
Age $\leq$ 30	0.0011	0.0011	0.0010	0.0011	0.0009	0.0013	0.0015	0.0011	0.0011
Age 31-50	0.0009	0.0007	0.0009	0.0007	0.0007	0.0013	0.0008	0.0009	0.0010
Age $>$ 50	0.0007	0.0004	0.0007	0.0005	0.0006	0.0010	0.0006	0.0008	0.0006
Diff. Ind.	0.0009	0.0007	0.0009	0.0007	0.0007	0.0013	0.0008	0.0009	0.0009
Same Ind.	0.0025	0.0034	0.0023	0.0014	0.0027	0.0013	0.0021	0.0031	0.0042

Panel B: Average Welfare Gain (Scaled in 2023 $

Worker		Industry							
Category	Avg.	Info.	Manu.	R/W Trd.	Prof. Bus.	Ed./Hlth	Lei/Hosp.	Gov.	Oth. Serv.
All	322	255	326	285	260	463	342	332	310
Unemployment	437	417	405	421	353	508	464	455	474
1st Earn Q.	322	237	281	282	249	492	379	325	331
2nd Earn Q.	295	219	301	251	228	471	304	326	260
3rd Earn Q.	306	237	346	258	266	434	296	318	291
4th Earn Q.	341	266	399	339	288	455	356	310	319
Age $\leq$ 30	346	306	341	326	279	421	442	329	323
Age 31-50	321	240	339	274	258	505	302	332	317
Age $>$ 50	298	217	295	260	252	464	275	339	285
Diff. Ind.	268	242	279	210	218	352	271	293	283
Same Ind.	1348	1906	1286	865	1220	942	912	2072	1583

Notes: Each cell in Panel A (Panel B) contains the increase in probability of being employed (average welfare gain) generated by a 250 job stimulus for workers initially employed in the previous year (or most recently employed) in the focal tract whose belong to the worker subpopulation defined by the row label. Each column averages results across four stimulus packages featuring jobs with establishments in the same industry supersector but in different categories of the establishment-level employment and average worker earnings distributions. Results are further averaged across 300 simulations featuring different target census tracts for each of the stimulus package specifications. See 1 for expanded definitions of the demographic groups and industries in the row and column labels.

Table 3: Expected Change in Employment Probability and Utility From New Stimulus Positions Among Workers Initially Employed in the Focal Tract among Different Worker Subpopulations by Firm Size Quartile/Firm Average Pay Quartile Combination (Averaged Across Industry Supersectors)

Worker	Change in P(Employed)				Avg. Welfare Gain (2023 $)			
Category	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi
All	0.0010	0.0010	0.0008	0.0007	325	330	320	311
Unemployment	0.0042	0.0036	0.0032	0.0025	539	513	384	313
1st Earn Q.	0.0010	0.0010	0.0007	0.0007	373	371	277	267
2nd Earn Q.	0.0005	0.0005	0.0004	0.0004	294	303	286	296
3rd Earn Q.	0.0003	0.0003	0.0004	0.0004	268	277	348	330
4th Earn Q.	0.0002	0.0002	0.0003	0.0003	235	243	438	449
Age $\leq$ 30	0.0013	0.0014	0.0010	0.0009	355	382	323	323
Age 31-50	0.0010	0.0009	0.0009	0.0007	314	309	337	323
Age $>$ 50	0.0008	0.0006	0.0007	0.0005	312	298	299	285
Diff. Ind.	0.0010	0.0009	0.0008	0.0007	279	276	268	250
Same Ind.	0.0034	0.0023	0.0023	0.0022	1578	1279	1321	1216

Notes: See Table 2 for expanded definitions of worker subpopulations defined by the row labels. The cells in the first four (next four) columns contain the change in employment probability (average job-related welfare gain, scaled to be equivalent to $ of 2023 annual earnings) generated by a 250 job stimulus for workers employed in the previous year (or most recently employed) in the focal tract who belong to the worker subpopulation listed by the row label. Each column averages results from eight stimuli that feature jobs with establishments from different industry supersectors but the same quartiles of the establishment-level employment and average worker earnings distributions (indicated by the column label). Results are further averaged across 300 simulations featuring different target census tracts for each of the stimulus package specifications. "Sm./Low": The 250 stimulus jobs are generated by establishments whose employment levels and average worker pay levels are below the respective worker-weighted medians among all firms. "Lg./Low": The 250 stimulus jobs are generated by establishments whose employment levels place them above the worker-weighted median among all firms and whose average worker pay levels place them below the worker-weighted median among all firms. "Sm./Hi": The 250 stimulus jobs are generated by establishments whose employment levels place them below the worker-weighted median among all firms and whose average worker pay levels place them in the highest quartile of firms. "Lg./Hi": The 250 stimulus jobs are generated by establishments whose employment levels place them above the worker-weighted median among all firms and whose average worker pay levels place them in the highest quartile of firms.

Table 4: Cumulative Share of Employment Gains within Bins of Distance from Focal Tract due to Stimulus among Subpopulations Defined by Initial Earnings, Age, and Initial Industry: Average Across All Stimulus Specifications Featuring 250 New Jobs

Distance from	Employment Status/Earnings Quartile				
Focal Tract	Unemp.	1st Q.	2nd Q.	3rd Q.	4th Q.
Target Tract	0.007	0.006	0.005	0.004	0.003
1 Tct Away	0.018	0.014	0.013	0.011	0.008
2 Tcts Away	0.032	0.025	0.022	0.021	0.015
3+ Tcts w/in PUMA	0.070	0.054	0.050	0.047	0.036
1 PUMA Away	0.103	0.081	0.075	0.072	0.055
2 PUMAs Away	0.161	0.129	0.122	0.118	0.093
3+ PUMAs w/in State	0.446	0.339	0.325	0.333	0.303
1 State Away	0.517	0.413	0.401	0.408	0.362
2+ States Away	0.658	0.561	0.546	0.551	0.479
Out of Sample	1	1	1	1	1
	Age			Industry Status	
	Age $<$ 30	Age 31-50	Age $>$50	Diff. Ind.	Same Ind.
Target Tract	0.006	0.005	0.005	0.005	0.011
1 Tct Away	0.016	0.014	0.014	0.014	0.021
2 Tcts Away	0.028	0.026	0.024	0.026	0.034
3+ Tcts w/in PUMA	0.062	0.056	0.054	0.057	0.068
1 PUMA Away	0.092	0.084	0.080	0.086	0.101
2 PUMAs Away	0.147	0.133	0.128	0.136	0.155
3+ PUMAs w/in State	0.403	0.370	0.353	0.380	0.379
1 State Away	0.478	0.441	0.421	0.452	0.451
2+ States Away	0.624	0.580	0.557	0.593	0.588
Out of Sample	1	1	1	1	1

Notes: See Table A3 for expanded definitions of the row labels. Each cell contains the share of employment gains in the subsequent year caused by a 250 job stimulus accruing to workers whose initial establishment's distance from the targeted census tract is closer than or within the row label's distance bin among those whose baseline age, earnings, or industry category matches the column label. "Unemp": Initially unemployed workers (no job with $<\$2,000$ in earnings). "1st/2nd/3rd/4th Q.": workers' baseline quartile in the 2012 annualized earnings distribution among dominant jobs. "Same (Diff) Ind.": Workers whose baseline industry is the same as (different than) the simulated job creation. Each cell averages results across 300 simulations with different target census tracts for each of 32 stimulus packages of new jobs in establishments with different combinations of industry supersector, firm size quartile, and firm average pay quartile.

