Contrasting the Local and National Demographic Incidence of Local Labor Demand Shocks

Richard K. Mansfield*
University of Colorado-Boulder

March 16, 2025

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.

*This paper has benefitted greatly from considerable early assistance from Evan Buntrock, Henry Hyatt, John Abowd, and Erika McEntarfer and valuable comments from Jeronimo Carballo, Terra McKinnish, Brian Cadena, Francisca Antman, Miles Kimball, Martin Boileau, Alessandro Peri, Murat Iyigun, Stephan Heblich, Yanos Zylberberg, and participants at the NBER Trade and Geography Conference, the SOLE/EALE Joint Conference, the IZA World of Labor Conference, the American and European Urban Economics Association Conferences, and seminars at the University of Illinois, Princeton University, Colorado State University, the University of Bristol, and the University of Colorado-Boulder. This research uses data from the US Census Bureau's Longitudinal Employer-Household Dynamics Program, which was partially supported by NSF grants SES-9978093, SES-0339191, and ITR-0427889; National Institute on Aging grant AG018854; and grants from the Alfred P. Sloan Foundation. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1802 (CBDRB-FY24-P1802-R11331). Any views expressed are those of the authors and not those of the U.S. Census Bureau.

1 Introduction

Billions of dollars in local aid are spent each year by state, federal, and local agencies to support city- or county-level economic development initiatives that seek to enhance labor market opportunities for local workers (Bartik (2004)). These often consist of local infrastructure spending, tax breaks to lure firms to relocate, or discounted loans or subsidies aimed at startup companies. To decide which types of initiatives to support, federal, state, and local policymakers must predict 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 an initiative at the city, county, or state level depends critically on the shares of its employment and welfare incidence expected to redound to workers within city, county, and state borders.

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 few location(s), these studies generally feature samples that are too small or too geographically focused to allow comparison of shocks with different labor demand compositions on locations with different local labor supply compositions, or to examine differential demographic incidence among local and less local areas.

Motivated by this challenge, we adapt the assignment model of Choo and Siow (2006) (hereafter CS) 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 demographic and spatial incidence.

Two key features of CS' assignment game make it particularly suitable for this analysis. First, it accommodates multidimensional heterogeneity based on unordered categorical characteristics for agents on both sides of the matching market. This permits the use of arbitrary spatial links between workers and establishments in different geographic units (both large and small) that vary flexibly across 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 worker-position pairs 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, suffice to compute changes in match outcomes and expected welfare for both worker and firm types for any counterfactual change in labor supply and/or demand composition. This sufficient statistics

approach does not require specifying 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 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 feature three key properties that make it suitable for a rich assignment model: 1) they capture the (near) universe of job matches from the available states, mitigating selection problems; 2) their huge scale allows 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 provides the model with the necessary inputs to compute shares of employment and welfare gains or losses from alternative local labor demand shocks that accrue to particular demographic groups in locations both near and far from the shock.

Our simulations evaluate counterfactual establishment openings and closings that create or destroy 250 positions of a given establishment size, average pay, and supersector composition in particular U.S. census tracts. While we model how labor market competition diffuses labor demand shocks 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 2023 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.

Second, despite very disproportionate per-capita gains for the most local workers, the cumulative shares of welfare and particularly employment gains accruing to non-local workers can be quite large, since the local workforce makes up a very small share of the national labor market. We find that only 9.9% of job-related utility gains and 5.8% of net employment gains from such local stimuli

[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 containing at least 100,000 residents. We use PUMAs instead of Commuting Zones to better match the level at which local economic initiatives are formed, since PUMAs are smaller and more consistent in population (there are 2,378 PUMAs vs. 741 CZs), do not cross state lines, and often distinguish large suburbs from city centers.

accrue to workers initially or most recently working in the surrounding PUMA, while 35.9% and 62.0% of such 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 locally. We also find that the within-PUMA share of employment gains is over twice as high for shocks to rural rather than urban areas (8.7% vs. 3.7%).

Third, the primary beneficiaries of local shocks vary widely with the establishment composition of the newly created jobs, suggesting opportunities for local officials to craft initiatives that target specific local subpopulations. 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 low-paid workers benefit most ($573) from positions at large, low-paying education & health firms and least ($167) from positions at large, high paying information firms, and the highest-paid workers benefit most ($641) from positions at large, high-paying education & health firms and least ($154) from positions at large, low paying information firms. Thus, whether shocks reduce local income inequality depends critically on which types of jobs are brought to town. Incorporating existing job multiplier estimates only slightly alters these findings.

Fourth, in contrast to the sensitivity of local incidence to shock composition, 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 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 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 that uses the estimated model to forecast the real-

ized 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 one-sided parametric models that fit firm or worker conditional choice probabilities with over 100 parameters. Thus, the model's 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, 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).[3]

Two recent papers in this vein are notable for incorporating spillovers to non-targeted locations driven by worker mobility. Like this paper, Sprung-Keyser et al. (2022) finds that heterogeneity in geographic mobility by demographic group substantially affects 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. Hornbeck and Moretti (2024) show as we do 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 depend on 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.

Our approach also complements a sub-literature on local job multipliers from an initial job stimulus due to increased product demand and agglomeration/congestion externalities (e.g. Moretti (2010) or Bartik and Sotherland (2019)). Such papers do not assess which types of workers from which initial locations benefit from the net change in local job opportunities, while our model takes the new spatial distribution of positions (possibly reflecting multipliers) as an input and evaluates the resulting skill and spatial incidence. We demonstrate this complementarity by evaluating a shock combining 250 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 gain insight by imposing additional structure on the sources of match surpluses (e.g. Schmutz and Sidibé (2019)),

[3]A 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.

incorporating housing and/or product markets (e.g. Monte et al. (2018)), or introducing dynamics (e.g. Caliendo et al. (2019)), while our model offers a richer and more flexible labor market.[4] 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.

These papers each aggregate locations to at least the county level. Manning and Petrongolo (2017), by contrast, use a search model to assess the geographic employment incidence of exogenous vacancy creation within a single British census (tract-sized) ward. Like Marinescu and Rathelot (2018), they find that labor markets are quite local, in that moderate distance to vacancies markedly decreases the probability of applying. But like us, they find that ripple effects from overlapping markets cause little of the employment gain to accrue to the targeted ward.

These papers feature no worker heterogeneity beyond initial location, and only Caliendo et al. (2019) has observable firm heterogeneity (industries) beyond location. Several spatial labor market models (e.g. Diamond (2016), Piyapromdee (2021)) feature imperfect substitutability among observable worker types but only index firms by location. Lindenlaub (2017), Bonhomme et al. (2019), and Chan et al. (2024) estimate multidimensional two-sided labor market models, but without any geographic dimension. Thus, only our model can evaluate differential incidence across both space and skill/demographic groups from local labor demand shocks of varying firm composition.

Indeed, Nimczik (2018) 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 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 to define labor markets 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.[5]

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), Menzel (2015), Chiappori and Salanié (2016), Mourifié and Siow (2021), and Galichon and Salanié (2022)).

We make three contributions to this literature. First, 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

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

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

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 unobserved on one or both sides of the market.[6] 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 and establishes identification of the joint surplus parameters. Section 3 describes the LEHD database and presents summary statistics that motivate the subsequent analysis. Section 4 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 incidence. Section 5 presents the main findings and the model validation results. Section 6 concludes.

2 The Two-Sided Assignment Model

We model the labor market as a static assignment game played by workers and establishments. Our adaptation of CS' model to a labor market setting shares many features with Chan et al. (2024): 1) workers and positions assigned to many worker and position types defined by combinations of observable characteristics; 2) unrestricted type-level horizontal and vertical differentiation on both sides of the market, so that worker types may differ in their rankings of firm types' non-pecuniary amenities and position types may differ in their rankings of worker types' productivity; and 3) unobserved idiosyncratic match quality among individual worker-position pairs. These features allow the model to flexibly accommodate a variety of frictions from switching locations, industries, and even jobs that differ across demographic groups. We show that such frictions are reflected in observed mobility patterns and shape the worker incidence of local labor demand shocks. We follow the recent spatial literature by minimally restricting links between different locations (e.g. Caliendo et al. (2019), building on Dekle et al. (2007)).[7]. Like Lamadon et al. (2022), firms are non-strategic, so that they ignore the impact of their job offers on equilibrium compensation. Unlike both papers, firms can wage discriminate based on idiosyncratic worker tastes or productivity, and thus offer worker-specific compensation rather than posting a common wage within a worker type.

2.1 Defining the Assignment Game

The exposition of the model follows Galichon and Salanié (2022) (hereafter GS), which generalizes CS. Suppose that in a given year there are I workers comprising the set $\mathcal{I}$ who participate in the labor market. Each worker i enters the market in an existing job match with a position $j(i)$ at establishment $m(j(i))$ taken from the set of possible positions $\mathcal{J}$, with $m(j) = 0$ representing a "match" with unemployment. Each worker i belongs to an observed worker type $l(i) \in \mathcal{L}$. In the

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

[7] We abstract from endogenous firm location choices (Bilal, 2023) and human capital dynamics (Martellini, 2022)

empirical work, worker types will be 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 initial establishment.[8]

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.[9] 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.[10]

We assign each potential job match (i, k) to a group g among a set $\mathcal{G}$ of G mutually exclusive groups. Let $g(i,k) \equiv g(l(i), f(k), z(i,k)) \equiv g(l, f, z)$ denote match (i,k)'s group assignment, with $z(i,k)$ capturing match characteristics defining the group beyond the worker and position's types. Below, the only z characteristic will be a trichotomous indicator that equals one for continued employment at the same establishment, two for employer changes within the same supersector, and zero otherwise.[11] 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 $U(i,k)$. We adopt a money-metric form for $U(i,k)$, with w_{ik} denoting the worker's potential earnings at k:[12]

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

$\pi^i_{ik} \equiv \theta^l(g(i,k)) + \epsilon^i_{ik}$ 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. That said, such components might include worker i's valuation of various non-pecuniary amenities offered by position k (including its location's appeal). In addition, though assignment games traditionally have been characterized and parameterized as "frictionless", π^i_{ik} could in principle also capture any deterministic or even stochastic search, moving, or training costs paid by worker i to find, move to, or settle into position k from initial position j.[13] 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.[14]

[8] A lack of residential location data requires initial (i.e. past) establishment locations to be used as worker locations.

[9] This might occur if the cost of coordinating multiple hires outweighs the gains from better exploiting production complementarities. Roth and Sotomayor (1992) discuss difficulties arising from preferences over collections of workers.

[10] There is no inconsistency in using prior earnings to proxy for both a worker's skill and an employer's skill requirements, since earnings have been widely shown to contain persistent worker and firm components.

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

[12] We have data on earnings but not wages or hours, so we assume that jobs are full-time and salaried.

[13] Menzel (2015) makes a deterministic assignment model stochastic by adding a probability that i and k meet that is independent of other payoff determinants and assigning their joint surplus to $-\infty$ if the pair does not meet.

[14] 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

$\theta^l(g)$ captures the part of position k's value to worker i that is common to any type $l(i)$ worker accepting a type $f(k)$ position 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 keeping their current position to avoid search and training costs. ϵ^i_{ik} captures the part of k's value to i that is specific to (i,k) within (l,f,z). For example, a worker may prefer to return to a familiar location.

Let $V(i,k)$ denote the value to establishment $m(k)$ of hiring (or retaining) worker i in position k. $V(i,k)$ is assumed to have an analogous form to $U(i,k)$:

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

Akin to π^i_{ik}, π^k_{ik} 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 the next year with i already in position k. As with π^i_{ik}, π^k_{ik} contains a common group-level component $\theta^f(g)$ and an idiosyncratic component $\epsilon^k_{i,k}$. $\theta^f(g)$ might capture smaller per-position costs for larger firms of recruiting distant workers, or greater revenue generation by highly skilled workers at high-paying firms. ϵ^k_{ik} might capture skills required by position k that worker i uniquely possesses.

We define the joint surplus from match (i,k) as the combined worker and position valuations:

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

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 assignment 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. Shapley and Shubik (1972) show that transferable utility guarantees a unique competitive equilibrium assignment (or, equivalently, stable matching) of workers to positions as long as preferences are strict on both sides of the market.

Two key features of the equilibrium assignment should be noted. First, it is fully determined by the joint surplus values $\{\pi_{ik}\}$ (See Appendix A1); no separate information on the worker and firm components π^i_{ik} and π^k_{ik} 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 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 equilibrium assignment μ is generally unique, the equilibrium payoffs and transfers are not: all r_i

are likely to generate similar incidence predictions, particularly for short-run impacts from small shocks.

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. 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 increases in utility bids r_i following demand shocks can always be scaled in terms of earnings gain equivalents (even when they involve taking an earnings cut to get a position offering higher non-pecuniary values).

Since we wish to examine shock incidence at the worker type level rather than predict exact worker-position matches, we follow CS in analyzing group-level equilibria that are consistent with the underlying disaggregated 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 $\theta_g \equiv \theta^l(g) + \theta^f(g)$ and $\epsilon_{ik} \equiv \frac{\epsilon^i_{ik}+\epsilon^k_{ik}}{\sigma}$. σ is a scaling parameter that captures the relative importance of idiosyncratic vs. group-level surplus components in producing the variation in match surpluses across potential pairs (i, k). We show below that counterfactual assignments do not depend on σ, but σ governs the scale of changes in utility bids r_i needed to re-equilibrate the market.

The goal is to use the observed matching μ to recover the set of group mean surplus values $\{\theta_g\}$ that govern the frequencies of different kinds of job matches. We secure identification of $\{\theta_g\}$ by assuming that ϵ_{ik} draws are i.i.d across all potential (i, k) pairs and follow a Type 1 extreme value distribution.[15] Unlike in CS and GS, the idiosyncratic component in equation (4) is 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 pair-specific idiosyncratic components eliminates violations of the independence of irrelevant alternatives (IIA) axiom, so that subdividing worker or position types in arbitrary ways does not change estimated joint surplus values or match probabilities. Sections 5.6 and 5.7 compare simulated forecasts and out-of-sample prediction accuracy between our model and the CS model. As discussed in Section 2.3 and Appendix A5, allowing such heterogeneity prevents the use of observed transfers to recover the worker and position subcomponents $\theta^l(g)$ and $\theta^f(g)$. Fortunately, this decomposition is not needed to generate key measures of worker-level incidence.

[15] Menzel (2015) shows that imposing i.i.d draws is the key assumption rather than the Type 1 EV distribution. Sprung-Keyser et al. (2022) provide quasi-experimental support for a key property of i.i.d EV models: origins' changes in mobility rates to a location receiving a labor demand shock increase monotonically with their baseline mobility rates.

2.2 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 $\mu_{ik} = 1$ only if $i \in \arg\max_{i \in \mathcal{I}} \pi_{ik} - r_i$. 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 expression 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 type f position 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.[16] 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 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, because mean surplus values among (l, f) pairs are identified without Assumptions 1 and 2, these assumptions are only necessary to isolate the surplus from hiring a within-firm incumbent relative to a worker of the same type from another firm in the same census tract. Since we focus on characterizing expected values of incidence by worker type rather than full distributions, restricting within-type heterogeneity is of minimal consequence. Violations (discussed in Appendix A1) lead to slight over or understatements of deviations among (l, f) type combinations from the samplewide 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, $\hat{n}(l)$ denote the fraction of matches featuring type l workers, and $\hat{h}(f)$ denote the fraction featuring type f positions.[17] One

[16] 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.

[17] Because we do not observe unfilled vacancies, in the empirical work we augment $\mathcal{K}$ to include a sufficient number

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 $\{\hat{P}(g|f)\}$ will 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 (given full population data), to establish identification one can treat their terms as known and move them to the left hand side. These adjusted log odds still conflate the re-scaled mean surplus difference between g_1 and g_2, $(\frac{\theta_{g_1}-\theta_{g_2}}{\sigma})$, with the log ratio of mean exponentiated re-scaled utilities among 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 $l(g_4) = l(g_2)$. 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:

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

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 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 high-paying firms often hire other firms' high-paid workers and low-paying firms often hire other firms' low-paid workers but not vice versa, there must be greater combined surpluses from the first two kinds of matches than their alternatives. Whether due to complementarities in production, or in tastes for/provision of amenities, or reduced moving costs, we show that one need not identify the source of this comparative advantage to evaluate the incidence of counterfactual labor demand shocks as long as it is minimally affected by the shock.

of unemployment "positions" to ensure that each match will have both a worker and a "position".

2.3 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 possible supply shock). 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 how the opening will change job-related utilities and employment rates among existing workers/job seekers in the chosen and nearby neighborhoods.

We assume that the counterfactual assignment also satisfies the assumptions of Proposition A1 above, and 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.[18] 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 $\mathbf{C}^{CF} \equiv \{C_1^{CF} \dots C_L^{CF}\}$ are ex ante unknown equilibrium objects of interest affected by the counterfactual changes reflected in $(\theta_g^{CF}, n^{CF}(l), h^{CF}(f))$. Thus, each counterfactual CCP must be treated as a function of $\mathbf{C}^{CF}$.

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.[19] 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 the job assignment is unaffected. 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 they negligibly affect the division of surplus and can be safely ignored. A related concern is that firms might alter how many positions they choose to fill when wages change following labor demand shocks. However, for relatively small and localized shocks, firms' extensive margin response may be highly inelastic

[18] These shares can be directly computed when $n^{CF}(l)$ and $h^{CF}(f)$ are set equal to those from some observed year y.

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

if the costs of adjusting its number of positions (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 worker composition of a fixed set of positions.

While our baseline approach assumes a perfectly inelastic extensive margin, we explicitly incorporate endogenous extensive margin responses as a robustness check in Section 5.6. This requires using existing wage elasticity and multiplier estimates and iterating between assignment model equilibria and calibrated extensive margin responses to changes in a position's expected profitability until a fixed point is found. The final $h^{CF}(f)$ can be interpreted as a post-adjustment distribution.

Treating the set of filled positions 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 that identify extensive margin labor demand elasticities by type. Instead, surplus diff-in-diffs Θ^{D-in-D} (along with σ) only determine equilibrium elasticities of substitution for each position type among different worker types. These elasticities are fully determined by *relative* prices, so valid sources of relative price variation among workers from different initial locations include shifts in the spatial composition of workers seeking positions as well as shifts in the spatial composition of positions seeking workers. So there is no inconsistency in using the full set of year-to-year job flows that are 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 eases the computation of counterfactual equilibria. With unknown counts of unmatched partners, 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 $\{C_f^{CF}\}$.[20] 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 $n^{CF}(l)$, allowing worker-side clearing.[21] Because relative payoffs among worker types fully determine the equilibrium assignment and the worker type distribution $n^{CF}(*)$ must sum to one, one can normalize $C_1^{CF} = 0$, leaving $L - 1$ market clearing conditions for $L - 1$ remaining members of $\{C_l^{CF}\}$.

Given $\mathbf{C^{CF}}$, one can directly recover the counterfactual probability for any match group via $P^{CF}(g) = \sum_f h^{CF}(f) P^{CF}(g|f, \mathbf{C^{CF}})$. Since this solution also satisfies the stability and feasibility conditions, it must be the unique aggregate counterfactual stable assignment. Appendix A3 proves this result. Because only $\min\{L - 1, F\}$ equations must be solved, this approach provides considerable computational savings when $L \gg F$ or $F \gg L$. Below we present results that average over 300 counterfactual allocations featuring around 5,000 worker and 10,000 position types.

[20] 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.

[21] 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)).

2.4 Interpreting the Counterfactual Simulations

We generally use data from the 2012-2013 set of job matches 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 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 2.2 suggests that only the set of diff-in-diffs $\Theta^{D-in-D,2012}$ is identified. In Appendix A2, we prove Proposition A2, which 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 $\{\Delta_l\}$ 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. A2 imply that baseline utility differences 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.[22] 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 and size decisions. To highlight heterogeneity in incidence by firm size, average pay, and industry,

[22] 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 $\ln(C_l) \equiv \ln(\frac{1}{|l|}\sum_{i:l(i)=l} e^{\frac{r_i}{\sigma}}) \approx \ln(\frac{1}{|l|}\sum_{i:l(i)=l} e^{\frac{\overline{r}_l}{\sigma}}) = \frac{\overline{r}_l}{\sigma}$.

we 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)).[23] 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.[24] Nonetheless, to assess sensitivity to unmodeled changes in local continuation values, we consider in Section 5.6 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 the new positions to have the same expected duration as other positions of their type.[25] 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. To justify the focus on one-year transitions between static equilibria, in Appendix A8 we use data on each supersector's mean vacancy durations and shares of new hires from each other supersector and from unemployment to calibrate simulations suggesting that at least 98% of vacancy chains generated by a shock creating a small set of new positions are completed via a U-to-E hire within one year, regardless of the supersector responsible for the original job creation.

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 sug-

[23] 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.

[24] In Appendix A2, such surplus changes only affect Δ^2_f, 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.

[25] Analyzing shocks of varying duration requires an explicitly dynamic assignment model akin to Choo (2015).

gest that increases in rent and other local services' prices 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. 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. 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.[26] 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.

2.5 Identifying σ

The share of welfare gains or losses by worker type can be recovered without estimating σ. However, since utility is additive in earnings, knowledge of σ allows estimated utility gains $\frac{\overline{r}_l^{CF1} - \overline{r}_l^{CF2}}{\sigma}$ to be scaled in dollars, making it easy to gauge the economic importance of shock-induced welfare changes. Conditional on Θ, σ sets the elasticity of matching choices with respect to relative wages or required utility bids, which governs the scale of utility reallocation due to labor demand shifts.[27]

As Galichon et al. (2017) discuss, identifying 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 strong assumptions, estimates of σ are likely to be quite rough, though they are fairly consistent across years.[28] We use the mean of $\hat{\sigma}^y$

[26] 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}.

[27] Intuitively, when position type C disproportionately chooses workers of type A over type B compared to position type D, it could be because $\theta_{AC} - \theta_{AD} \gg \theta_{BC} - \theta_{BD}$ and σ is large, or because $\theta_{AC} - \theta_{AD}$ slightly exceeds $\theta_{BC} - \theta_{BD}$ but σ is tiny. In the first case, large changes in utility bids are needed to induce enough substitution across worker types to produce the required reallocation. In the second case, small utility changes suffice to re-equilibrate the market.

[28] 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 utilities $\ln(C_l^{CF,y}) - \ln(C_{l'}^{CF,y}) \approx (\overline{r}_l^{CF,y} - \overline{r}_{l'}^{CF,y})/\sigma^y$. These utility predictions stem from counterfactual simulations in which worker and position type distributions evolve as they actually did but surpluses are fixed at 2003-2004 values. The coefficient on $(\overline{r}_l^{CF,y} - \overline{r}_{l'}^{CF,y})/\sigma^y$ approximates σ^y as long as a) other determinants of actual utility changes by worker type, namely changes in relative joint surpluses Θ, are roughly orthogonal to the predicted utility changes based on changes in supply and demand composition, and b) mean utility gains for each worker type in year y generally consisted of increases in earnings rather than amenities or continuation values.

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.

3 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.[29] The worker IDs are then linked across states, and the data are augmented with establishment assignments, establishment characteristics (notably location and industry) , and worker demographics (including age, race and sex but not occupation nor education for most workers).[30]

3.1 Sample Selection

We take a 50% random subsample of all workers ever observed as employed between 2002 and 2013 within the 19 U.S. states that opted to provide data to our project.[31] 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 our 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. Preliminary work suggested that the shock incidence forecasts were quite insensitive to the years chosen.[32]

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.[33] Workers are considered nonemployed in a given year if they did not earn above $2,000 at any job in

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

[30] 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 due to 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.

[31] 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.

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

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

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 below rely on surplus values 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 some 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 the model's, 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 then use BLS national unemployment counts by age group to align the scale of the labor force with standard measures. Appendix A6 details these imputation procedures.

Rather than exclude workers from the remaining 31 states, which would cause us to overstate the geographic concentration of shock incidence, we aggregate all out-of-sample employment into a single out-of-sample "state". 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 sensitive to observing worker flows to and from states adjacent to the focal state, we sample target tracts only from 10 states in the west/southwest/great plains area where almost all adjacent states are observed.[34]

3.2 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 4.1), their age-adjusted earnings quartile based on their $y-1$ earnings at this establishment, their age category ($\leq$ 30, 31-50, or $>$ 50), and whether their $y-1$ supersector matches that of the simulated job creation or destruction.[35] For workers who were not employed in $y-1$, their most recent establishment's location is used (for new entrants, the location is imputed using ACS/LEHD data) and the earnings quartile is replaced

[34] ACS residential mobility data suggest that we observe about 47% of year-to-year worker inflows into these 10 states and 92% of total job-to-job changes (including within-state flows) ending in these 10 states.

[35] 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.

by a separate "unemployed" category. Workers' year y positions are assigned to position types based on the combination of their establishment's location, supersector, employment (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 observed and likely to be key determinants of productivity complementarities, recruiting, search and moving costs, and the other match surplus components. 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 ($z(i,k) = 1$), industry but not firm ($z(i,k) = 2$), or neither ($z(i,k) = 0$).

3.3 Summary Statistics

Figure 1a and column 1 of Table A4 present 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 our 24.2M sample observations from 2012-2013 involved job-to-job transitions, with 8.3% changing supersector. 69.5% of workers kept 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.

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 ≤ 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, age ≤ 30 workers 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 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 25% for finance to 42% for natural resources).

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.

4 Estimation

4.1 Collapsing the Type Space for Distant Geographic Areas

Since groups g are defined by several worker and position characteristics besides their respective locations, treating each census tract as a separate location would generate trillions of groups. Given elevated interest in the incidence of alternative shocks among nearby locations, we combine initial types (and thus groups) with the same worker and position characteristics that are near each other but far from the shock. Specifically, beyond a five tract radius around the targeted tract, a type's location is defined by its PUMA. Beyond the targeted state, a type's location is defined 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 worker and position tracts are in different states would otherwise be weakly identified. This approach still uses all observed job matches and all locations in the 19 state sample plus the out-of-sample "state", so each local labor market remains nested within a single national market.

Even after combining types, there are relatively few observed matches per group, 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 A2 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.

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.[36] However, because the type aggregation procedure implies that type and group spaces vary by target tract, we must redefine match groups post-simulation prior to averaging by replacing worker and position type locations with bins of distance to the targeted tract.[37] We then report incidence estimates for various distance rings around the shock.[38] 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.

4.2 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".[39] This represents about a 10% change in labor demand for an average tract with around 2,500 jobs. For each chosen 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 A3 details each shock's composition. We then consider packages that require the new positions to be filled only by workers from the surrounding PUMA so as to assess the value of analogous stipulations in some economic development contracts between cities and incoming firms.[40] 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.

[36] A tract is only eligible to be a target tract if it contains ≥ 250 jobs, so that surplus parameters for local matches are well-identified. We use the same 300 target tracts for each shock specification to ease comparison across specifications.

[37] 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.

[38] 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.

[39] 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.

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

4.3 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 5.7 to assess the combined contribution of sampling error and misspecification to out-of-sample forecast accuracy.[41]

5 Results

5.1 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 and utilities 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 A4, col. 3) displays the mean probability of taking one of the 250 stimulus jobs among workers 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: a target tract worker's probability of taking a stimulus job is three times higher (.0054) than one in an adjacent tract (.0017) and about 8 and 20 times those of workers 2 tracts away (.0007) or 3+ tracts away within the same PUMA (.0003). Additional 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+ PUMAs away within the state, and 4,279 and 26,566 times more likely than a random worker one state or 2+ 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 A4 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 A4, 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

[41] The first few results tables do provide standard errors reflecting only sampling error from averaging over a 300 target tract sample instead of all available tracts. These standard errors are tiny, suggesting little value to enlarging this sample.

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 A4, col. 5) reports the change in the probability of any employment, relative to a no-stimulus counterfactual, by distance from the target 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 A4, 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 A4, 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, which varies with shock composition but is generally young, unemployed workers in a distant state. 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 A4, 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% 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 A5) 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 A6) 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.

5.2 Heterogeneity in Local Incidence by Worker and Firm Characteristics

Figures 3a and 3b plot the shares of focal tract employment and welfare gains that accrue to local subpopulations defined by initial earnings, age, or same/different industry category against their respective baseline local employment shares, while Table 2 (col. 1) reports their per-worker gains.

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) allows 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.

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 actively searching for jobs. However, further disaggregation reveals that among the initially unemployed, younger workers 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 gains for younger employed workers than older ones 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.

Table 2 shows that the pattern of local employment and welfare gains by subpopulation varies substantially with the industry of the new job creation. For example, young local workers benefit most from leisure & hospitality positions ($442) and least from professional & business services (PBS) positions ($279), while workers over 50 benefit most from education/health positions ($464) and least from information positions ($217). Manufacturing and PBS stimuli 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 the bottom quartile and the second highest gains for the top quartile. Information stimuli yield smaller local utility gains in general due to its greater propensity to hire distant workers, while education/health, which tends to hire locally and at all skill and experience levels, produces 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, new positions at the highest paying quartile of firms (regardless

of size) yield much larger gains for local high-paid workers ($\sim$ \$445) than low-paid ($\sim$ \$270) and unemployed ($\sim$ \$350) workers. By contrast, job creation at firms with below median average pay raises utility by $\sim$ \$240, $\sim$ \$370, and $\sim$ \$525 for the same three groups. Unemployed local workers in particular gain more utility from jobs created at small firms, suggesting that local officials can help them more by supporting 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.

5.3 Heterogeneity in National Incidence by Worker and Firm Characteristics

Tables 4 and 5 display cumulative shares of subpopulation-specific employment and welfare gains 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% and 44.6% of their nationwide employment gains accrue to those in the target tract's PUMA and state, respectively. 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 workers aged $\leq$ 30 (40.3%) than those 31-50 (37.0%) and over 50 (35.3%), while within-focal tract shares are twice as large for workers from the shock's industry as from other industries. Welfare has very similar patterns, with unemployed, low-paid, younger, and same industry workers all having much more locally concentrated gains than their high-paid, older, different industry counterparts.

Such heterogeneity in spatial decay rates suggests that the sizable variation in local incidence across subpopulations and shock compositions need not translate to the state or national level. To this end, column 1 of Table A7 reports the national shares of net employment gains by subpopulation, while Figure 4a plots these shares against their national workforce shares. Like its local counterpart 3a, Figure 4a shows that younger, lower-paid, and especially unemployed workers enjoy disproportionate shares of national employment gains from job stimuli. However, in contrast to 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 two factors. First, the set of positions vacated by workers taking stimulus jobs better approximates the U.S. establishment distribution than the original shock,

and each successive ripple of shock-induced reallocation yields an ever more generic vacancy composition. So workers from the targeted industry have an ever smaller advantage in securing vacated positions as distance from the shock grows. Second, due to the large surplus premium from staying at a job, most employed workers rarely seek other jobs, making them inelastic to job opportunities relative to the unemployed. Indeed, the only industry whose workers' national employment gain share exceeds its population share when getting 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 subpopulation shares of national welfare gains. Again, while the slightly disproportionate shares for higher-paid and younger workers match the local results, same-industry workers' share of national welfare gains (11.5%) is far smaller than their local share (34.8%), and only slightly exceeds their national workforce share (9.6%). Disaggregating to unemployed×age combinations (Figure A5b) reveals a second local vs. national discrepancy. Local mid-career and older unemployed workers disproportionately benefit relative to other locals, while nationally their utility gain share is below 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 state or national level incidence, again illustrating the value of the model.

The increasingly generic composition of vacated positions with greater distance also implies that which sector is targeted barely affects the shock's impact level nor its geographic, age, or initial earnings incidence beyond the surrounding PUMA; the within-PUMA share of net employment (utility) gains is between 5.3% and 6.6% (9.1% and 11.5%) regardless of chosen sector (Table A8). And shocks to all sectors yield shares of national employment and utility gains for 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 and particularly state and federal policymakers may safely ignore differences in demographic and geographic incidence when deciding among local initiatives targeting different sectors.

Changing the firm size/pay composition also barely shifts geographic, and more surprisingly, earnings and age incidence beyond nearby tracts (Tables A7 and A9). 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.

5.4 Local and National Incidence of Plant and Store Closings

Table 6's first row compares the average change in focal tract workers' employment rate (col. 1-2) and expected welfare (col. 5-6) 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 that we randomly selected from those with 500 or more positions of this type at baseline to ensure realistic targets for plant openings and closings. Since 250 new jobs is a smaller percent

change in these high employment tracts, it only raises the employment rate and welfare among focal tract workers by 0.03% and $150. However, a dramatic asymmetry is instantly apparent: the same-sized plant closing lowers these workers' employment rate by 0.59% and their annual welfare by a whopping $5,624. Focal tract workers account for 0.4% and 1.7% of national employment and welfare gains among plant openings and 8.6% and 35.7% of losses among closings.

What causes this asymmetry? Plant openings require new hires, and since local hiring still imposes hefty search and training costs, it only yields somewhat larger surplus than hiring more distant workers, so labor demand for locals only increases modestly. Our job creation simulations capture this by preventing new positions from being filled by "job stayers" (groups with $z(g) = 1$).[42] 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 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 versus aggregate job counts.