Table 5: Cumulative Share of Utility Gains within Bins of Distance from Focal Tract due to Stimulus among Subpopulations Defined by Initial Earnings, Age, and Initial Industry: Average Across All Stimulus Specifications Featuring 250 New Jobs

Distance from	Employment Status/Earnings Quartile				
Focal Tract	Unemp.	1st Q.	2nd Q.	3rd Q.	4th Q.
Target Tract	0.014	0.013	0.009	0.008	0.006
1 Tct Away	0.038	0.031	0.025	0.021	0.016
2 Tcts Away	0.069	0.053	0.044	0.038	0.030
3+ Tcts w/in PUMA	0.148	0.114	0.099	0.089	0.072
1 PUMA Away	0.214	0.170	0.151	0.137	0.113
2 PUMAs Away	0.329	0.272	0.243	0.224	0.190
3+ PUMAs w/in State	0.778	0.679	0.607	0.597	0.613
1 State Away	0.869	0.787	0.713	0.697	0.700
2+ States Away	0.920	0.856	0.785	0.773	0.772
Out of Sample	1	1	1	1	1
	Age			Industry Status	
	Age $<$ 30	Age 31-50	Age $>$50	Diff Ind.	Same Ind.
Target Tract	0.010	0.009	0.009	0.008	0.020
1 Tct Away	0.028	0.023	0.024	0.023	0.038
2 Tcts Away	0.049	0.041	0.042	0.042	0.061
3+ Tcts w/in PUMA	0.110	0.094	0.094	0.096	0.123
1 PUMA Away	0.165	0.142	0.143	0.146	0.181
2 PUMAs Away	0.263	0.231	0.233	0.237	0.280
3+ PUMAs w/in State	0.676	0.626	0.622	0.638	0.665
1 State Away	0.774	0.723	0.722	0.737	0.757
2+ States Away	0.840	0.793	0.797	0.807	0.825
Out of Sample	1	1	1	1	1

Notes: See Table A3 for expanded definitions of the row labels. See Table 4 for expanded definitions of the column labels. Each cell contains the average job-related welfare gain (scaled to be equivalent to $ of 2012 annual earnings) generated by a 250 job stimulus for workers whose distance between their origin establishment and the census tract receiving the stimulus package is closer than or within the distance bin indicated in the row label and whose employment status or earnings in the origin year placed them in the earnings/employment category listed by the column label. Each cell averages results across 32 stimulus packages featuring new jobs with establishments with different combinations of industry supersector, firm size quartile, and firm average pay quartile. Results are further averaged across 300 simulations for each of the 32 stimulus package specifications featuring different target census tracts.

Table 6: Heterogeneity by Establishment Composition in Local Employment and Welfare Gains and Losses Among Focal Tract Workers in Various Subpopulations: Comparing Establishment Openings and Closings Featuring High-Paying Manufacturing Plants vs. Low-Paying Retail Stores

	Change in P(Employed)				Change in E[Welfare]			
	Manufacturing		Retail		Manufacturing		Retail	
Subpop.	Open	Close	Open	Close	Open	Close	Open	Close
All	0.0003	-0.0059	0.0002	-0.0058	150	-5624	64	-3134
Unemp	0.0013	-0.0008	0.0008	-0.0006	164	-92	99	-65
1st Earn Q.	0.0003	-0.0030	0.0002	-0.0110	83	-915	68	-3920
2nd Earn Q.	0.0001	-0.0060	0.0001	-0.0079	118	-3574	63	-4313
3rd Earn Q.	0.0001	-0.0087	0.0001	-0.0047	176	-8382	56	-3431
4th Earn Q.	0.0001	-0.0083	0.0001	-0.0031	209	-11962	49	-2663
Age $<=$ 30	0.0003	-0.0053	0.0003	-0.0076	126	-2852	77	-2773
Age 31-50	0.0003	-0.0063	0.0002	-0.0053	161	-6503	57	-3046
Age $>$ 50	0.0002	-0.0057	0.0001	-0.0044	162	-7515	59	-3710
Diff. Ind.	0.0003	-0.0002	0.0002	-0.0002	106	-94	54	-58
Same Ind.	0.0001	-0.0165	0.0002	-0.0180	298	-18699	90	-8430

Notes: The table displays the change in employment probability (columns 1-4) and expected welfare (columns 5-8, scaled in $ of 2023 annual earnings) generated by simulated manufacturing plant or retail store openings or closings for local workers (those employed (or unemployed) in the previous year in the focal tract) who initially belong to the subpopulation indicated by the row label. See Table 4 for expanded definitions of the subpopulations indicated by the row labels. The column subheadings "Manufacturing" and "Retail" indicate whether the results displayed in the chosen column reflect the creation or destruction of 250 positions at large, high paying manufacturing firms or large, low-paying retail firms, respectively. The column subheadings "Open" and "Close" indicate whether the results displayed in the chosen column reflect simulated plant openings featuring the creation of 250 jobs from the focal tract or plant closings featuring the removal of 250 jobs.

Table 7: Heterogeneity by Establishment Composition in Local and National Incidence Among Workers in Various Subpopulations: Comparing Plant Openings and Closings Featuring High-Paying Manufacturing Positions vs. Low-Paying Retail Positions

Panel A: Shares of Local Incidence among only Focal Tract Workers

		Employment				Welfare			
	Nat. Pop.	Manufacturing		Retail		Manufacturing		Retail	
Subpop.	Share	Open	Close	Open	Close	Open	Close	Open	Close
Unemp	0.121	0.543	0.016	0.558	0.012	0.132	0.002	0.188	0.003
1st Earn Q.	0.210	0.183	0.107	0.198	0.396	0.116	0.034	0.223	0.263
2nd Earn Q.	0.215	0.108	0.222	0.107	0.291	0.169	0.137	0.214	0.296
3rd Earn Q.	0.217	0.096	0.321	0.071	0.174	0.255	0.323	0.191	0.237
4th Earn Q.	0.237	0.070	0.334	0.066	0.127	0.329	0.504	0.184	0.201
Age <= 30	0.313	0.377	0.286	0.455	0.411	0.264	0.159	0.380	0.277
Age 31-50	0.425	0.425	0.458	0.361	0.389	0.454	0.491	0.378	0.413
Age > 50	0.262	0.198	0.256	0.184	0.200	0.282	0.350	0.242	0.310
Diff. Ind.	0.904	0.959	0.118	0.927	0.085	0.769	0.045	0.849	0.060
Same Ind.	0.096	0.041	0.882	0.073	0.915	0.231	0.955	0.151	0.940