Figure 5a (Table 7) plots each subpopulation's share of all within-tract employment losses against its local workforce share among all 200 plant closing simulations. 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 targets 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. Notably, 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 workers 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 A10 displays estimates of cumulative shares of employment and welfare incidence within

[42] In Carballo and Mansfield (2025), 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.

various worker subpopulations by distance. Differences in spatial decay rates are even more striking for closings than openings, in part due to much greater local losses from closings. Only 10.6% and 14.4% of 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, those within 10 miles account for 27.1% of welfare losses among young workers compared to 44.3% and 49.6% for those 31-50 and over 50, and 82.0% among manufacturing workers versus 8.8% in non-manufacturing. This heterogeneity partly reflects the shock becoming more 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 aggregate incidence. However, distant workers who are higher-paid, older, and already in manufacturing also tend to have high rates of job staying, suggesting that their job matches are creating large surpluses that generally insulate them from the shock.[43] 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 subpopulations' 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 the latter are more likely to be initially employed at the closing plants. 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, high-paid and experienced workers outcompete low-paid and inexperienced workers for now scarcer positions, so that 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% nationally, as they tend to be the labor force's marginal workers. For welfare, national shares of losses do increase with initial pay quartile (Figure 6b), but 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 employment rate decrease (1.1%) and share of local employment losses (39.6%) for the lowest-paid quartile than the highest-paid (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 quartiles ($3,920 and $4,313) than the top two ($2,663 and $3,431), since low-paid

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

workers' greater exposure is partly offset by smaller baseline retention rates, so that more would have left jobs even absent the shock. The lower retention rate in retail/wholesale than manufacturing also explains why the average welfare loss among all 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, store and plant closings' patterns of national incidence are far more similar than local incidence would suggest (Fig. 9). As with plant closings, much smaller shares of employment and welfare losses stay within the target industry at the national than local level (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 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 communities also care about flows of workers away from targeted sites. Thus, Figures A7a and A7b (Tab. A11) display the change in focal tract workers' probability of ending up employed in each distance bin. The share who continue to work in the tract only falls by 4.5% and 3.6% for plant and store closings, even though both usually reduce tract employment by $\sim$10%. This is because local workers are better able to retain or obtain remaining jobs than would-be job movers from afar, but also because a large minority of locals would have taken jobs elsewhere anyway. Indeed, the store closing's lower displacement 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 the PUMA's other tracts 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. A11) presents destination distributions by subpopulation. Among those induced to switch locations by plant closings, high earners are much more likely than low earners to find distant jobs, with 87.7% vs. 72.6% finding work in a different PUMA and 46.6% vs. 23.1% changing states (Panel A), reflecting their respective baseline tendencies to make such moves from Table 1. For store closings (Panel B), which target 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 displace more younger workers due to their greater presence in retail, a smaller share of those displaced become unemployed compared to older workers, who are less able or willing to move to more distant jobs. We also see a small added outflow by local unemployed workers who would otherwise have found local jobs, illustrating the need to examine equilibrium reallocation rather than just the destinations of initially laid-off workers.

Finally, Figure A9 (Table A12) shows how spatial incidence evolves as the shock 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 size. Closings do exhibit a slight convexity in local employment rate changes with scale, as 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. 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.

5.5 Heterogeneity in Incidence by Target Tract Characteristics

Heterogeneity in geographic incidence also stems from the choice of target tract. To save space, here we summarize how the characteristics of the tract targeted by a plant or store opening shock affects its spatial and skill incidence, and provide a more complete analysis in Appendix A9.

We find that welfare and employment gains are considerably more geographically concentrated for shocks to rural relative to urban tracts. Within-PUMA workers enjoy 15.2% (8.8%) of welfare (employment) gains for targeted tracts in the bottom quintile of population density versus 5.4% (3.7%) for tracts in the top quintile, with similar differences when remoteness is instead measured using # of jobs within 5 miles or rent for an average two-bedroom apartment. Given that mean impacts increase nearly linearly with shock size, this suggests that targeting several rural areas with small development initiatives might produce larger local employment and welfare gains per job created than a single large plant opening in a dense urban area. High-poverty tracts also exhibit larger local welfare gains and within-PUMA shares of gains, suggesting that targeting poorer areas may yield greater local labor market benefits than for a typical tract.

Regressions relating shock incidence measures to several focal tract characteristics simultaneously confirm that these results hold conditionally as well. The regression results also echo two key findings discussed above. First, a one s.d. increase in a PUMA's share of manufacturing workers predicts small (0.68%) and trivial (0.04%) increases in within-PUMA shares of welfare and employment gains from a plant opening, consistent with shocks rapidly becoming generic with distance.

Second, in yet another discrepancy between local and national incidence, tract characteristics that predict greater employment gains for local low-paid workers tend to predict smaller gains for low-paid workers nationwide. In this case, reduced-form estimates of larger local treatment effects for low-paid workers could cause incorrect inferences about which focal area choices would best alleviate poverty, since larger gains for the local poor in certain local areas captured by such regressions would be outweighed by smaller expected gains among many less proximate workers.[44]

5.6 Robustness Checks

Table A17 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 utility results from a model with job multipliers. We adopt Bartik and Sother-

[44] One possible explanation is that these characteristics predict higher search costs that lead firms to hire local rather than distant low-paid workers (or distant high-paid workers whose vacated jobs are taken by lower-paid neighbors.)

land (2019)'s estimate that each new high-tech manufacturing job (presumably at large, high paying firms) generates 0.71 extra jobs after one year. We assume that increased product demand for local services is the dominant source of the multiplier, and add 250*.71 = 177 additional retail/wholesale and leisure/hospitality jobs. We distribute these jobs across within-PUMA tracts in proportion to their workers' shares of expected earnings gains from the baseline results. The augmented shock increases employment and welfare gains within the PUMA by only slightly more than the 171% multiplier, with modest shifts in employment and welfare gain shares toward surrounding tracts and away from the target tract. These results indicate that explicitly introducing job multipliers instead of treating 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 stock 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 position stock 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 by 4 positions (250 to 246), with a standard deviation of 7. This adjustment slightly reduces the scale of employment and welfare gains, but it barely changes distance bins' shares of gains, mitigating concerns about bias from treating the set of filled positions as exogenous.

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 distance distribution of employment rate changes and gain shares among workers are surprisingly similar to their baseline counterparts, reflecting very similar worker reallocation. 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. Since the Choo-Siow specification does not require analogues to Assumptions 1 and 2, this finding also provides reassurance about robustness to violations of the assumptions needed to aggregate CCPs to the group level.

Columns 5 and 9 examine sensitivity to allowing plant openings to change relative joint surpluses among job matches involving within-PUMA worker and firm types, perhaps due to differen-

tial changes in expected future local opportunities. 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 below), and re-scaling 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 5% smaller and 1% larger, respectively, than their baseline counterparts, perhaps because such establishments compete more with existing within-PUMA businesses. This suggests that unmodeled changes in continuation values and other surplus components may be small outside manufacturing, at least in the short run.

Table A18 assesses sensitivity of model predictions to restricting surplus heterogeneity in various ways. We focus on local welfare changes across initial earnings and industry categories, where worker and firm heterogeneity were shown to matter most. In column 2, we equalize joint surplus values across all categories of firm industry, size, and pay, so that location is the only firm characteristic. Because this removes complementarity between high-skilled workers and high-paying firms, it mistakenly predicts larger local welfare gains for unemployed and low-paid workers whose low baseline retention rates suggest they are more open to new 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 among those switching firms. This halves same-industry workers' welfare gain, thus understating the concentration of local gains. Finally, column 5 removes the surplus premium from job staying/retention, so that new jobs immediately create the same surplus as existing jobs. By 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 the baseline model's two-sided heterogeneity is needed to generate the disparities in local welfare gains presented above.

5.7 Model Validation

The estimated surplus parameters 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, parameters governing ordinary worker flows may not fully capture the response to sizable, locally focused demand 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 workers 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 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 A10 offers further detail.

Row 1 of Table 8 shows that on average the model would need to reallocate 35.1% of job matches of workers starting in the target PUMA to other groups g to perfectly match the true within-PUMA distribution. However, given the group space's granularity within PUMA, an accurate prediction often requires forecasting the exact employer a worker moved to, since wrongly predicting any of a worker's destination tract, industry, firm size, or firm pay category results in an incorrect group assignment. Thus, matching the full group distribution $P(g)$ is arguably an unreasonable standard.

When worker and position locations are collapsed post-simulation to 14 distance bins from the target tract, the share of job matches that must be reallocated across groups falls to 11.1% (row 2), suggesting that the model predicts the distance of workers' job transitions well, just not the exact destination tract. Moreover, 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 $P(g|f)$. With many million observed job matches, the risk of overfitting from using a highly saturated, just-identified model is far outweighed by the inability of a more parsimonious model (still featuring ~ 200 parameters!) to capture the data's rich matching patterns. The two-sided model also outperforms (though by much less) 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 has additional predictive value, even for smallish shocks. The baseline model also outperforms the Choo-Siow model, which assumes perfect instead of zero correlation in workers' preferences for positions within position types, particularly for more aggregated predictions. It also generates much more accurate predictions than the alternatives from Table A18 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 employment statuses after substantial local labor demand shocks.

6 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 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 for workers even 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 demographic and spatial composition of worker welfare gains 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 worked anyway without a positive shock, but have jobs at risk from negative local shocks, and removing the option to keep one's job creates large welfare losses.

Methodologically, we show that one can still forecast welfare incidence on both sides of a market from changes in either side's composition even when singles are unobserved on one or both sides. 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 fully exploits the massive scale of the LEHD data to capture multidimensional heterogeneity on both sides of a market without placing unjustified structure on the job matching technology. This approach could easily be 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.

Bilal, Adrien, "The geography of unemployment," *The Quarterly Journal of Economics*, 2023, *138* (3), 1507–1576.

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 K Mansfield, "Firm Trade Exposure, Labor Market Competition, and the Worker Incidence of Trade Shocks," Technical Report, National Bureau of Economic Research 2025.

Chan, Mons, Kory Kroft, Elena Mattana, and Ismael Mourifié, "An empirical framework for matching with imperfect competition," Technical Report, National Bureau of Economic Research 2024.

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.

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.

Davis, Steven J, R Jason Faberman, and John C Haltiwanger, "The establishment-level behavior of vacancies and hiring," *The Quarterly Journal of Economics*, 2013, *128* (2), 581–622.

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.

Martellini, Paolo, "Local labor markets and aggregate productivity," *Manuscript, UW Madison.[470]*, 2022.

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.

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 ≤ 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 ≤ 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 ≤ 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 A4 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 A4 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			
	Loc. 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.

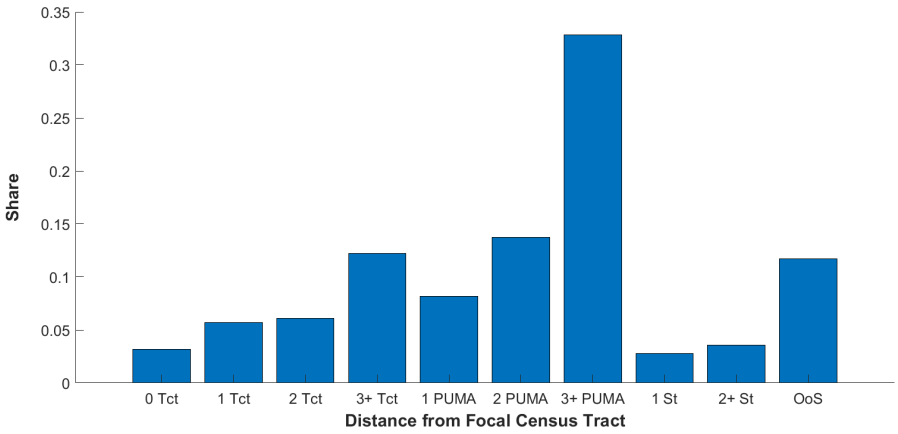
Table 8: Model Validation Results: Dissimilarity Index Values Comparing Forecasted and Actual Worker Reallocations Following Large Local Labor Demand Shocks Using Alternative Match Group Definitions and Methods for Generating Forecasts

	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)

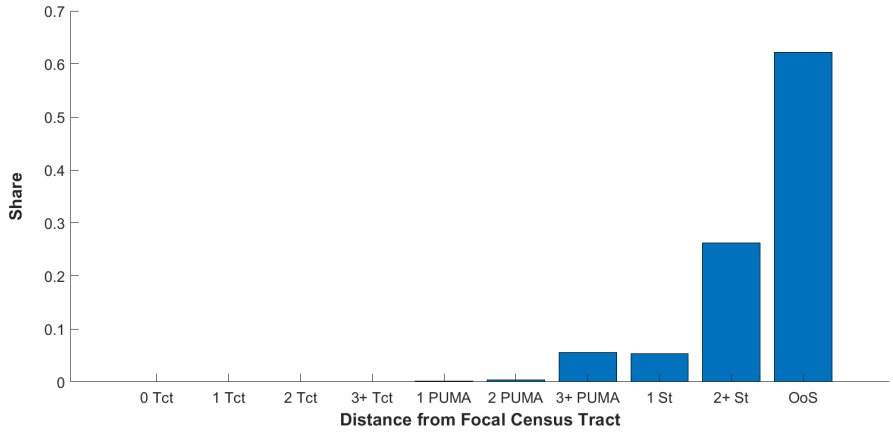
Notes: This table examines the fit of model-based predicted worker reallocations to the actual reallocations that occurred following a set of large establishment openings and closings in particular census tracts in particular years spanning 2003-2012. See Section A10 for a detailed description of the 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 A10 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



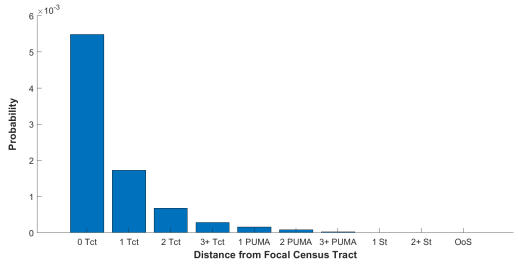
(b) Distribution of the Distance Between Workers' Initial Position and the Census Tract Targeted by the Simulated Stimulus Package: Average across All Simulated Stimuli



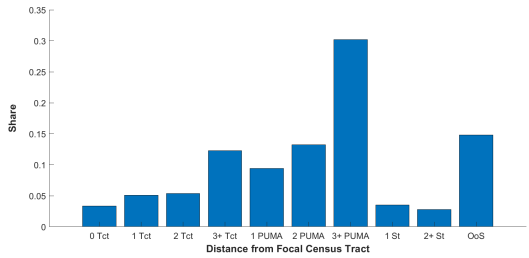
Notes: The bar heights in Figure 1a capture the shares of all worker transitions between dominant positions in 2012 and 2013 in which the geographic distance between these positions' establishments fell into the distance bins indicated by the bar labels. The bar heights in Figure 1b capture the shares 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

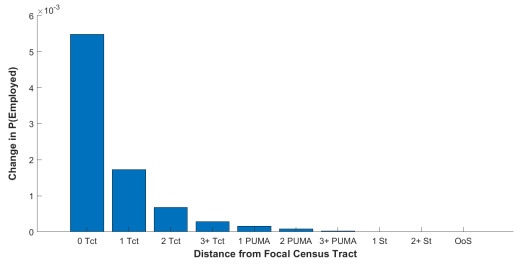
(a) Probability of Obtaining a Stimulus Job



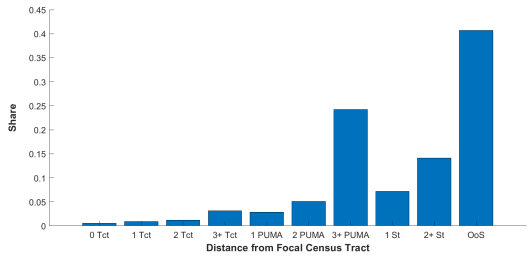
(b) Share of All Stimulus Jobs



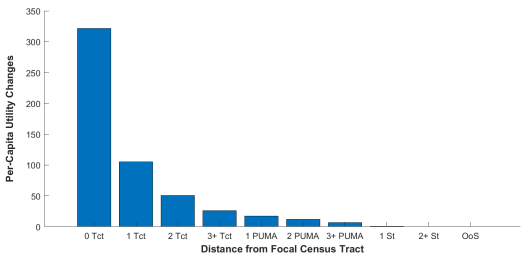
(c) Change in P(Employed)



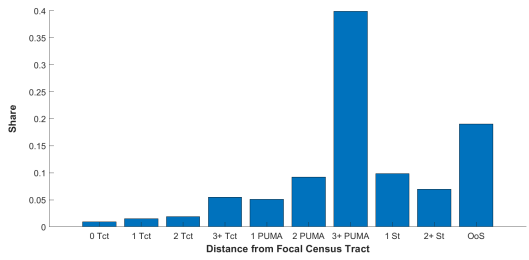
(d) Share of Employment Gains



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



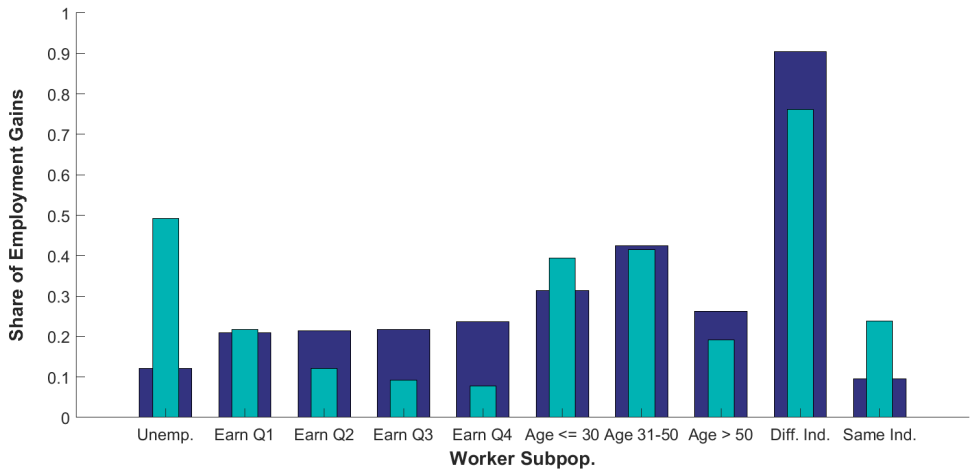
(f) Share of Welfare Gains



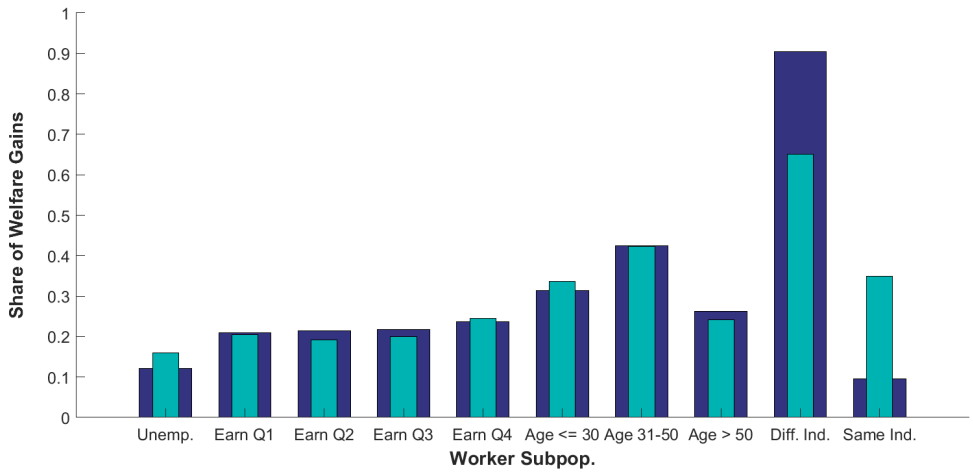
Notes: The bar heights in Figure 2a capture the average probability of obtaining a stimulus job among workers whose geographic distance between their initial establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. These probabilities 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



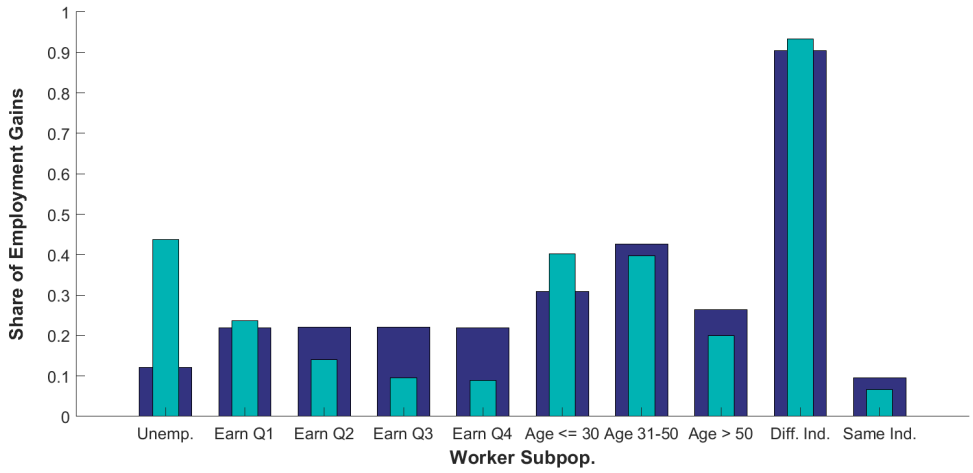
(b) Share of Focal Tract Utility Gains



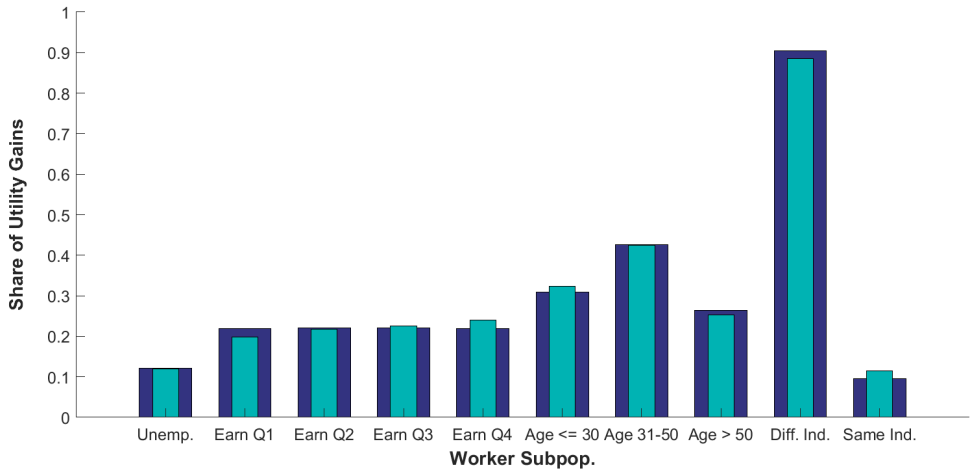
Notes: The heights of the wider bars within a particular group in Figures 3a and 3b capture the initial share of the focal tract workforce associated with the labeled worker subpopulation, while the heights of the narrower bars capture the subpopulation's share of the employment and job-related utility 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



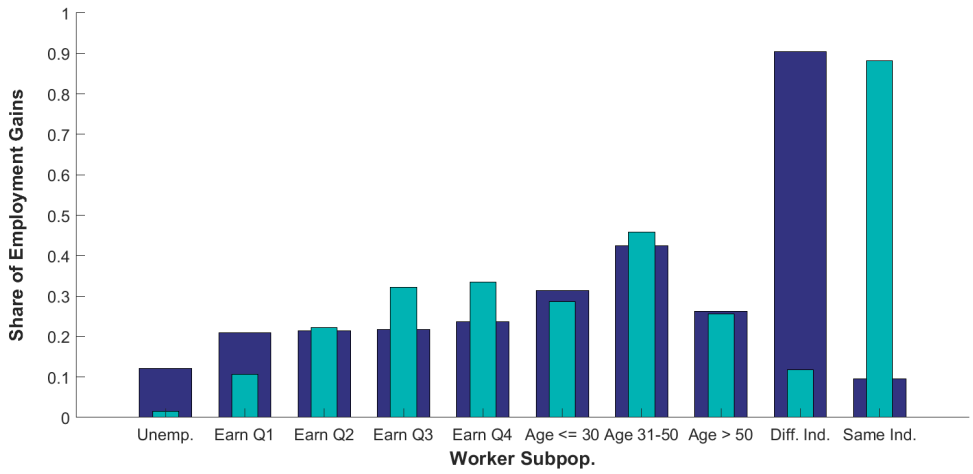
(b) Share of Total Utility Gains



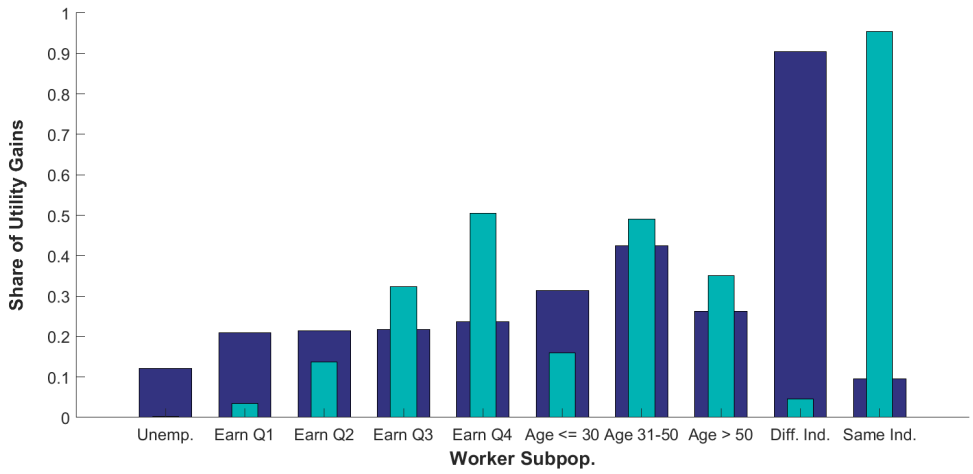
Notes: The heights of the wider bars within a particular group in Figures 4a and 4b capture the initial share of the national workforce associated with the labeled worker subpopulation, while the heights of the narrower bars capture the subpopulation's share of national employment and job-related utility 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



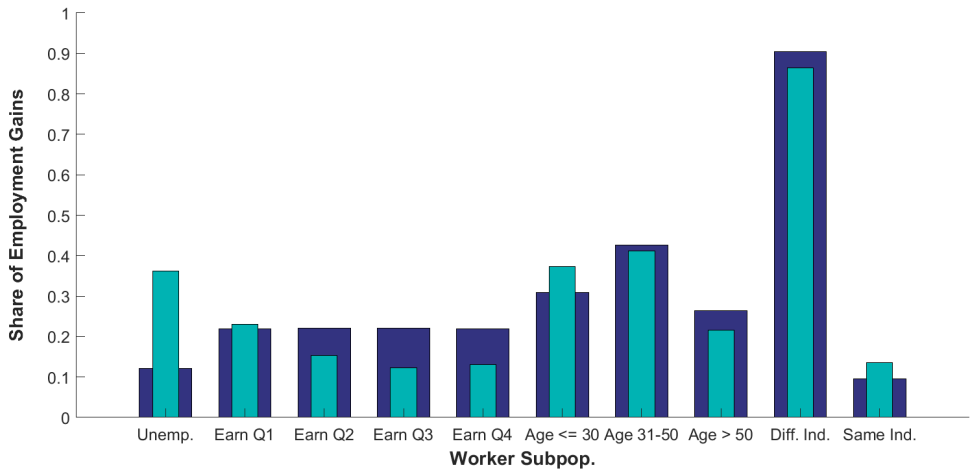
(b) Share of Focal Tract Utility Losses



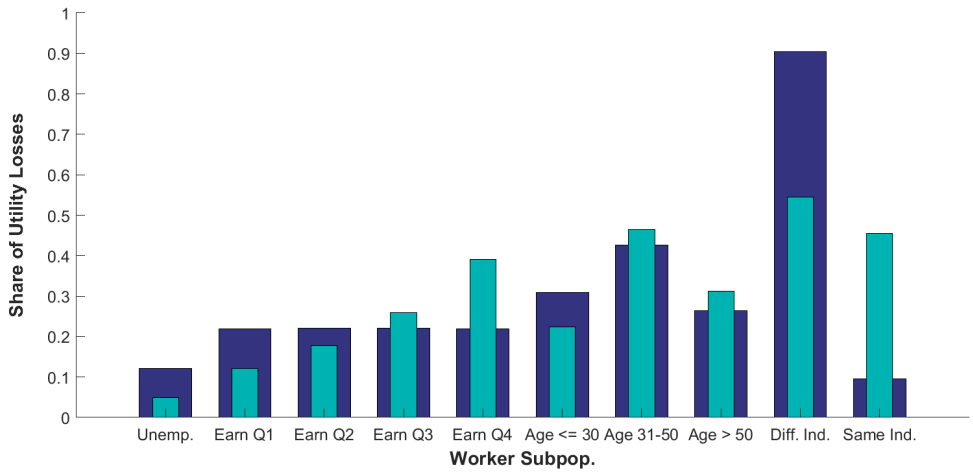
Notes: The heights of the wider bars within a particular group in Figures 5a and 5b capture the initial share of the focal tract workforce associated with the labeled worker subpopulation, while the heights of the narrower bars capture the subpopulation's share of the local employment and job-related 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



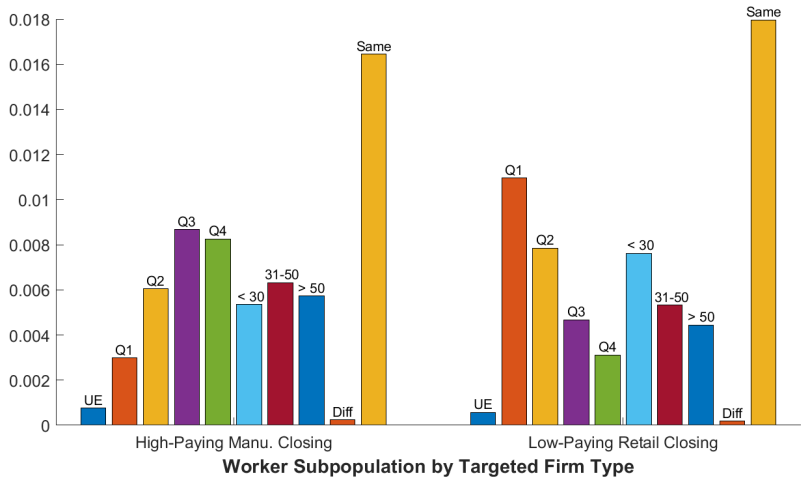
(b) Share of Focal Tract Utility Losses



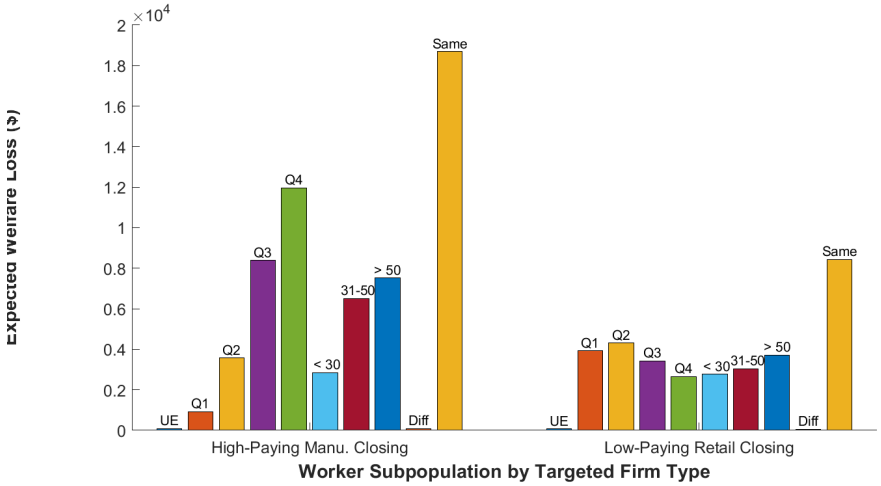
Notes: The heights of the wider bars within a particular group in Figures 6a and 6b capture the initial share of the national workforce associated with the labeled worker subpopulation, while the heights of the narrower bars capture the subpopulation's share of the national employment and job-related 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)



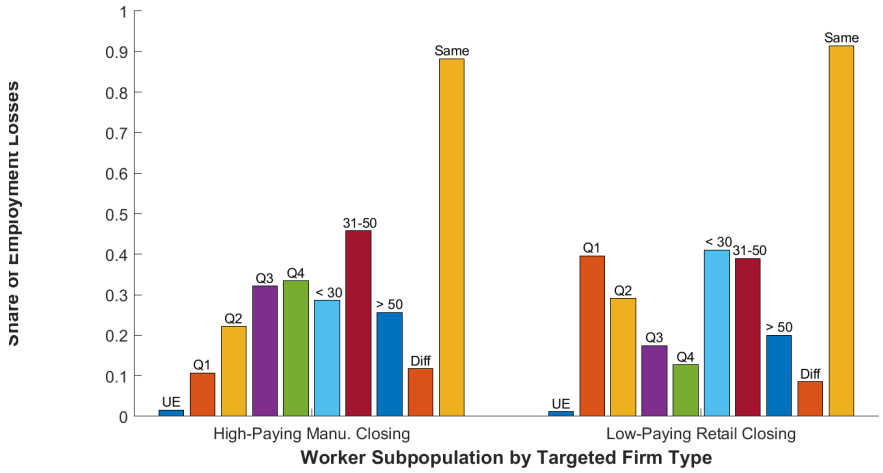
(b) Expected Utility Loss (in 2023 $)



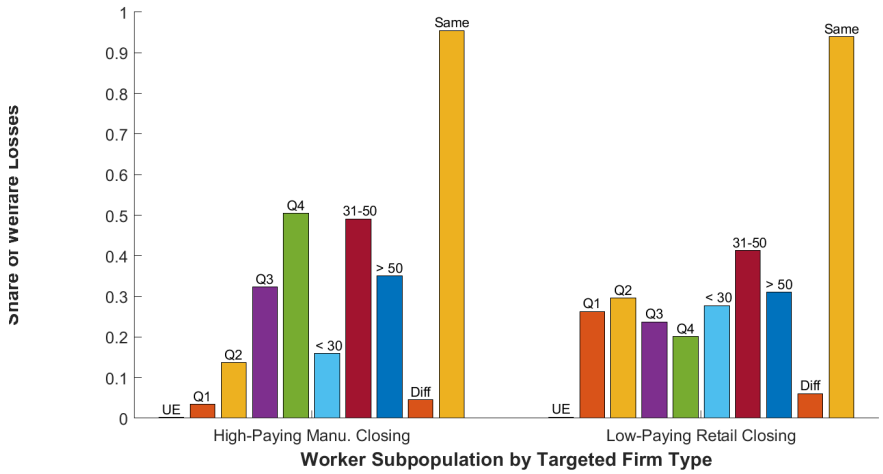
Notes: The bar heights within a particular group in Figures 7a and 7b capture the change in unemployment rate and expected utility, respectively, from two sets of simulated plant/store closings among workers who were employed (or unemployed) in the focal tract in the previous year and who belong to the 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