Panel B: Shares of National Incidence among All Workers

		Employment				Welfare			
	Nat. Pop.	Manufacturing		Retail		Manufacturing		Retail	
Subpop.	Share	Open	Close	Open	Close	Open	Close	Open	Close
Unemp	0.120	0.363	0.391	0.050	0.086	0.404	0.444	0.082	0.134
1st Earn Q.	0.220	0.230	0.266	0.122	0.235	0.241	0.241	0.170	0.210
2nd Earn Q.	0.220	0.153	0.157	0.178	0.247	0.149	0.140	0.215	0.224
3rd Earn Q.	0.220	0.123	0.100	0.259	0.225	0.104	0.093	0.240	0.219
4th Earn Q.	0.220	0.131	0.085	0.391	0.206	0.102	0.082	0.294	0.212
Age <= 30	0.310	0.372	0.414	0.224	0.323	0.391	0.420	0.285	0.344
Age 31-50	0.426	0.412	0.388	0.464	0.412	0.405	0.384	0.448	0.407
Age > 50	0.265	0.215	0.198	0.312	0.264	0.204	0.196	0.266	0.249
Diff. Ind.	0.904	0.865	0.798	0.546	0.632	0.943	0.868	0.827	0.781
Same Ind.	0.096	0.135	0.202	0.454	0.368	0.057	0.132	0.173	0.219

Notes: Panel A displays the shares of all employment and welfare gains or losses (in columns labeled "Employment" and "Welfare", respectively) generated by the simulated plant openings or closings that accrue to all workers nationally who initially belong to the subpopulation indicated by the row label. Panel B displays the expected change in employment probability and job-related welfare (scaled in $ of 2012 annual earnings) from these openings and closings that accrue to local workers (those employed (or unemployed) in the previous year in the focal tract) who initially belong to the subpopulation indicated by the row label. See Table 4 for expanded definitions of the subpopulations indicated by the row labels. The column subheadings "Manufacturing" and "Retail" indicate whether the results displayed in the chosen column reflect the creation or destruction of 250 positions at large, high paying manufacturing firms or large, low-paying retail firms, respectively. The column subheadings "Open" and "Close" indicate whether the results displayed in the chosen column reflect simulated plant openings featuring the creation of 250 jobs from the focal tract or plant closings featuring the removal of 250 jobs.

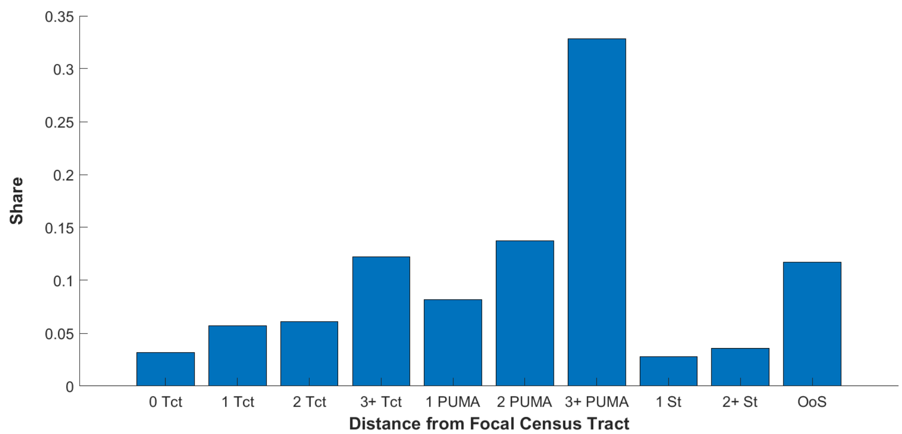
Table 8: Model Validation Results: Dissimilarity Index Values Comparing Forecasted and Actual Worker Reallocations Following Large Local Labor Demand Shocks Using Alternative Match Group Definitions and Methods for Generating Forecasts

	Alternative Models					Alternative Surplus Restrictions				
Level of Group Aggregation	Two-Sided Matching	Param. Logit	Raw CCP	Smoothed CCP	Choo-Siow	Loc. only (firm)	Loc. only (worker)	Loc. only (both)	No Same Ind.	No Same Firm
Full Group Space	0.351	0.458	0.353	0.356	0.351	0.389	0.344	0.447	0.353	0.847
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Dist. Bins	0.111	0.362	0.115	0.108	0.119	0.257	0.173	0.332	0.126	0.735
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Dist. Bins & No Firm Char.	0.023	0.266	0.038	0.037	0.037	0.130	0.086	0.192	0.023	0.024
	(3.8E-04)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(3.8E-04)	(0.001)
E-NE & NE-E Only & All Loc.	0.033	0.230	0.092	0.090	0.042	0.039	0.051	0.049	0.033	0.032
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
E-NE & NE-E Only & Dist. Bins	0.010	0.206	0.026	0.026	0.015	0.022	0.039	0.037	0.010	0.010
	(2.0E-04)	(0.001)	(0.001)	(0.001)	(3.5E-04)	(3.1E-04)	(3.6E-04)	(2.5E-04)	(2.0E-04)	(2.3E-04)