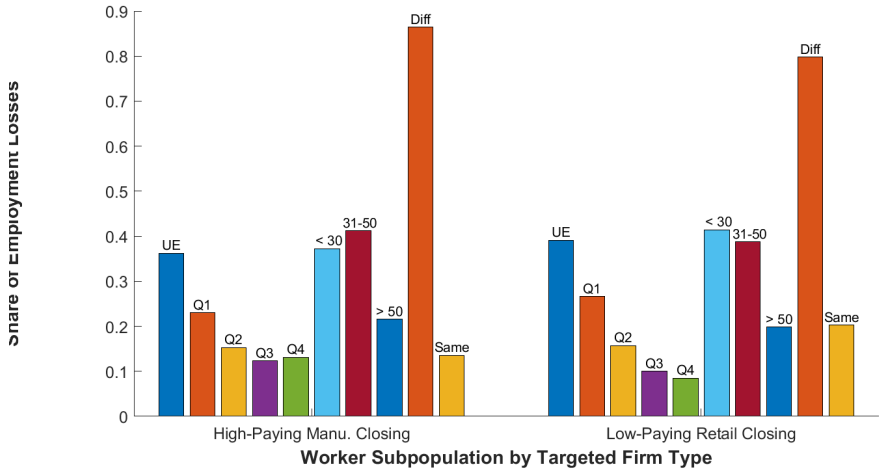
(b) Share of Focal Tract Welfare Losses



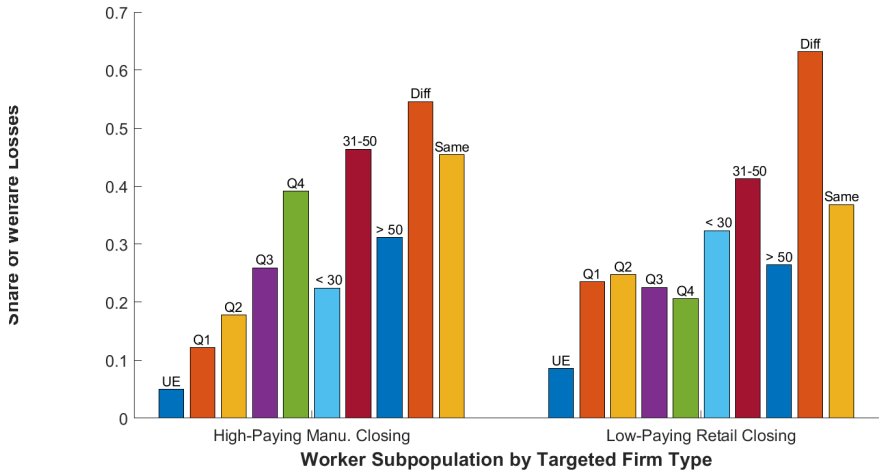
Notes: The bar heights within a particular group in Figures 8a and 8b capture the average share of local employment and welfare losses accruing to the worker subpopulations labeled over the bar among workers from the focal tract from two sets of simulated plant/store closings. See Figure 3 for expanded 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.

Online Appendix

A1 Proof of Proposition A1

Proposition A1:

Let $|l|$ and $|g_k|$ denote, respectively, the number of workers classified as worker type l and the number of workers whose job match would be classified as group g (either stayers or new hires among those in l) if hired by position k (a subset of the workers in $l(g)$). In addition, let $n(l)$ denote the share of all workers assigned to worker type l, so that $|l| = n(l)I$. Further, define C_l as the mean value of $e^{-\frac{r_i}{\sigma}}$ for a given worker type l. Define $S_{g|l,k} \equiv \frac{|g_k|}{|l|}$ as the share of workers of worker type l who would be assigned to group g if they filled position k (i.e. the share of workers who are incumbents at the firm if $z(g) = 1$, the share who would be within-industry job movers if $z(g) = 2$, and the share who would be industry switchers if $z(g) = 0$), and define $\overline{S}_{g|l,f} \equiv \frac{1}{|f|}\sum_{k\in f} S_{g|l,k}$ to be the mean of $S_{g|l,k}$ among all k assigned to position type f. Suppose the following assumptions hold:

$$\textit{Assumption 1:}\ \ \frac{1}{|g_k|}\sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}} \approx \frac{1}{|l|}\sum_{i:l(i)=l(g)} e^{-\frac{r_i}{\sigma}} = C_{l(g)}\ \forall (g,k) \tag{8}$$

$$\textit{Assumption 2:}\ \ S_{g|l,k} \approx \overline{S}_{g|l,f}\ \forall\ k,\ \forall\ g \tag{9}$$

Then the equilibrium aggregate group-level choice probabilities can be written as follows:

$$P(g|f) = \frac{e^{\frac{\theta_g}{\sigma}}\overline{S}_{g|l,f}n(l)C_l}{\sum_{l'\in\mathcal{L}}\sum_{g'\in(l,f)} e^{\frac{\theta_{g'}}{\sigma}}\overline{S}_{g'|l',f}n(l')C_{l'}} \tag{10}$$

Proof: Building off the second welfare theorem, Shapley and Shubik (1972) show that Walrasian equilibrium assignment in this game maximizes a linear programming problem. This then implies that the unique stable assignment can also be found by solving the dual problem: identifying a set of worker utility values $\{r_i\}$ and position profit values $\{q_k\}$ that minimize the total "cost" of all workers and positions, $\sum_{i\in\mathcal{I}} r_i + \sum_{k\in\mathcal{K}} q_k$, subject to the constraint that these values cannot violate the underlying joint surplus values: $r_i + q_k \geq \pi_{ik}\ \forall\ (i,k)$. Crucially, inspection of the problem reveals that the stable assignment is fully determined by the joint surplus values $\{\pi_{ik}\}$; no separate information on the worker and firm components π^i_{ik} and π^k_{ik} is needed. Following GS, this dual problem yields the following conditions that define the optimal assignment:

$$\mu_{ik} = 1 \text{ iff } k \in \arg\max_{k\in\mathcal{K}\cup 0} \pi_{ik} - q_k \text{ and } i \in \arg\max_{i\in\mathcal{I}\cup 0} \pi_{ik} - r_i \tag{11}$$

Given optimal worker and position payoffs $\{r_i\}$ and $\{q_k\}$ from the dual solution, Shapley and

Shubik (1972) show how to decentralize this optimal assignment via a set of earnings transfers w_{ik}:

$$w_{ik} = \pi^k_{ik} - q_k \tag{12}$$

Because $r_i + q_k = \pi_{ik} \equiv \pi^i_{ik} + \pi^k_{ik}$ for any matched pair (i, k) in the stable match, this implies:

$$w_{ik} = r_i - \pi^i_{ik} \tag{13}$$

Using (12) and (13), the conditions (11) can be rewritten as the standard requirements that worker and establishment choices must be utility- and profit-maximizing, respectively:

$$\mu_{ik} = 1 \text{ iff } k \in \arg\max_{k \in \mathcal{K} \cup 0} \pi^i_{ik} + w_{ik} \text{ and } i \in \arg\max_{i \in \mathcal{I} \cup 0} \pi^k_{ik} - w_{ik} \tag{14}$$

Given candidate equilibrium payoffs $\{r_i\}$ combined with the i.i.d. Type 1 EV assumption for ϵ_{ik}, Decker et al. (2013) show that the probability that hiring (or retaining) i maximizes k's payoff is given by:

$$P(i|k) = \frac{e^{\frac{\theta_g - r_i}{\sigma}}}{\sum_{i' \in \mathcal{I}} e^{\frac{\theta_{g'} - r_{i'}}{\sigma}}} \tag{15}$$

Next, note that the law of total probability implies:

$$\begin{aligned} P(g|f) &= \sum_{k \in f} P(g|f,k)P(k|f) = \frac{1}{|f|} \sum_{k \in f} P(g|k) = \frac{1}{|f|} \sum_{k \in f} \sum_{i:g(i,k)=g} P(i|k) \\ &= \frac{1}{|f|} \sum_{k \in f} \sum_{i:g(i,k)=g} \frac{e^{\frac{\theta_g - r_i}{\sigma}}}{\sum_{i' \in \mathcal{I}} e^{\frac{\theta_{g'} - r_{i'}}{\sigma}}} = \frac{1}{|f|} \sum_{k \in f} \frac{(e^{\frac{\theta_g}{\sigma}})(\sum_{i:g(i,k)=g} e^{\frac{-r_i}{\sigma}})}{\sum_{i' \in \mathcal{I}} e^{\frac{\theta_{g'} - r_{i'}}{\sigma}}}, \end{aligned} \tag{16}$$

where $|f|$ captures the number of positions k assigned to position type f.

Assumption 1 imposes that the mean exponentiated worker utility values $e^{\frac{-r_i}{\sigma}}$ vary minimally across groups g featuring the same worker type $l(g)$. Given the characteristics used to define l and g in the empirical application, this states that existing employees (potential stayers) and non-employees of each establishment (both from the same industry and from other industries) have approximately the same mean value of r_i among workers in the same age category whose initial jobs were in the same local area and pay category. In other words, the payoffs that workers in the same initial earnings and age class require in equilibrium will not differ systematically across establishments within a small local area. This becomes a better approximation as more characteristics and categories are used to define a worker type $l(i)$.

Assumption 2 imposes that the share of potential stayers vs. new hires from the same and from different industries among workers from each worker type l is common across establishments within position type f. In the chosen context, this means that establishments in the same geographic area, industry supersector, and establishment size and average pay categories have roughly the same

number and past pay and age composition of employees. This assumption is necessary because the probability of filling a position with an existing employee depends on how many employees one already has, so that without it the worker retention rate depends on sizes and worker type compositions of establishments that are at risk of retaining a worker.

Importantly, mean surpluses among (l, f) pairs are identified without imposing Assumptions 1 and 2, so that the assumptions are only necessary to isolate the surplus from hiring a within-firm incumbent relative to a worker from another firm within the same census tract. Violations lead to slight over or understatement of deviations among (l, f) type combinations from the average surplus premium for job staying in the population. In the absence of Assumption 1, the probability that a worker is an establishment stayer instead of a mover depends on the kinds of workers within the worker type that sorted into particular establishments within the firm type. If the establishments that are experiencing job loss within the firm type are particularly populated at baseline by workers with higher required r_i levels relative to the mean for their type, then the job retention rate predicted under Assumption 1 might slightly overstate how much retention would really occur. As long as there is not extreme segregation by workers' unmodeled utility requirements across establishments within a firm type, the predicted probabilities are unlikely to be very sensitive to Assumption 1.

Similarly, when Assumption 2 fails, it is possible that some establishments within a position type have disproportionate shares of type l workers at baseline relative to others. Because the job retention rate does not increase linearly with the share of potential job stayers for given surplus premium from job retention, changes in the concentration of potential stayers within particular establishments slightly changes the predicted share of (l, f) matches that consist of job stayers rather than job movers.

Note first that Assumption 2 implies that $|g_k| \equiv S_{g|l,k} n(l(g))I \approx \overline{S}_{g|l,f} n(l(g))I$. Thus, Assumptions 1 and 2 together imply:

$$\sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}} \approx \overline{S}_{g|l(g),f(g)} n(l(g))(I) C_{l(g)}. \tag{17}$$

Applying this result to the last expression in (16), one obtains:

$$\begin{aligned} P(g|f) &= \sum_{k \in f} \left(\frac{1}{|f|}\right) \frac{e^{\frac{\theta_g}{\sigma}} \sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}}}{\sum_{i' \in \mathcal{I}} e^{\frac{\theta_{g'} - r_{i'}}{\sigma}}} = \sum_{k \in f} \left(\frac{1}{|f|}\right) \frac{e^{\frac{\theta_g}{\sigma}} \sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}}}{\sum_{l' \in \mathcal{L}} \sum_{g' \in (l,f)} \sum_{i':g(i',k)=g'} e^{\frac{\theta_{g'} - r_{i'}}{\sigma}}} \\ &= \sum_{k \in f} \left(\frac{1}{|f|}\right) \frac{e^{\frac{\theta_g}{\sigma}} \overline{S}_{g|l,f} n(l)(I) C_l}{\sum_{l' \in \mathcal{L}} \sum_{g' \in (l,f)} e^{\frac{\theta_{g'}}{\sigma}} \overline{S}_{g'|l',f} n(l')(I) C_{l'}} \\ &= \frac{e^{\frac{\theta_g}{\sigma}} \overline{S}_{g|l,f} n(l)(I) C_l}{\sum_{l' \in \mathcal{L}} \sum_{g' \in (l,f)} e^{\frac{\theta_{g'}}{\sigma}} \overline{S}_{g'|l',f} n(l')(I) C_{l'}} \sum_{k \in f} \left(\frac{1}{|f|}\right) = \frac{e^{\frac{\theta_g}{\sigma}} \overline{S}_{g|l,f} n(l) C_l}{\sum_{l' \in \mathcal{L}} \sum_{g' \in (l,f)} e^{\frac{\theta_{g'}}{\sigma}} \overline{S}_{g'|l',f} n(l') C_{l'}} \end{aligned} \tag{18}$$

This concludes the proof.

A2 Proof of Proposition A2

Proposition A2:

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 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* $h^{CF}(*)$.

Proof: We prove Proposition A2 by construction.

Let $z(i,k)$ represent a trichotomous variable that takes on the value of 1 if the firms associated with positions $j(i)$ and k are the same $(1(m(j)=m(k)))$, 2 if the industries (but not the firms) associated with positions $j(i)$ and k are the same $(1(s(j)=s(k))\&m(j)\neq m(k))$ and 0 otherwise. Recall also that all job matches assigned to the same match group g share values of the worker and establishment characteristics that define the worker and position types l and f, respectively, as well as the value of $z(i,k)$. Thus, one can write $l(g)$, $f(g)$ and $z(g)$ for any group g. Let the worker types be ordered (arbitrarily) from $l=1\ldots l=L$, and let the position types be ordered (arbitrarily) from $f=1\ldots f=F$. Let $g(l,f,z)$ denote the group associated with worker type l, position type f, and existing worker indicator z. Assume that the set $\Theta^{D-in-D} = \{\frac{(\theta_g-\theta_{g'})-(\theta_{g''}-\theta_{g'''})}{\sigma} \forall\, (g,g',g'',g''')\}$ is known, since a consistent estimator for each element of the set can be obtained via adjusted log odds ratios, as described in Section 2.2. Consider defining the set of alternative group-level joint surplus values $\tilde{\Theta} = \{\tilde{\theta}_g\}$ as follows:

$$\tilde{\theta}_{g'} = 0 \;\forall\; g' : (l(g')=1 \text{ and/or } f(g')=1) \text{ and } z(g')=0 \tag{19}$$

$$\tilde{\theta}_{g'} = \frac{(\theta_{g'} - \theta_{g(1,f(g'),0)}) - (\theta_{g(l(g'),1,0)} - \theta_{g(1,1,0)})}{\sigma} \;\forall\; g' : (f(g')\neq \text{ and } l(g')\neq 1) \text{ and/or } z(g')\neq 0 \tag{20}$$

Under the definitions in (19) and (20), we have:

$$\frac{(\tilde{\theta}_g - \tilde{\theta}_{g'}) - (\tilde{\theta}_{g''} - \tilde{\theta}_{g'''})}{\sigma} = \frac{(\theta_g - \theta_{g'}) - (\theta_{g''} - \theta_{g'''})}{\sigma}$$
$$\forall\, (g,g',g'',g''') : l(g)=l(g''), l(g')=l(g'''), f(g)=f(g'), f(g'')=f(g''') \tag{21}$$

Thus, the appropriate difference-in-differences using elements of $\tilde{\Theta}$ match their analogues among the true surpluses in Θ^{D-in-D}, so that all the information about Θ in the identified set Θ^{D-in-D} is

retained. And unlike the true set Θ, the construction of $\tilde{\Theta}$ only requires knowledge of Θ^{D-in-D}.

Next, note that the elements of $\tilde{\Theta}$ can be written in the following form:

$$\tilde{\theta}_g = \theta_g + \Delta^1_{l(g)} + \Delta^2_{f(g)} \; \forall \; g \in \mathcal{G}, \text{ where} \tag{22}$$

$$\Delta^1_{l(g)} = \theta_{g(l(g),1,0)} - \theta_{g(1,1,0)} \quad \text{and} \quad \Delta^2_{f(g)} = \theta_{g(1,f(g),0)} \tag{23}$$

where $\mathcal{G}$ is the set of all possible match groups. In other words, each alternative surplus $\tilde{\theta}_g$ equals the true surplus θ_g plus a constant $(\Delta^1_{l(g)})$ that is common to all groups featuring the same worker type and a constant $(\Delta^2_{f(g)})$ that is common to all groups featuring the same position type.