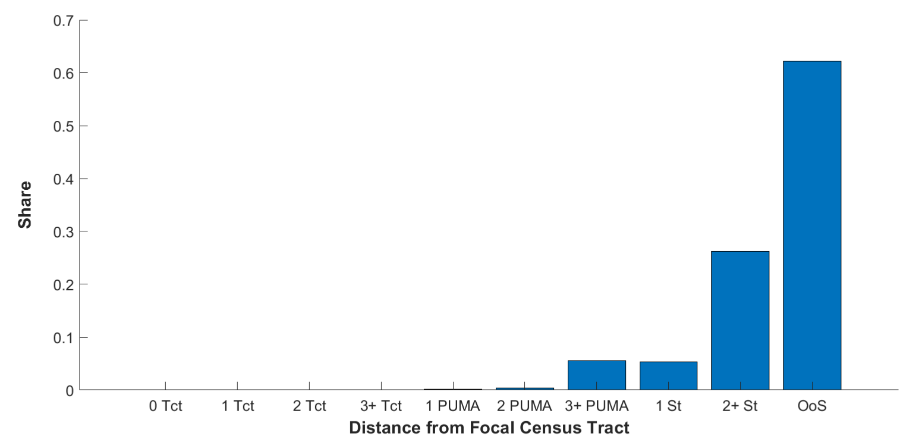
Notes: This table examines the fit of model-based predicted worker reallocations to the actual reallocations that occurred following a set of large establishment openings and closings in particular census tracts in particular years spanning 2003-2012. See Section A8 for a detailed description of the model validation exercise. Each row of the table considers a different metric for measuring model fit, while each column considers a different matching model. Columns 1-5 examine alternative matching models, while columns 6-10 consider aggregated versions of the baseline model from column 1. Each entry averages the fit metric across all 421 local shocks identified. For each shock, predictions are based on parameters estimated using local data from the year before the shock occurred. "Two-sided Matching" refers to the preferred two-sided matching model presented in this paper. "Param. Logit" refers to a one-sided parametric conditional logit model (See A8 for a list of the predictor variables). "Raw CCP" refers to a prediction that holds the previous year's conditional choice probability (CCP) distribution constant for each position type, but updates the position type marginal distribution to reflect the shock, while "Smoothed CCP" does the same but smooths the CCPs across similar position types before constructing the predicted reallocation. None of those three alternative models impose market clearing. "Choo-Siow" uses Choo and Siow (2006)'s version of the assignment model to generate predicted allocations. This model replaces the idiosyncratic surplus component ϵ_{ik} with the sum of two components $\epsilon^1_{lk} + \epsilon^2_{if}$. "Loc. only (firm/worker/both)" consider specifications that remove surplus heterogeneity among non-location firm characteristics, worker characteristics, or both, respectively. "No Same Firm" and "No Same Ind." remove surplus heterogeneity among match groups based on whether a worker is staying in the same firm and whether a moving worker is staying within the same industry, respectively. "Full Group Space" evaluates model fit using the index of dissimilarity between the actual and predicted distribution across match groups associated with workers from the PUMA targeted by the shock. "Dist. Bins", and "Dist. Bins & No Firm Char" evaluate the index of dissimilarity on aggregated group spaces in which origin and destination locations are each aggregated to 14 distance bins relative to the focal tract, and, in the latter case, position types featuring the same distance bin but different non-location characteristics are combined. "E-to-UE and UE-to-E Only (All Loc.)" calculates the index of dissimilarity only among match groups featuring employment-to-unemployment and unemployment-to-employment transitions, while "E-to-UE and UE-to-E Only (Dist. Bins)" does the same but aggregates locations to large distance bins relative to the focal census tract.

Figure 1: The Distance Distributions of Job-to-Job Transitions and of Workers' Distance from the Target Tracts of Simulated Labor Demand Shocks

(a) Empirical Distribution of 2012-2013 Job Transitions



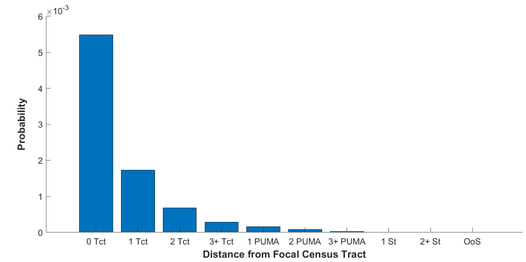
(b) Distribution of the Distance Between Workers' Initial Position and the Census Tract Targeted by the Simulated Stimulus Package: Average across All Simulated Stimuli



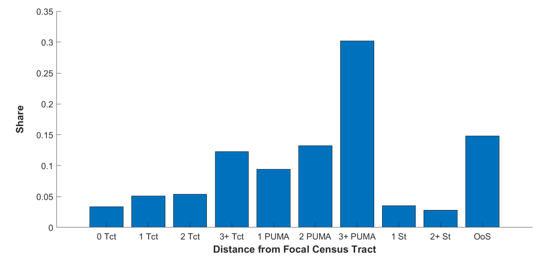
Notes: The bar heights in Figure 1a capture the shares of all worker transitions between dominant positions in 2012 and 2013 in which the geographic distance between these positions' establishments fell into the distance bins indicated by the bar labels. The bar heights in Figure 1b capture the shares of all workers for whom the geographic distance between their initial establishments and the census tract receiving the simulated stimulus package fell into the labeled distance bins (computed separately for each target tract, then averaged across all 300 target tracts). "0/1/2/3+ Tct" indicates that the two establishments (or, for Figure 1b, the establishment and the targeted tract) were in the same tract or one, two, or 3+ tracts away (by tract pathlength) within the same PUMA. "1/2/3+ PUMA" and "1/2+ State" indicate the PUMA pathlength (if within the same state) and state pathlength, respectively. "OoS" indicates that the worker's position was in an out-of-sample state.

Figure 2: Comparing the Spatial Distributions of P(Stimulus Job), Change in P(Employed), and Change in Average Welfare, along with Shares of Stimulus Jobs, Additional Employment and Additional Welfare: Average across All Simulated Stimuli

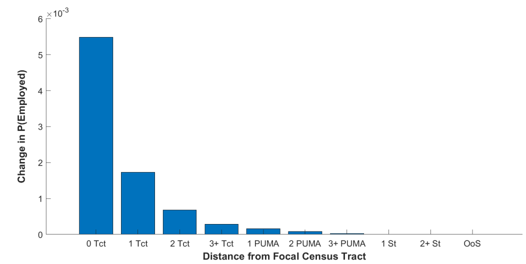
(a) Probability of Obtaining a Stimulus Job



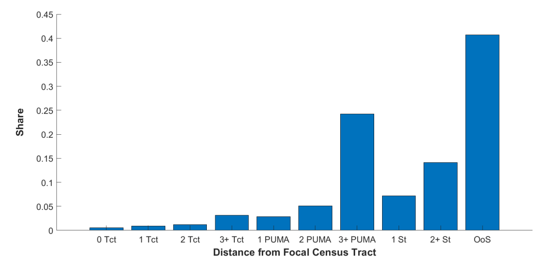
(b) Share of All Stimulus Jobs



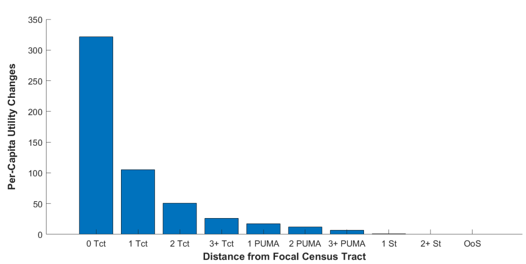
(c) Change in P(Employed)



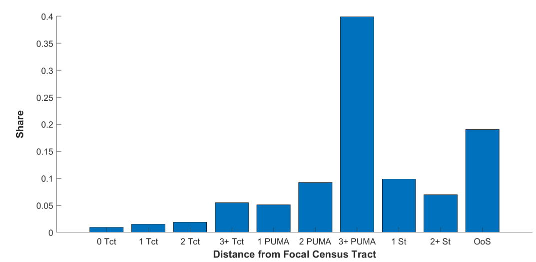
(d) Share of Employment Gains



(e) Avg. Welfare Gain (in 2023 $)



(f) Share of Welfare Gains

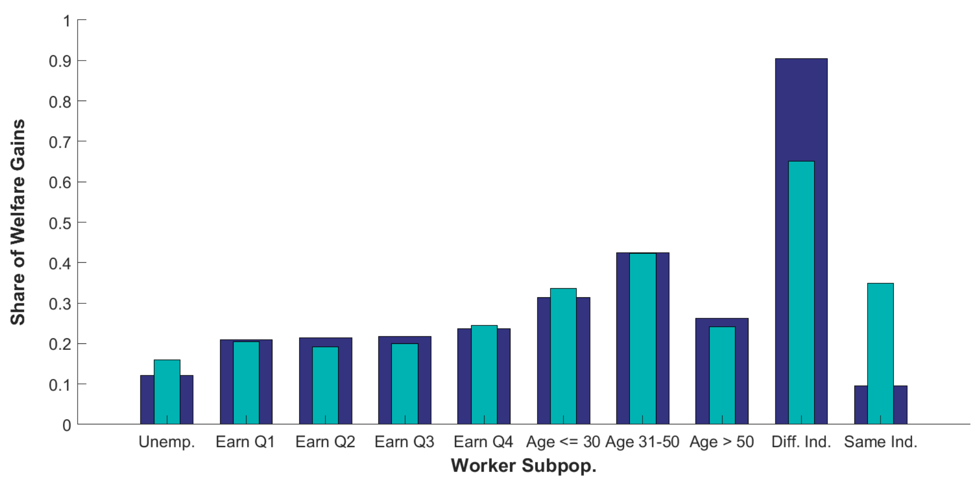


Notes: The bar heights in Figure 2a capture the average probability of obtaining a stimulus job among workers whose geographic distance between their initial establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. These probabilities average across different demographic categories and across stimulus packages featuring different firm compositions. Figure 2b displays the share of all stimulus jobs generated by the stimulus that redounds to workers in the chosen distance bin. Figures 2c and 2d display the corresponding gains in employment probability and shares of national employment gains accruing to workers in each distance bin, while Figures 2e and 2f display the corresponding expected welfare gains and shares of national welfare gains accruing to workers in each distance bin. Each bar represents an average over 300 simulations featuring different target census tracts as well as over 32 packages for each these 300 simulations featuring different firm composition (combinations of industry supersector and firm size and average pay categories). “0/1/2/3+ Tct” indicates that the origin establishment was in the target tract or was one, two, or 3 or more tracts away (by tract pathlength) within the same PUMA. “1/2/3+ PUMA” and “1/2+ State” indicate the PUMA pathlength (if within the same state) and state pathlength (if in different states), respectively. “OoS” indicates that the worker’s position was in an out-of-sample state.

Figure 3: Comparing Shares of Focal Tract Employment and Utility Gains with Initial Focal Tract Workforce Shares Among Workers from Different Subpopulations:
Average across All Simulated Stimuli

(a) Share of Focal Tract Net Employment Gains

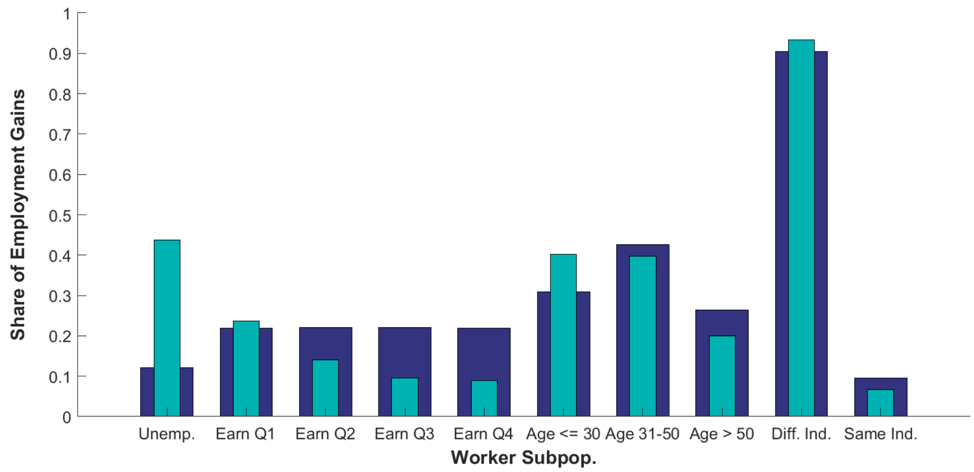
(b) Share of Focal Tract Utility Gains



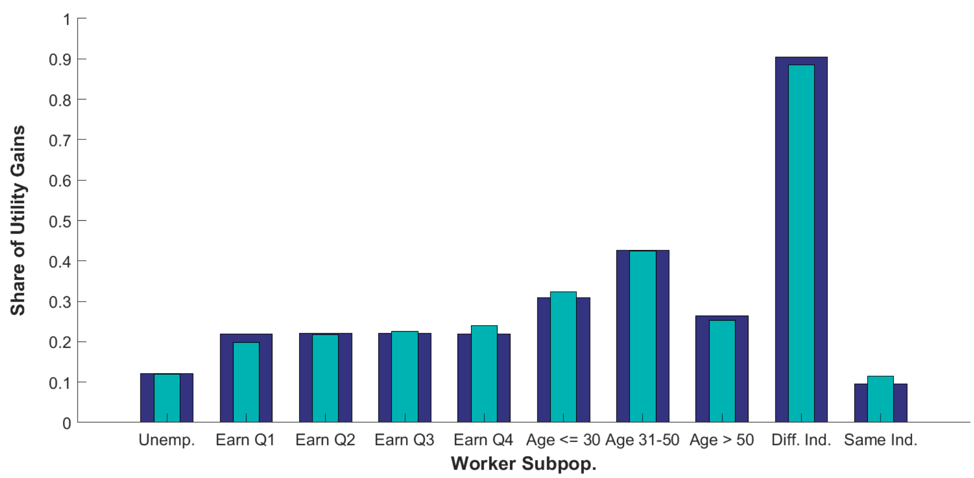
Notes: The heights of the wider bars within a particular group in Figures 3a and 3b capture the initial share of the focal tract workforce associated with the labeled worker subpopulation, while the heights of the narrower bars capture the subpopulation's share of the employment and job-related utility gains accruing to workers in the tract receiving the newly created jobs. Averages are taken across stimulus packages featuring different firm supersector/size/avg. pay compositions, as well as across 300 simulations featuring different targeted census tracts for each firm composition. "Unemp": Workers who were not initially employed. "Earn Q1"-"Earn Q4": Workers whose pay at their dominant job in the initial year placed them in the 1st/2nd/3rd/4th quartile of the national earnings distribution. "Same (Diff.) Ind". Workers whose position in the baseline year was in the same (different) industry as the jobs being created by the stimulus package.