Next, recall that there exists a unique aggregate assignment associated with each combination of marginal worker and position type distributions $n^{CF}(l)$ and $h^{CF}(f)$ and set of group-level surpluses, including $\tilde{\Theta}$. Let $\tilde{P}^{CF}(*) \equiv P^{CF}(*|\tilde{\Theta}, \tilde{C}^{CF}_2, \ldots, \tilde{C}^{CF}_L))$ represent the assignment that results from combining arbitrary marginals $n^{CF}(l)$ and $h^{CF}(f)$ with $\tilde{\Theta}$. $\tilde{\mathbf{C}}^{\mathbf{CF}} = [1, \tilde{C}^{CF}_2 \ldots \tilde{C}^{CF}_L]$ denotes the vector of mean exponentiated utility values for each worker type l (with $\tilde{C}^{CF}_1$ normalized to 1) that solves the system of excess demand equations below, and thus yields $\tilde{P}^{CF}(g) \; \forall \; g \; \in \mathcal{G}$ when plugged into equation (5) along with the elements of $\tilde{\Theta}$, n^{CF} and $\overline{S}^{CF}_{g'|l(g'),d}$:

$$\begin{aligned}
&\sum_{f \in \mathcal{F}} h^{CF}(f)\Big(\sum_{g:l(g)=2} P^{CF}(g|f, \tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}})\Big) = n^{CF}(2) \\
&\vdots \\
&\sum_{f \in \mathcal{F}} h^{CF}(f)\Big(\sum_{g:l(g)=L} P^{CF}(g|f, \tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}})\Big) = n^{CF}(l)
\end{aligned} \tag{24}$$

We wish to show that $\tilde{P}^{CF}(*) \equiv P^{CF}(*|\tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}})$ will be identical to the alternative unique counterfactual equilibrium assignment $P^{CF}(*|\Theta, \mathbf{C}^{\mathbf{CF}})$ that combines the same arbitrary marginal distributions $n^{CF}(l)$ and $h^{CF}(f)$ with the set Θ instead of $\tilde{\Theta}$. Here, $\mathbf{C}^{\mathbf{CF}} = [1, C^{CF}_2 \ldots C^{CF}_L]$ denotes a vector of l-type-specific mean exponentiated utility values that clears the market by satisfying the following alternative excess demand equations:[45]

$$\begin{aligned}
&\sum_{f \in \mathcal{F}} h^{CF}(f)\Big(\sum_{g:l(g)=2} P^{CF}(g|f, \Theta, \mathbf{C}^{\mathbf{CF}})\Big) = n^{CF}(2) \\
&\vdots \\
&\sum_{f \in \mathcal{F}} h^{CF}(f)\Big(\sum_{g:l(g)=L} P^{CF}(g|f, \Theta, \mathbf{C}^{\mathbf{CF}})\Big) = n^{CF}(l)
\end{aligned} \tag{25}$$

Since all other terms are shared by the systems (24) and (25), it suffices to show that $P^{CF}(g|f, \tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}}) =$

[45] Note that we have suppressed the dependence of $P^{CF}(*|\Theta, \mathbf{C}^{\mathbf{CF}}, n^{CF}(l), h^{CF}(f), \overline{S}_{g|l,f})$ on $n^{CF}(l)$, $h^{CF}(f)$, and $\overline{S}_{g|l,f}$ because these are held fixed across the two alternative counterfactual simulations.

$P^{CF}(g|f,\Theta,\mathbf{C^{CF}}))\ \forall\ g \in \mathcal{G}$ for some vector $\mathbf{C^{CF}}$. Consider the following vector $\mathbf{C^{CF}}$:

$$C_l^{CF} = \tilde{C}_l^{CF} e^{\frac{\Delta_l^1}{\sigma}}\ \forall\ l \in [2,\ldots,L] \tag{26}$$

where Δ_l^1 is as defined in (23). For an arbitrary choice of g, we obtain:

$$\begin{aligned}
P^{CF}(g|f(g),\tilde{\Theta},\tilde{\mathbf{C}}^{\mathbf{CF}}) &= \frac{e^{\frac{\tilde{\theta}_g^{CF}}{\sigma}} \overline{S}_{g|l(g),f(g)}^{CF} n^{CF}(l(g)) \tilde{C}_l^{CF}}{\sum_{l'\in\mathcal{L}} \sum_{g'\in(l',f)} e^{\frac{\tilde{\theta}_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l'(g'),f(g)}^{CF} n^{CF}(l') \tilde{C}_{l'}^{CF}} \\
&= \frac{e^{\frac{(\theta_g^{CF}+\Delta_{l(g)}^1+\Delta_{f(g)}^2)}{\sigma}} \overline{S}_{g|l(g),f(g)}^{CF} n^{CF}(l(g)) C_l^{CF} e^{\frac{-\Delta_l^1}{\sigma}}}{\sum_{l'\in\mathcal{L}} \sum_{g'\in(l',f)} e^{\frac{(\theta_{g'}^{CF}+\Delta_{l(g')}^1+\Delta_{f(g')}^2)}{\sigma}} \overline{S}_{g'|l'(g'),f(g)}^{CF} n^{CF}(l') C_{l'}^{CF} e^{\frac{-\Delta_{l'}^1}{\sigma}}} \\
&= e^{\frac{\Delta_{l(g)}^1}{\sigma}} e^{\frac{\Delta_{f(g)}^2}{\sigma}} e^{\frac{-\Delta_{l(g)}^1}{\sigma}} \frac{e^{\frac{\theta_g^{CF}}{\sigma}} \overline{S}_{g|l(g),f(g)}^{CF} n^{CF}(l(g)) C_l^{CF}}{e^{\frac{\Delta_{f(g)}^2}{\sigma}} \sum_{l'\in\mathcal{L}} e^{\frac{\Delta_{l(g')}^1}{\sigma}} e^{\frac{-\Delta_{l(g')}^1}{\sigma}} \sum_{g'\in(l',f)} e^{\frac{\theta_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l'(g'),f(g)}^{CF} n^{CF}(l') C_{l'}^{CF}} \\
&= \frac{e^{\frac{\theta_g^{CF}}{\sigma}} \overline{S}_{g|l(g),f(g)}^{CF} n^{CF}(l(g)) C_l^{CF}}{\sum_{l'\in\mathcal{L}} \sum_{g'\in(l',f)} e^{\frac{\theta_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l'(g'),f(g)}^{CF} n^{CF}(l') C_{l'}^{CF}} = P^{CF}(g|f,\Theta,\mathbf{C^{CF}})
\end{aligned} \tag{27}$$

This proves that $P^{CF}(g|f,\Theta,\mathbf{C^{CF}})$ also satisfies the market clearing conditions (25) above, and will therefore be the unique group-level assignment consistent with marketwide equilibrium and stability. Thus, we have shown that the counterfactual assignment that is recovered when using an alternative set of surpluses $\tilde{\Theta}$ derived from the identified set Θ^{D-in-D} will in fact equal the counterfactual assignment we desire, which is based on the true set of joint surplus values Θ. Furthermore, while worker-type specific mean utility values $\tilde{\mathbf{C}}^{\mathbf{CF}}$ that clear the market given $\tilde{\Theta}$ will differ for each worker type from the corresponding vector $\mathbf{C^{CF}}$ based on the true surplus set Θ, these differences are invariant to the marginal worker type and position type distributions $n^{CF}(l)$ and $h^{CF}(f)$ used to define the counterfactual. This implies that differences in utility gains caused by alternative counterfactuals among worker types are identified, permitting comparisons of the utility incidence of alternative labor supply or demand shocks. This concludes the proof.

A3 Proof of Proposition A3

Proposition A3:

Suppose the following assumptions hold:

1') *The assumptions laid out in section 2 continue to hold. Namely, each joint surplus π_{ik} is additively separable in the group-level and idiosyncratic components, the vector of idiosyncratic*

components ϵ_{ik} is independently and identically distributed, and follows the type 1 extreme value distribution, and Assumptions 1 and 2 hold.

2') *The set of destination positions $k \in \tilde{\mathcal{K}}$ that will be filled in the stable counterfactual assignment are known in advance, and the set of destination positions $k \in \tilde{\mathcal{K}}$ that will remain unfilled in the stable counterfactual assignment are ignorable, in the sense that their existence does not change the assignment nor the division of surplus among the remaining set of positions $\mathcal{K}$ and set of workers $\mathcal{I}$.*

3') $\frac{1}{|g_i|}\sum_{k:g(i,k)=g} e^{-\frac{q_k}{\sigma}} \approx \frac{1}{|f|}\sum_{k:f(k)=f(g)} e^{-\frac{q_k}{\sigma}} = C_{f(g)}\ \forall (g,i).$

4') $P(g|i,f(g)) \approx P(g|l(g),f(g))\ \forall (g,i).$

Then the group-level assignment $P^{CF}(g)$ that satisfies the following $L-1$ excess demand equations represents the unique group-level equilibrium assignment $P^{CF^*}(g)$ consistent with the unique worker/position level stable matching μ^{CF}:

$$\begin{aligned}
&\sum_{f\in\mathcal{F}} h^{CF}(f)\Big(\sum_{g:l(g)=2} P^{CF}(g|f,C_2^{CF},\ldots,C_L^{CF}) = n^{CF}(2)\\
&\vdots\\
&\sum_{f\in\mathcal{F}} h^{CF}(f)\Big(\sum_{g:l(g)=L} P^{CF}(g|f,C_2^{CF},\ldots,C_L^{CF}) = n^{CF}(L)
\end{aligned} \tag{28}$$

where $P^{CF}(g|f,C_2^{CF},\ldots,C_L^{CF})$ is given by:

$$P^{CF}(g|f) = \frac{e^{\frac{\theta_g^{CF}}{\sigma}}\overline{S}^{CF}_{g|l(g),f} n^{CF}(l(g)) C_l^{CF}}{\sum_{l'\in\mathcal{L}}\sum_{g'\in(l,f)} e^{\frac{\theta_{g'}^{CF}}{\sigma}}\overline{S}^{CF}_{g'|l(g'),d} n^{CF}(l') C_{l'}^{CF}} \ \forall\ f\in[1,\ldots,F] \tag{29}$$

Proof: Proposition A3 states that assignment $P^{CF}(g)$ implied by the vector of mean utility values $\mathbf{C^{CF}} = [1, C_2, \ldots, C_L^{CF}]$ that solves the system of equations (28) in fact represents the unique group-level stable (and equilibrium) assignment $P^{CF^*}(g)$.

First, note that if unfilled positions are ignorable for the counterfactual assignment, then we can focus on finding a stable assignment of a restricted version of the assignment game in which only remaining K positions need to be considered. As discussed in section 2.3, Assumption 2' implicitly requires that no position that remains unfilled is ever the second-best option for any worker who takes a job in the destination period.

Furthermore, Assumption 2' imposes that each of the remaining positions will be filled in any

stable matching. Recall that stability in the individual-level matching μ^{CF} requires:

$$\mu_{ik}^{CF} = 1 \text{ iff } k \in \arg\max_{k \in \tilde{\mathcal{K}} \cup 0} \pi_{ik} - q_k^{CF} \text{ and } i \in \arg\max_{i \in \tilde{\mathcal{I}} \cup 0} \pi_{ik} - r_i^{CF} \tag{30}$$

Assumption 2' allows us to replace $i \in \arg\max_{i \in \tilde{\mathcal{I}} \cup 0} \pi_{ik} - r_i^{CF}$ with $i \in \arg\max_{i \in \tilde{\mathcal{I}}} \pi_{ik} - r_i^{CF}$. In other words, we assume in advance that the individual rationality conditions that any proposed match yield a higher payoff to the position than remaining vacant, $\pi_{ik} - r_i > \pi_{0k}$ when $\mu_{ik} = 1$, are satisfied and can be ignored. Implicitly, this requires that the joint surpluses to workers and firms from matching up are sufficiently large relative to both workers' and firms' outside options.[46] Imposing Assumption 2' may cause utility losses among local workers from negative local labor demand shocks to be overstated, since some workers would likely find jobs at positions that were not willing to hire at the original wage level but would enter the labor market at lower wage levels. Conversely, gains to local workers from positive shocks may be understated, since some local firms that filled positions at the original wage levels might choose to remain vacant (or move to other locations) when competition for local workers becomes more fierce.

In our applications the number of positions that will be filled is greater than the number of workers seeking positions (I). In order to be able to consistently allocate workers to match groups, even when they move to (or remain in) nonemployment, we define a "nonemployment" position type as the last position type F. Because the number of workers who end up nonemployed is assumed to be known, we allocate enough "nonemployment" positions within type F, $h^{CF}(F)$, so that the number of workers I equals the number of "positions" K, once K includes the dummy nonemployment positions. We then normalize this common number of workers and firm positions (assumed to be very large) to be 1, and reinterpret $n^{CF}(l)$ and $h^{CF}(f)$ as probability mass functions providing shares of the relevant worker and position populations rather than counts.

As discussed in section 2.2, Assumption 1', when combined with the stability conditions (30), implies that the probability that a given position k will be filled by a particular worker i is given by the logit form (15). When combined with Assumptions 1 and 2 (also cited by Assumption 1'), this implies that the group-level conditional choice probability $P(g|f)$ takes the form (29) for any position types f that are composed of positions k (as derived in section A1).

However, the statement of Proposition A3 makes clear that the form (29) also holds for the last type F, which contains the "dummy" unemployment positions whose "choices" will be workers becoming unemployed. The stability conditions (30) do not provide any justification for why these dummy nonemployment positions should be filled via the same logit form as the other position types that consist of actual positions at firms. Thus, these dummy positions, and the assumption that the probability distribution over alternative groups representing different worker and job match characteristics $(l(g), z(g))$ follows the logit form, are mere computational devices to calculate the equilibrium assignment. That this computational device in fact yields the unique stable assignment

[46] This implicitly requires that the unobserved draws ϵ_{0k} for position vacancy values are taken from a bounded distribution rather than the Type 1 extreme value distribution.

for the counterfactual labor market is the primary reason Proposition A2 requires a proof.

However, the stability conditions and Assumption 1' imply that the probability that a given worker i will choose a particular position k (where $k = 0$ represents nonemployment) is also given by the logit form (Decker et al. (2013)):

$$P^{CF}(k|i) = \frac{e^{\frac{\theta_g^{CF} - q_k^{CF}}{\sigma}}}{\sum_{k' \in \mathcal{K} \cup 0} e^{\frac{\theta_{g'}^{CF} - q_{k'}^{CF}}{\sigma}}} \tag{31}$$

This can then be aggregated (using the same steps as in section A1) to provide an expression for the probability that a randomly chosen worker from a given worker type l matches with a position that yields a transition assigned to group g:

$$P^{CF}(g|l) = \frac{1}{|l|} \sum_{i \in l} \frac{(e^{\frac{\theta_g^{CF}}{\sigma}})(\sum_{k:g(i,j(i),k)=g} e^{\frac{-q_k^{CF}}{\sigma}})}{\sum_{k' \in \mathcal{K} \cup 0} e^{\frac{\theta_{g'}^{CF} - q_{k'}^{CF}}{\sigma}}} \tag{32}$$

Assumptions 3' and 4', which are analogues to Assumptions 1 and 2 in section A1, allow us to simply this expression to the following:

$$P^{CF}(g|l) = \frac{e^{\frac{\theta_g^{CF}}{\sigma}} \overline{S}_{g|l(g),d}^{CF} h^{CF}(f(g)) \tilde{C}_f^{CF}}{\sum_{f' \in \mathcal{F}} \sum_{g' \in (l,f')} e^{\frac{\theta_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l(g'),d}^{CF} h^{CF}(f') \tilde{C}_{f'}^{CF}} \quad \forall\, l \in [1, \ldots, L] \tag{33}$$

Assumption 3' states that the discounted profits of alternative positions k of the same position type f are roughly the same. This implies that the profit share that workers must provide to the position in a stable matching is approximately the same for their existing positions as for other positions in the same local area with the same industry and establishment size and establishment average pay categories, and can be summarized by a parameter C_f^{CF} that is defined at the position type level.

Taken literally (given the characteristics we use to define groups), Assumption 4' states that every worker of the same worker type starts the year in firms with the same number of destination positions, which clearly does not hold. More broadly, though, Assumptions 3' and 4' allow us to replace the term $\sum_{k:g(i,k)=g} e^{\frac{-q_k^{CF}}{\sigma}}$ that depends on the individual i with an expression $P^{CF}(g|l, f(g)) h^{CF}(f(g)) \tilde{C}_{f(g)}^{CF}$ that depends on only group and destination-type level terms. Essentially, we assume that ignoring within-worker type variation in the number of positions at which they would be stayers (due to different establishment sizes of initial job matches) when aggregating is not generating significant bias in the counterfactual assignment and incidence estimates.

Under Assumptions 1' through 4', the group-level stable matching must satisfy the following market clearing conditions, which specify that supply must equal demand for each position type f:

$$\sum_{l \in \mathcal{L}} n^{CF}(l) \Big(\sum_{g:f(g)=2} P^{CF*}(g|l, \tilde{\mathbf{C}}^{\mathbf{CF}}) = h^{CF}(2) \tag{34}$$

$$\vdots \tag{35}$$

$$\sum_{l \in \mathcal{L}} n^{CF}(l)(\sum_{g:f(g)=F} P^{CF^*}(g|l, \tilde{\mathbf{C}}^{\mathbf{CF}}) = h^{CF}(F) \tag{36}$$

where $\tilde{\mathbf{C}}^{\mathbf{CF}}$ represents the $F-1$ length vector = $[1, \tilde{C}_2^{CF}, \ldots, \tilde{C}_F^{CF}]$ and each conditional probability $P^{CF^*}(g|l, \tilde{\mathbf{C}}^{\mathbf{CF}})$ takes the form in (33).