Figure 4: Comparing Shares of National Employment and Utility Gains with Initial National Workforce Shares Among Workers from Different Subpopulations:
Average across All Simulated Stimuli

(a) Share of Additional Employment



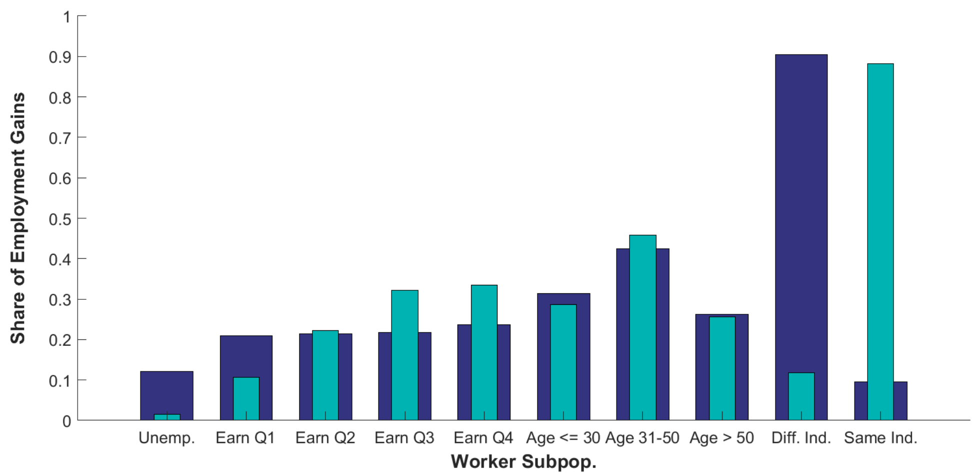
(b) Share of Total Utility Gains



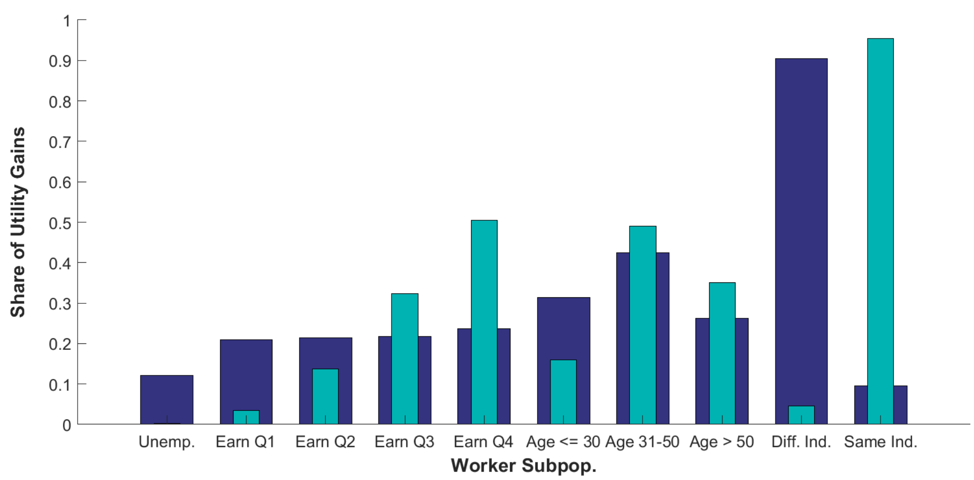
Notes: The heights of the wider bars within a particular group in Figures 4a and 4b capture the initial share of the national workforce associated with the labeled worker subpopulation, while the heights of the narrower bars capture the subpopulation's share of national employment and job-related utility gains created by the local job creation package. Averages are taken across job creation packages featuring 250 positions from different firm supersector/size/avg. pay compositions, as well as across 300 simulations featuring different targeted census tracts for each firm composition. "Unemp": Workers who were not initially employed. "Earn Q1"-"Earn Q4": Workers whose pay at their dominant job in the initial year placed them in the 1st/2nd/3rd/4th quartile of the national age-adjusted annualized earnings distribution. "Same (Diff.) Ind". Workers whose position in the baseline year was in the same (different) industry as the jobs being created by the stimulus package.

Figure 5: Comparing Shares of Focal Tract Employment and Utility Losses Produced by the Removal of 250 Positions at Large, High Paying Manufacturing Firms with Initial Focal Tract Workforce Shares Among Workers from Different Subpopulations

(a) Share of Focal Tract Net Employment Losses



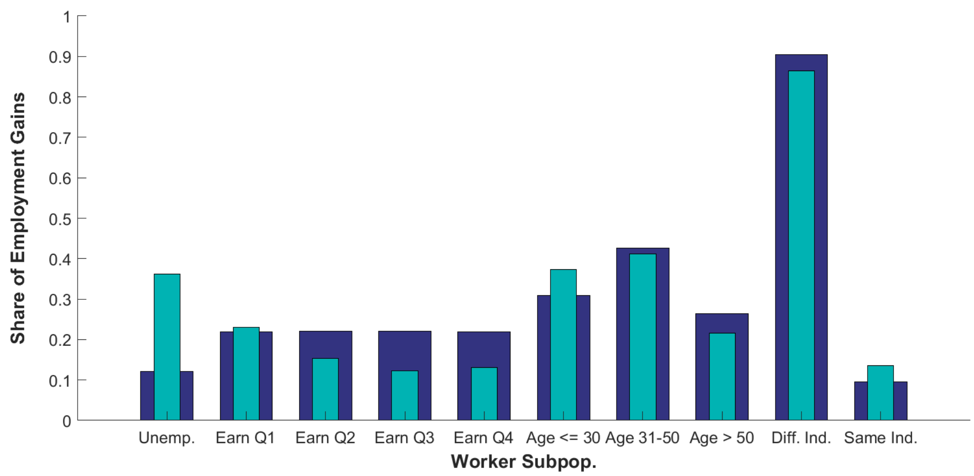
(b) Share of Focal Tract Utility Losses



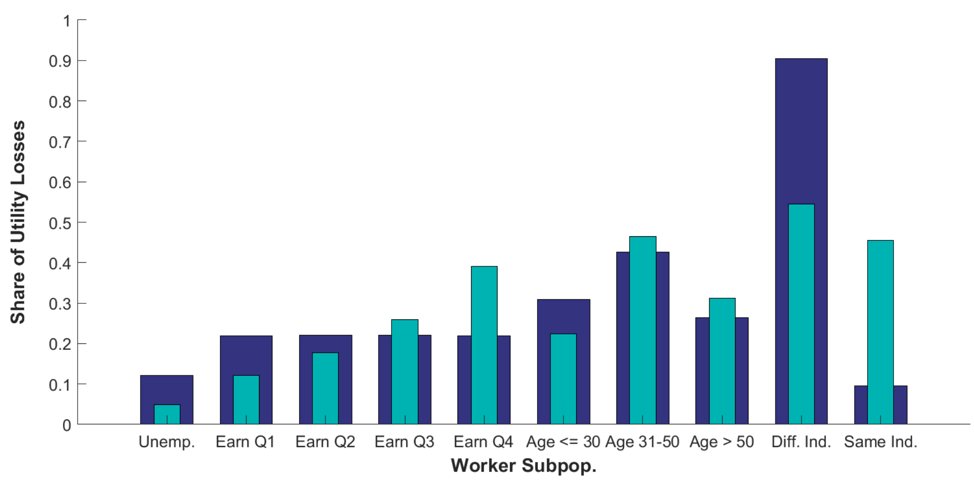
Notes: The heights of the wider bars within a particular group in Figures 5a and 5b capture the initial share of the focal tract workforce associated with the labeled worker subpopulation, while the heights of the narrower bars capture the subpopulation's share of the local employment and job-related utility losses accruing to workers in the tract experiencing the removal of 250 positions at large, high-paying manufacturing firms. Averages are taken across 200 simulations featuring different targeted census tracts for each firm composition. "Unemp": Workers who were not initially employed. "Earn Q1"-"Earn Q4": Workers whose pay at their dominant job in the initial year placed them in the 1st/2nd/3rd/4th quartile of the national age-adjusted annualized earnings distribution. "Same (Diff.) Ind". Workers whose position in the baseline year was in the same (different) industry as the jobs being created by the stimulus package.

Figure 6: Comparing Shares of National Employment and Utility Losses Produced by the Removal of 250 Positions at Large, High Paying Manufacturing Firms with Initial National Workforce Shares Among Workers from Different Subpopulations

(a) Share of Focal Tract Net Employment Losses



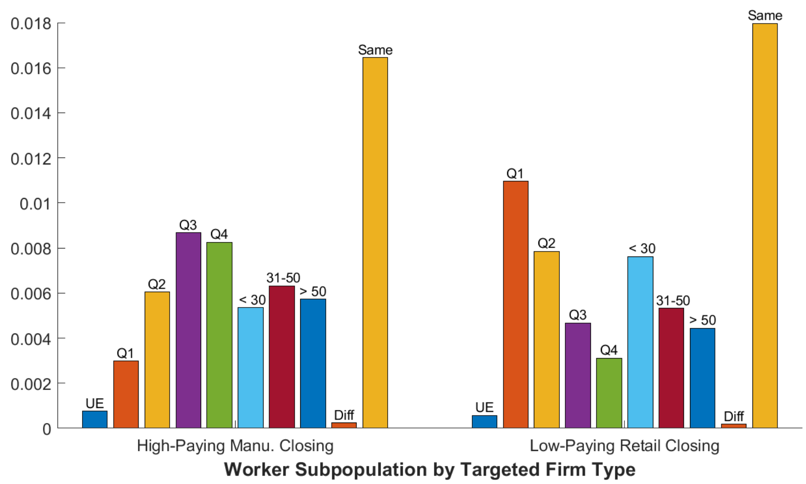
(b) Share of Focal Tract Utility Losses



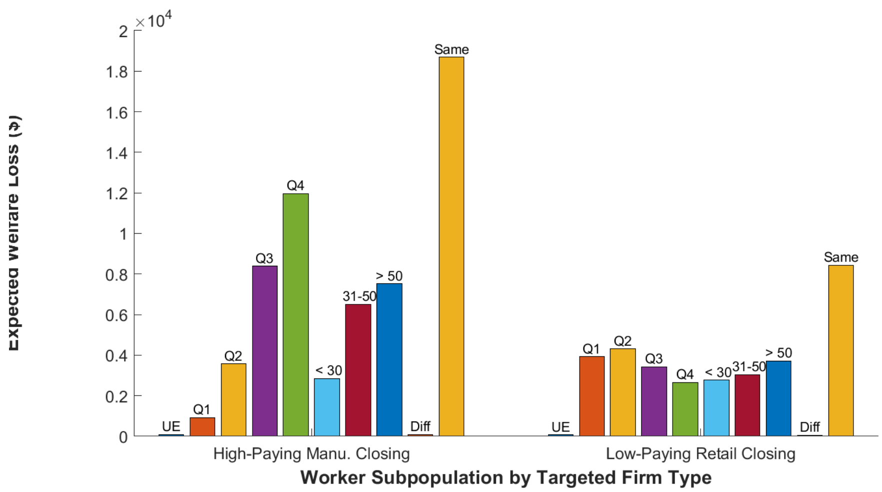
Notes: The heights of the wider bars within a particular group in Figures 6a and 6b capture the initial share of the national workforce associated with the labeled worker subpopulation, while the heights of the narrower bars capture the subpopulation's share of the national employment and job-related utility losses forom the removal of 250 positions at large, high-paying manufacturing firms in a single tract. Averages are taken across 200 simulations featuring different targeted census tracts for each firm composition. "Unemp": Workers who were not initially employed. "Earn Q1"-"Earn Q4": Workers whose pay at their dominant job in the initial year placed them in the 1st/2nd/3rd/4th quartile of the national age-adjusted annualized earnings distribution. "Same (Diff.) Ind". Workers whose position in the baseline year was in the same (different) industry as the jobs being created by the stimulus package.

Figure 7: Changes in Unemployment Rates and Expected Job-Related Utility for Workers from the Focal Tract Produced by Plant/Store Closings Featuring either High-Paying Manufacturing Establishments or Low-Paying Retail Establishments by Worker Subpopulation

(a) Change in P(Unemployed)



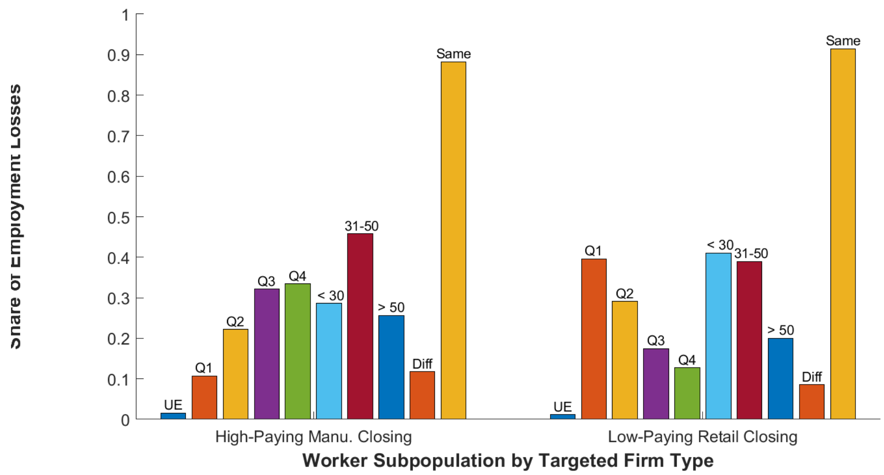
(b) Expected Utility Loss (in 2023 $)