Assumption 2' allows us to ignore the possibility that supply might exceed demand for some position types (implying some vacant positions). In this alternative position-side system of equations, the expressions for each conditional probability $P^{CF^*}(g|l)$ do in fact stem directly from the necessary stability conditions. And all of the feasibility conditions for a stable matching are incorporated into the zero-excess demand equations (since $P^{CF^*}(g|l)$ sum to 1 by construction, the assignment $P^{CF^*}(g)$ that satisfies this system necessarily sums to the worker-type PMF $n^{CF}(l)$). Thus, one can apply the proof by Decker et al. (2013) that there exists a unique group-level assignment that satisfies all of the group-level feasibility and stability conditions (and is thus consistent with a stable matching in the assignment game defined at the level of worker-position matches).

If one wished, one could directly compute the unique group-level counterfactual assignment $P^{CF^*}(g|l)$ by finding a $F-1$ length vector $\tilde{\mathbf{C}}^{\mathbf{CF}}$ that solved this system, and constructing the implied assignment by plugging this vector into the conditional probability expressions (33). However, when $F \gg L$, solving this system is considerably more computationally burdensome than solving the worker-side counterpart (28), which features $L-1$ equations. Thus, the remainder of this proof is devoted to showing that any assignment $P^{CF}(g)$ implied by a solution to (28) must equal the assignment $P^{CF^*}(g)$ implied by a solution to (36). And since we know that the latter solution represents the unique group-level matching consistent with stability in the assignment game, the former solution must also be unique, and must also represent the group-level matching consistent with stability in the assignment game. Essentially, this amounts to showing that the device of adding "dummy" nonemployment positions present in (28) appropriately incorporates the surpluses π_{i0} that workers obtain from staying single.

Consider an L length vector $\mathbf{C}^{\mathbf{CF}} = [1, C_2^{CF}, \ldots, C_L^{CF}]$ that solves (28) and yields assignment $P^{CF}(g)$. We will show that one can use $\mathbf{C}^{\mathbf{CF}}$ to construct an alternative F length vector $\tilde{\mathbf{C}}^{\mathbf{CF}} = [1, \tilde{C}_2^{CF}, \ldots, \tilde{C}_F^{CF}]$ that solves (36), and that the assignment it generates, $P^{CF^*}(g)$, equals $P^{CF}(g)$.

We propose the following vector $\tilde{\mathbf{C}}^{\mathbf{CF}}$:

$$\tilde{C}_f^{CF} = \frac{\sum_{l=1}^{L} \sum_{g':(l(g'),f(g'))=(l,F)} e^{\frac{\theta_{g'}}{\sigma}} n^{CF}(l) \overline{S}_{g'|l,F} C_l^{CF}}{\sum_{l=1}^{L} \sum_{g':(l(g'),f(g'))=(l,f)} e^{\frac{\theta_{g'}}{\sigma}} n^{CF}(l) \overline{S}_{g'|l,f} C_l^{CF}} \; \forall \; f \in [1, \ldots, F] \tag{37}$$

Here, the numerator captures the inclusive value (as defined by Menzel (2015)) associated with the nonemployment position type F, while the denominator captures the inclusive value for the chosen position type f. This implies that $\tilde{C}_F^{CF} = 1$. While any position type could be chosen as the one whose mean exponentiated profit value is normalized, normalizing the nonemployment type is

particularly appealing, since it implies "profit" values of 0 for the dummy nonemployment position type F ($\tilde{C}_F^{CF} = e^{\overline{q}_F} = e^0 = 1$).

Let λ represent the inclusive value of the unemployment position type F, the numerator in (37):

$$\lambda = \sum_{l=1}^{L} \sum_{g':(l(g'),f(g'))=(l,F)} e^{\frac{\theta_{g'}}{\sigma}} n^{CF}(l) \overline{S}_{g'|l,F}^{CF} C_l^{CF} \tag{38}$$

Note that λ is independent of position type. We begin by showing that the assignments implied by the vectors $[C_1^{CF}, \ldots, C_L^{CF}]$ and $[C_1^{CF}, \ldots, \tilde{C}_F^{CF}]$ are identical: $P^{CF}(g) = P^{CF^*}(g)$.

Since $\mathbf{C}^{CF}$ solves the worker-side system of excess demand equations (28), we know that

$$\begin{aligned}
&\sum_{f' \in \mathcal{F}} h^{CF}(f') \sum_{g' \in (l,f')} \frac{e^{\frac{\theta_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l,f'}^{CF} n^{CF}(l) C_l^{CF}}{\sum_{l'=1}^{L} \sum_{g':(l(g'),f(g'))=(l',f)} e^{\frac{\theta_{g'}}{\sigma}} n^{CF}(l') \overline{S}_{g'|l',f} C_{l'}^{CF}} = n^{CF}(l) \; \forall \, l \in [1, L] \\
&\Rightarrow \sum_{f' \in \mathcal{F}} \sum_{g' \in (l,f')} \frac{e^{\frac{\theta_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l,f'}^{CF} h^{CF}(f')}{\sum_{l=1}^{L} \sum_{g':(l(g'),f(g'))=(l,f)} e^{\frac{\theta_{g'}}{\sigma}} n^{CF}(l) \overline{S}_{g'|l,f} C_l^{CF}} = \frac{1}{C_l^{CF}} \; \forall \, l \in [1, L] \\
&\Rightarrow \sum_{f' \in \mathcal{F}} \sum_{g' \in (l,f')} \frac{e^{\frac{\theta_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l,f'}^{CF} h^{CF}(f')}{\frac{\lambda}{\tilde{C}_{f'}^{CF}}} = \frac{1}{C_l^{CF}} \; \forall \, l \in [1, L] \\
&\Rightarrow \sum_{f' \in \mathcal{F}} \sum_{g' \in (l,f')} e^{\frac{\theta_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l,f'}^{CF} h^{CF}(f') \tilde{C}_{f'}^{CF} = \frac{\lambda}{C_l^{CF}} \; \forall \, l \in [1, L]
\end{aligned} \tag{39}$$

We can now proceed:

$$\begin{aligned}
P^{CF^*}(g) &= n^{CF}(l) P^{CF^*}(g|l) = n^{CF}(l) \frac{e^{\frac{\theta_g^{CF}}{\sigma}} \overline{S}_{g|l,f}^{CF} h^{CF}(f) \tilde{C}_f^{CF}}{\sum_{f' \in \mathcal{F}} \sum_{g' \in (l,f')} e^{\frac{\theta_{g'}^{CF}}{\sigma}} \overline{S}_{g'|l,f'}^{CF} h^{CF}(f') \tilde{C}_{f'}^{CF}} \\
&= \frac{n^{CF}(l) e^{\frac{\theta_g^{CF}}{\sigma}} \overline{S}_{g|l,f}^{CF} h^{CF}(f) \tilde{C}_f^{CF} C_l^{CF}}{\lambda} \\
&= h^{CF}(f) \frac{e^{\frac{\theta_g^{CF}}{\sigma}} n^{CF}(l) \overline{S}_{g|l,f}^{CF} \lambda C_l^{CF}}{\lambda \sum_{l'=1}^{L} \sum_{g':(l(g'),f(g'))=(l',f)} e^{\frac{\theta_{g'}}{\sigma}} f^{CF}(l') \overline{S}_{g'|l',f} C_{l'}^{CF}} \\
&= h^{CF}(f) P^{CF}(g|f) = P^{CF}(g)
\end{aligned} \tag{40}$$

It remains to show that the chosen $\tilde{\mathbf{C}}^{\mathbf{CF}}$ vector (37) solves (36). Consider the left-hand side of the excess demand equation for an arbitrary position type f in the system (36). One can write:

$$\sum_{l=1}^{L} \sum_{g:(l(g),f(g))=(l,f)} n^{CF}(l) P^{CF^*}(g|l, \Theta^{CF}, \tilde{\mathbf{C}}^{CF})$$

$$
\begin{aligned}
&= \sum_{l=1}^{L} \sum_{g:(l(g),f(g))=(l,f)} h^{CF}(f) P^{CF}(g|f, \Theta^{CF}, \mathbf{C}^{CF}) \\
&= h^{CF}(f) \sum_{l=1}^{L} \sum_{g:(l(g),f(g))=(l,f)} P^{CF}(g|f, \Theta^{CF}, \mathbf{C}^{CF}) \\
&= h^{CF}(f) \sum_{g:f(g)=f} P^{CF}(g|f, \Theta^{CF}, \mathbf{C}^{CF}) \\
&= h^{CF}(f) \qquad (41)
\end{aligned}
$$

where the last line imposes that $P^{CF}(g|f)$ is a (conditional) probability distribution and thus sums to one. Since we have proved that the implied "demand" by workers for positions of an arbitrary position type equals the "supply" $h^{CF}(f)$, we have proved that $\tilde{C}^{CF}$ solves the system (36).

Notice that the expression for the proposed equilibrium mean ex post profit vector (37) has value beyond its use in proving proposition A1. Once the L-vector of mean ex post utilities $\{C_l^{CF}\}$ for each worker type have been computed, one can use (37) to directly calculate the mean ex post profit vector for each position type f without having to solve a system of $F-1$ equations. This is quite valuable when $F \gg L$, as it is in our application. Of course, the equivalent mapping can be inferred by symmetry for the opposite case where $L \gg F$:

$$
C_l^{CF} = \frac{\sum_{f=1}^{F} \sum_{g':(l(g'),f(g'))=(L,f)} e^{\frac{\theta_{g'}}{\sigma}} h^{CF}(f) \overline{S}_{g'|L,f} \tilde{C}_f^{CF}}{\sum_{f=1}^{F} \sum_{g':(l(g'),f(g'))=(l,f)} e^{\frac{\theta_{g'}}{\sigma}} h^{CF}(f) \overline{S}_{g'|l,f} \tilde{C}_f^{CF}} \; \forall \, l \in [1, \ldots, L] \qquad (42)
$$

In section 2.3 we showed that these vectors are sufficient to determine both the worker and position type-level incidence of any counterfactual shocks to the composition or spatial distribution of labor supply and/or labor demand. Thus, at least in cases where the proposed model is a reasonable approximation of the functioning of the labor market (and housing supply is sufficiently elastic and agglomeration effects and other product market spillovers are second order), a proper welfare analysis of such shocks only requires solving at most $\min\{L, F\}$ non-linear excess demand equations. Since an analytical Jacobian can be derived and fed as an input to non-linear equations solvers, relatively large scale assignment problems featuring thousands of types on one side of the market (and perhaps more on the opposite side) can be solved within a matter of minutes.

A4 Estimating the Value of σ

We attempt to estimate σ, the standard deviation of the unobserved match-level component ϵ_{ik}, by exploiting the evolution in the composition of U.S. worker and position types $n^y(l)$ and $h^y(f)$ across years y. Specifically, we first estimate the set of group-level surpluses $\{\theta_g^{2003}\}$ from the observed 2003-2004 matching. Then, holding these surplus values fixed, we combine $\{\theta_g^{2003}\}$ with $n^y(l)$ and $h^y(f)$ from each other year $y \in [2004, 2012]$ to generate counterfactual assignments and changes

in scaled mean (exponentiated) utility values $\{C_l^{CF}\}$ for each worker type. These counterfactuals predict how mean worker utilities by skill/location combination would have evolved given the observed compositional changes in labor supply and demand had the underlying surplus values $\{\theta_g\}$ been constant and equal to $\{\theta_g^{2003}\}$ throughout the period. Thus, if most of the changes in the moving costs, recruiting costs, tastes, and relative productivities that compose the joint surplus values $\{\theta_g\}$ are nearly orthogonal to these counterfactual utility predictions, the relationship between the predicted utility changes and the observed earnings changes pins down σ, which plays the role of scaling utility changes in dollar equivalents.

Specifically, recall that $C_l^{CF} \approx \frac{1}{|l|}\sum_{i:l(i)=l} e^{\frac{-r_i^{CF}}{\sigma}}$. Thus, if ex post utility r_i^{CF} does not vary too much across individuals within a worker type, so that Jensen's inequality is near equality and $\frac{1}{|l^y|}\sum_{i:l(i)=l} e^{\frac{-r_i^{CF,y}}{\sigma^y}} \approx e^{\frac{\overline{r}_l^{CF,y}}{\sigma^y}}$, then taking logs yields $ln(C_l^{CF,y}) \approx \frac{\overline{r}_l^{CF,y}}{\sigma^y}$.

Next, we form the corresponding changes in observed annual earnings for each worker type in each year, $\overline{Earn}_l^{y+1} - \overline{Earn}_l^{y}$.[47] We then run the following regression at the l-type level for each year $y \in [2004 - 2012]$:

$$\overline{Earn}_l^{y+1} - \overline{Earn}_l^{y} = \beta_0^y + \beta_1^y(ln(C_l^{CF,y+1}) - ln(C_l^{CF,y})) + \nu_l^y \tag{43}$$

Recall that the $\overline{r}_l^{CF,y}$ values represent predicted money metric utility gains, and are thus denominated in dollars. Thus, if $\overline{r}_l^{CF,y+1} - \overline{r}_l^{CF,y}$ approximately equals $\overline{Earn}_l^{y+1} - \overline{Earn}_l^{y}$, then $\beta_1^y \approx \sigma^y$. In addition to the approximations above, this requires two conditions to hold.

First, other determinants of realized money metric utility changes by worker type, namely changes in joint surplus values, must be roughly orthogonal to the predicted money metric utility changes, $ln(C_l^{CF,y+1}) - ln(C_l^{CF,y}$, that are based purely on shifts in supply and demand composition with fixed surplus values. This might be violated, for example, if the kinds of workers who benefit from a favorable change in demand composition also experience increased productivity (which could possibly cause some of the change in demand composition), which would cause an overestimate of σ.

Second, realized earnings changes must move roughly one-for-one with realized money metric utility changes. This might be violated if its workers whose equilibrium utility increased systematically moved to jobs featuring better or worse amenities, avoided more moving/recruiting training costs, or moved to jobs featuring better or worse continuation values, so that their earnings gain over- or understated their utility gain. For example, firms might respond to an decrease in supply of a particular worker type by improving the amenities they offer rather than paying more, so that the earnings change understates the true utility change, causing an underestimate of σ.

Given the strong assumptions required, this estimator of σ is unlikely to be fully unbiased. However, this approach should generate a reasonable calibration of σ as long as most of the covariance between realized earnings changes and counterfactual utility changes is attributable to supply and

[47] Note that while worker earnings in initial job matches were used to assign workers to skill categories, to this point we have not used observed worker earnings in destination positions to identify any other parameters.

demand shifts rather than systematic correlations between the counterfactual utility changes and changes in joint surplus components.

Further efforts could conceivably be taken to exclude worker types l' whose surplus values $\{\theta_g : l(g) = l'\}$ were known to be changing over the chosen time period, or to allow θ_g to evolve in a particular parametric fashion. In fact, GS discuss how a vector of σ values associated with different types or combinations of types based on observed characteristics might potentially be jointly estimated with other model parameters (thereby allowing heteroskedasticity across types in the idiosyncratic match component). Since the focus of this paper is primarily on examining relative incidence across different worker types from shocks featuring different changes in labor demand composition, we opted for the simpler, more transparent approach.

As noted in Section 4.1, the worker type space depends on which location is considered the target location for the shock, with the geographic units that partially define worker types becoming more aggregated farther from the shock. To address this issue, in practice we constructed separate true and counterfactual earnings changes and estimated equation (43) for the collapsed worker type spaces associated with each possible target PUMA among the sample states, and averaged the estimates of β_1 across all regressions satisfying a minimum R^2 threshold of .1 to obtain $\hat{\beta}_1^y$.[48] The estimates of $\hat{\beta}_1^y$ are fairly consistent across years, so we use the mean estimate across all years, $\overline{\sigma} = 18,420$, to produce dollar values for all the results relating to utility gains presented in the paper.

A5 Using Transfers to Decompose the Joint Surpluses $\{\theta_g\}$

This appendix examines whether observing equilibrium transfers, denoted w_{ik}, allows the identification of additional parameters of interest. In CS's assignment model, the unobserved match-level heterogeneity is assumed to take the form $\epsilon_{ik} = \epsilon^1_{l(i)k} + \epsilon^2_{if(k)}$, so that aggregate surplus is left unchanged when two pairs of job matches (i,k) and (i',k') belonging to the same group g swap partners. The elimination of any true (i,k) match-level surplus component implies that equilibrium transfers cannot vary among job matches belong to the same group g, so that $w_{ik} = w_{g(i,k)}\ \forall\,(i,k)$.[49] GS show that under this assumption, observing the (common) group-level transfers w_g would be sufficient to decompose the group-level mean joint surplus θ_g into the worker and position's respective pre-transfer payoffs, which we denote θ^l_g and θ^f_g, respectively.

Because the model proposed in section 2 does not impose the additive separability assumption $\epsilon_{ik} = \epsilon^1_{l(i)k} + \epsilon^2_{if(k)}$, equilibrium transfers will in general vary among (i,k) pairs within the same group g. Given the substantial earnings variance within observed groups g regardless of the worker,

[48] A few PUMAs and states experienced relatively little year-to-year change in the distribution of employment across position types, so that the counterfactual earnings forecasts predicted true earnings changes poorly. In this case, the R^2 from the regression was very low, and β_1^y was badly identified. The results become far more stable across the remaining alternative type spaces when a minimum R^2 was imposed to eliminate the few badly identified estimates, which tended to produce outliers.