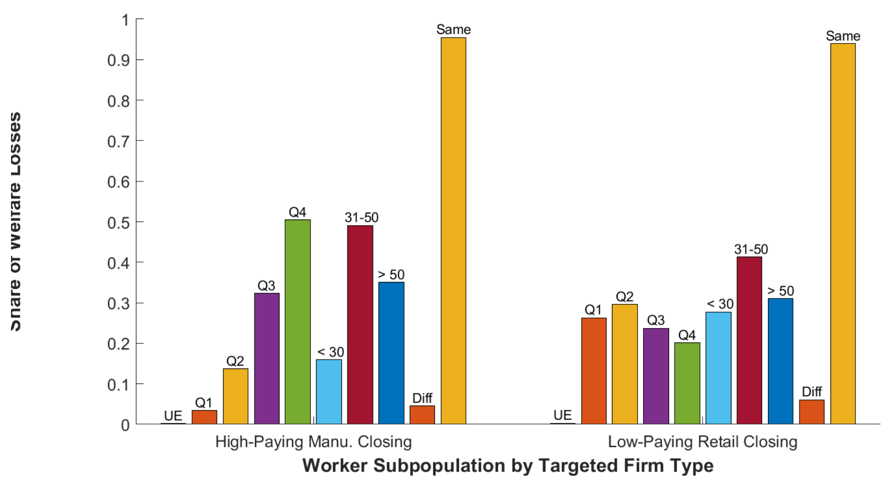
Notes: The bar heights within a particular group in Figures 7a and 7b capture the change in unemployment rate and expected utility, respectively, from two sets of simulated plant/store closings among workers who were employed (or unemployed) in the focal tract in the previous year and who belong to the subpopulation labeled atop the bar. See Figure 6 for more detailed descriptions of the labeled subpopulations. For each outcome, the left group of bars depicts the incidence of the removal of 250 positions at large, high paying manufacturing firms, while the right group depicts the corresponding incidence of the removal of 250 positions at large, low-paying retail firms. For each plant or store closing, averages are taken across 200 simulations featuring different targeted census tracts.

Figure 8: Shares of Focal Tract Employment and Welfare Losses Produced by Plant/Store Closings Featuring either High-Paying Manufacturing Establishments or Low-Paying Retail Establishments among Different Worker Subpopulations

(a) Share of Focal Tract Employment Losses



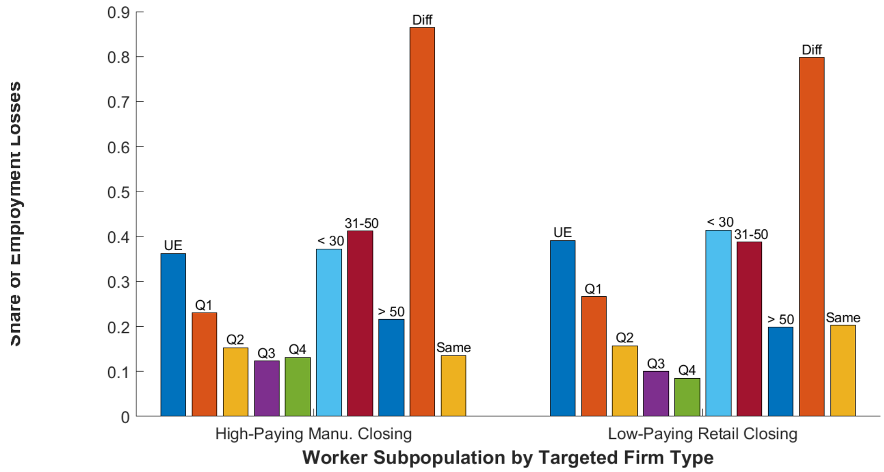
(b) Share of Focal Tract Welfare Losses



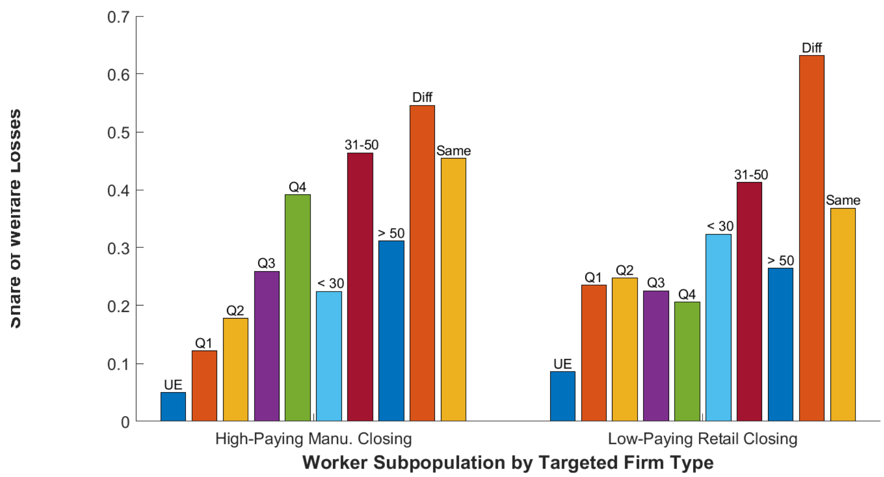
Notes: The bar heights within a particular group in Figures 8a and 8b capture the average share of local employment and welfare losses accruing to the worker subpopulations labeled over the bar among workers from the focal tract from two sets of simulated plant/store closings. See Figure 3 for expanded subpopulation definitions. For each outcome, the left group of bars depicts the incidence of the removal of 250 positions at large, high paying manufacturing firms, while the right group depicts the incidence of removing 250 positions at large, low-paying retail firms. For each plant or store closing, averages are taken across 200 simulations featuring different targeted census tracts.

Figure 9: Shares of National Employment and Welfare Losses Produced by Plant/Store Closings Featuring either High-Paying Manufacturing Establishments or Low-Paying Retail Establishments among Different Worker Subpopulations

(a) Share of National Employment Losses



(b) Share of National Welfare Losses



Notes: The bar heights within a particular group in Figures 9a and 9b capture the share of national employment and welfare losses, respectively, from a set of simulated plant/store closings accruing to the worker subpopulations labeled over the bar. See Figure 3 for expanded subpopulation definitions. For each outcome, the left group of bars depicts the incidence of the removal of 250 positions at large, high paying manufacturing firms, while the right group depicts the incidence of removing 250 positions at large, low-paying retail firms. For each plant or store closing, averages are taken across 200 simulations featuring different targeted census tracts.