[49] If $w_{ik} > w_{i'k'}$ for any two matched pairs (i,k) and $(i'k')$ such that $g(i,k) = g(i',k')$, then (i',k) would form a blocking pair by proposing a surplus split between them featuring a transfer between w_{ik} and $w_{i'k'}$, thus undermining the stability of the proposed matching.

position, and job match characteristics used to define g, the CS restriction on the nature of unobserved match-level heterogeneity would be strongly rejected in the labor market context.

However, one can still consider whether the observed transfers $\{w_{ik}\}$ identify additional objects. From section 2.1, equilibrium transfers are related to equilibrium worker and position payoffs via:

$$w_{ik} = \pi^f_{ik} - q_k \tag{44}$$

$$w_{ik} = r_i - \pi^l_{ik} \tag{45}$$

Next, recall from equation (6) that under Assumptions 1 and 2 in Proposition A1 the log odds that a randomly chosen position from arbitrary position type f will choose a worker whose hire would be assigned to group g_1 relative to g_2 are given by:

$$\ln(\frac{P(g_1|d)}{P(g_2|f)}) = \ln(P(g_1|f)) - \ln(P(g_2|f)) = \frac{\theta_{g_1}}{\sigma} + \ln(\overline{S}_{g_1|l(g_1),f}) + \ln(n(l(g_1))) + \ln(C_{l(g_1)}) - \frac{\theta_{g_2}}{\sigma} - \ln \overline{S}_{g_2|l(g_2),f}) - \ln(n(l(g_2))) - \ln(C_{l(g_2)}) \tag{46}$$

Since $\ln(\overline{S}_{g_1|l(g_1),f})$, $\ln(\overline{S}_{g_2|l(g_2),f})$, $\ln(n(l(g_1)))$, and $\ln(n(l(g_2)))$ are all observed (or, if a large sample is taken, extremely precisely estimated), one can form adjusted log odds:

$$\ln(\frac{\hat{P}_{g_1|f}/(\overline{S}_{g_1|l(g_1),f})n(l(g_1)))}{\hat{P}_{g_2|f}/(\overline{S}_{g_2|l(g_2),f}n(l(g_2)))}) = (\frac{\theta_{g_1} - \theta_{g_2}}{\sigma}) + (\ln(C_{l(g_1)}) - \ln(C_{l(g_2)})) \tag{47}$$

Under Assumption 1, C_l is the mean of exponentiated (and rescaled) equilibrium utility payoffs owed to workers $i : l(i) = l$:

$$C_l = \frac{1}{|l|} \sum_{i:l(i)=l(g)} e^{-\frac{r_i}{\sigma}} \approx \sum_{\frac{1}{g_k}} \sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}} \; \forall \; k \tag{48}$$

Plugging (45) into (48) and then (48) into (47) yields:

$$\ln(\frac{\hat{P}_{g_1|f}/(\overline{S}_{g_1|l(g_1),f}n(l(g_1)))}{\hat{P}_{g_2|f}/(\overline{S}_{g_2|l(g_2),f}n(l(g_2)))}) = (\frac{\theta_{g_1} - \theta_{g_2}}{\sigma}) + (\ln(\frac{1}{|l|} \sum_{i:l(i,j(i))=l(g_1)} e^{-\frac{w_{ik}+\pi^l_{ik}}{\sigma}}) - \ln(\frac{1}{|l|} \sum_{i:l(i,j(i))=l(g_2)} e^{-\frac{w_{ik}+\pi^l_{ik}}{\sigma}})) \tag{49}$$

It is not immediately obvious how to use equation (49) to recover parameters of interest. Only when one adds further assumptions that are at odds with the structure of the model can one recover an expression that mirrors the one in CS. Specifically, suppose the following assumptions hold:

$$r_i \approx r_{l(i)} \; \forall \; i : l(i) = l \; \forall l \in \mathcal{L}$$

$$\pi^l_{ik} = \pi^l_{g(i,k)} \equiv \theta^l_g \; \forall \; (i,k) : g(i,k) = g \; \forall \; g \in \mathcal{G}$$

$$w_{ik} = w_{g(i,k)} \forall\ (i,k) : g(i,k) = g\ \forall\ g \in \mathcal{G} \tag{50}$$

These assumptions are extremely unlikely to hold in any stable matching if there is meaningful variance in ϵ_{ik} among the (i,k) pairs within the same group g. Nonetheless, they yield:

$$\begin{aligned}
&\ln\left(\frac{\hat{P}_{g_1|f}/(\overline{S}_{g_1|l(g_1),f}n(l(g_1)))}{\hat{P}_{g_2|f}/(\overline{S}_{g_2|l(g_2),f}n(l(g_2)))}\right) = \left(\frac{\theta_{g_1}-\theta_{g_2}}{\sigma}\right) + \left(\ln(e^{-r_{l(g_1)}}) - \ln(e^{-r_{l(g_2)}})\right) \\
&= \left(\frac{\theta_{g_1}-\theta_{g_2}}{\sigma}\right) + \frac{-r_{l(g_1)}+r_{l(g_2)}}{\sigma} = \left(\frac{\theta_{g_1}-\theta_{g_2}}{\sigma}\right) + \left(\frac{-(w_{g_1}+\theta^l_{g_1})+(w_{g_2}+\theta^l_{g_2})}{\sigma}\right) \\
&= \frac{\theta^f_{g_1}-\theta^f_{g_2}+(w_{g_2}-w_{g_1})}{\sigma}
\end{aligned} \tag{51}$$

Given an estimate of σ based on multiple markets (as described in Appendix A4) and data on mean annual earnings for each match group $g \in \mathcal{G}$, one could identify the difference in the position component of the joint surplus for arbitrary groups g_1 and g_2. This provides information about the relative profit contributions of different types of workers for each type of position before such workers salaries are considered. Note that one could still not separate the training cost, recruiting cost, current revenue contribution, and continuation value components of θ^f_g without additional data.

A similar progression using adjusted log odds based on the worker side conditional probabilities $P(g_1|l_1)$ and $P(g_2|l_1)$ would yield an estimate of the corresponding difference in the worker components of the joint surplus $\theta^l_{g_1} - \theta^l_{g_2}$ for any two groups featuring the same worker type. Since one such group could represent nonemployment, this approach would provide estimates of the desirability of working at various types of firms in various locations for zero pay relative to nonemployment. These values identify the reservation salary necessary to convince each worker type to take (or continue) a position of each type. Again, one could not disentangle the moving cost, search cost, non-wage amenity value, and continuation value components of the surplus without further data.

Because 1) we deem the assumptions (50) to be antithetical to the spirit of the model and at odds with the data, and 2) other than estimating σ, the use of transfers is not necessary to fulfill the primary aim of the paper, evaluating the utility and employment incidence across worker types of alternative local labor demand shocks, we do not make further use of the observed annual earnings distributions in the destination period.

A6 Imputing Missing Transitions Involving Unemployment and Missing Match Group Characteristics

Recall that nonemployed workers are only included in the sample in a given year if they are observed resuming work in a future year. This requirement is imposed so as to better distinguish unemployed workers from those exiting the labor force, but it creates the likely possibility of undercounting employment-to-unemployment (E-to-U) and unemployment-to-unemployment (U-to-U) transitions toward the end of the sample, when high shares of unemployment spells are right-censored due

to data availability.[50] In addition, the inability to observe the characteristics of those working in states that did not approve the use of their LEHD data creates a further need for imputation for employment-to-employment (E-to-E) and unemployment-to-employment (U-to-E) transitions originating in out-of-sample states. This appendix describes how data from the harmonized American Community Survey (hereafter ACS) series created by IPUMS along with official unemployment statistics from the Bureau of Labor Statistics (hereafter BLS) were used to address these problems.

A6.1 E-to-U and U-to-U Transitions

Note first that a match count must be generated for each match group g classified as an E-to-U transition, which consists of a combination of origin location, age category, and earnings quartile (since there is a single unemployment position type). Because the 1% ACS sample is too small to generate accurate E-to-U counts at even the PUMA level, we construct population-weighted E-to-U ACS counts by initial state, age, and earnings group, and re-scale each count so that the aggregate stock of E-to-U transitions matches the count of workers unemployed between more than 26 and less than 52 weeks from the BLS for the chosen year.[51] For in-sample initial states, we impute a match group g for each implied individual from the rescaled ACS counts by combining the ACS characteristics with a draw from the observed conditional empirical distribution of origin tracts given origin state among E-to-U transitions in the LEHD. E-to-U counts of transitions from the ACS that originate out of sample are aggregated across states, leaving 12 groups corresponding to the combinations of the three age categories and four initial earnings quartiles.

For U-to-U transitions, we begin with age-specific counts of long-term unemployment ($>$ 52 weeks) from the BLS, and distribute them across origin and destination states according to the joint distribution of state pairs among ACS U-to-U counts. We then impute an origin tract for each U-to-U transition from an in-sample state by drawing from the conditional empirical distribution of origin tracts given origin state and age group among the combined pool of E-to-U and E-to-E LEHD transitions that end in the observed state, so as to ensure appropriate support among origin tracts.

A6.2 E-to-E and U-to-E Transitions

Note that full match group counts are observed for all E-to-E transitions among in-sample states. Since we aggregate out-of-sample destination positions to a single type, in-sample to out-of-sample E-to-E match counts are also fully observed (by combining the absence of observed earnings with the provided indicator for non-zero earnings somewhere in the U.S.). E-to-E match counts among out-of-sample states require an initial earnings quartile and age to be assigned. We draw this using the distribution of initial earnings quartiles$\times$ age combinations among LEHD in-sample observations. E-to-E match counts from out-of-sample to in-sample states are completed similarly, except

[50] Since nearly all states enter the sample well before the years used for this analysis, the analogous risk of undercounting unemployment-to-employment transitions is negligible.

[51] Because we use a 50% random sample of LEHD transitions, we multiply estimated E-to-U counts (and U-to-U counts) by .5.

that the distribution of earnings quartiles also conditions on destination state, industry, firm size, and firm average pay as well as on being a state switcher.

U-to-E transitions in a fashion analogous to that of E-to-E transitions, except that an initial location must be imputed as well. If the worker has worked in-sample previously, we use the most recently observed employer tract as the worker's initial location. For those without previously observed employers (mostly young new entrants to the labor market), we use the same method for drawing origin tracts that was detailed for U-to-U transitions in the previous subsection.

A7 Smoothing Procedure

In this appendix we describe how we smooth the empirical distribution of job matches across groups, $\hat{P}(g)$, prior to estimation in order to generate accurate estimates of the set of identified joint surplus difference-in-differences Θ^{D-in-D}. We smooth for two reasons. First, such smoothing serves as a "noise infusion" technique that removes the risk that individual or establishment identities could be revealed by any estimates presented in the paper, as required of all research results generated from confidential microdata in Federal Statistical Research Data Centers (FSRDCs). Second, smoothing is necessary because there are sufficiently few observations per match group such that many match groups are rarely (or never) observed in a given matching despite substantial underlying matching surpluses simply due to sampling error. Essentially, $\hat{P}(g)$ is only a consistent estimator of $P(g)$ as the number of observed job matches per group I/G approaches infinity.

We overcome this sampling error problem by assuming that the underlying frequency $P(g)$ with which a job match belongs to a particular match group is a smooth function of the observed characteristics that define group g (following Hotz and Miller (1993) and Arcidiacono and Miller (2011)). This permits the use of a kernel density estimator that computes a weighted average of the empirical probabilities $\hat{P}(g')$ of "nearby" groups g' that feature "similar" vectors of characteristics to generate a well-behaved approximation of $P(g)$ from the noisy empirical distribution $\hat{P}(g)$.

Such smoothing introduces two additional challenges. First, excessive smoothing across other match groups erodes the signal contained in the data about the degree of heterogeneity in the relative surplus from job matches featuring different combinations of worker characteristics, establishment characteristics, and origin and destination locations. Since highlighting the role of such heterogeneity in forecasting the incidence of labor market shocks is a primary goal of the paper, decisions about the appropriate bandwidth must be made with considerable thought. The second, related challenge consists of identifying which of the worker and position characteristics that defines other groups makes them "similar", in the sense that the surplus $\{\theta_{g'}\}$ is likely to closely approximate the surplus θ_g whose estimate we wish to make more precise.

Recall that each group $g \equiv g(l, f, z)$ is a combination of 1) the origin establishment location (which we denote $loc(l)$), workers' initial age category (denoted $age(l)$), workers' initial earnings quartile (or unemployment status) at the origin establishment (denoted $earn(l)$), and an indicator for whether the worker's initial industry matches that of the job stimulus ($sameind(l)$); 2)

the destination establishment's location ($loc(f)$), size category ($f\_size(f)$), average earnings category ($f\_earn(f)$), and industry supersector ($ind(f)$); and 3) the trichotomous indicator $z(i,k)$ that equals '1' when establishment $j(i)$ and establishment k are the same ($z(i,k)=1$), '2' when these establishments are different but share an industry, and '0' when $j(i)$ and k are in different industries.

Given the goal of accurately characterizing incidence at a very high spatial resolution, we wish to preserve as accurately as possible any signal in the data about the structure of spatial ties between nearby local areas. Thus, wherever possible the kernel estimator should place non-zero weight only on alternative groups g' that share the same origin and destination locations ($loc(l(g)) = loc(l(g'))$ and $loc(f(g)) = loc(f(g'))$). Similarly, we posit that an establishment's combination of size, average pay, and industry is more important than its location in determining the initial earnings and age categories of the worker that generates the most surplus. Let $wchar(l(g)) \equiv [earn(l), age(l), sameind(l)]$ denote the non-location worker characteristics. To develop a smoothing approach that embodies these principles, we first exploit the fact that $P(g)$ can be written as $P(g|f) * h(f(g))$, and then decompose $P(g|f)$ via:

$$\begin{aligned} P(g|f) &= P([l(g), f(g), z(g)]|f) \\ &= P([loc(l(g)), wchar(l(g)), z(g)]|f) \\ &= P(loc(l(g))|wchar(l(g)), z(g), f) P([wchar(l(g)), z(g)]|f) \end{aligned} \tag{52}$$

where we use the definition of g, the set of characteristics that define $l(g)$ and $z(g)$, and the low of total probability. We use separate kernel density estimator procedures to estimate $P(loc(l(g))|wchar(l(g)), z(g), f(g))$ and $P(wchar(l(g)), z(g))|f(g))$.

Consider first the estimation of $P(loc(l(g))|wchar(l(g), z(g), f(g))$. For job stayer groups ($z(g)=1$), $P(loc(l(g))|wchar(l(g)), 1(z(g)=1), f) = 1(loc(l(g)) = loc(f(g)))$, since a potential stayer associated with a particular position type must have already been working at the same location in the origin period (since we treat establishments that switch locations as different establishments for computational reasons). Thus, no smoothing of this component is necessary for such groups. For groups with $z(g)=0$ or $z(g)=2$, this is the conditional probability that a particular new hire would be originally located at location $loc(l)$, given the hired worker's initial earnings, age, the position's type f, and whether the worker would be an industry stayer or switcher. Let $K^{dist}(g,g')$ denote the metric capturing the similarity of alternative groups g' and g for the purpose of estimating the propensity for establishments of type f to hire workers from a particular location (conditional on the other worker characteristics). As discussed above, wherever possible we only assign finite distance $K^{dist}(g,g') < \infty$ (i.e. non-zero weight) to empirical conditional probabilities $P(loc(l(g'))|wchar(l(g')), z(g'), f(g'))$ of alternative groups g' that feature both the same origin location $loc(l(g')) = loc(l(g))$ *and* destination location $loc(f(g')) = loc(f(g))$.[52]

$K^{dist}(g,g')$ assigns the smallest distance to alternative groups g' that also feature the same

[52]There are a very small number of worker and position types that are never observed in any job match. By necessity, we put positive weight on groups featuring nearby origin or destination locations in such cases.

position type ($f(g') = f(g)$), so that g and g' only differ in the non-location characteristics of hired workers. The closer $wchar(l(g'))$ is to $wchar(l(g'))$ (based on mahalanobis distance for the naturally ordered earnings and age categories), the smaller is the assigned distance $K^{dist}(g, g')$, but the profile flattens so that all groups g' that differ from g' only due to $wchar(l(g'))$ contribute to the weighted average. $K^{dist}(g, g')$ assigns larger (but still finite) distance to groups g' whose position types also differ on establishment size, avg. pay, or industry dimensions. The more different the establishment composition of the group, the smaller is its weight, with the profile again flattening so that all groups g' featuring the same origin and destination locations receive non-zero weight. Thus, groups with less similar worker and establishment characteristics receive non-negligible weight only when there are too few observations from groups with more similar worker and establishment characteristics to form reliable estimates. The weight assigned to a particular alternative group g' also depends on the number of observed new hires made by $f(g')$ at a particular combination of non-location worker characteristics $wchar(l(g'))$, denoted $N^{dist}(g')$ below, since this determines the signal strength of the empirical CCP $P(loc(l(g')|wchar(l(g')), z(g), f(g'))$. Thus, we have:

$$P(loc(l(g))|wchar(l(g)), z(g), f(g)) \approx$$
$$\sum_{g'}\left(\frac{\phi(K^{dist}(g', g)N^{dist}(g'))}{\sum_{g''}\phi(K^{dist}(g'', g)N^{dist}(g''))}\hat{P}(loc(l(g'))|wchar(l(g')), z(g'), f(g'))\right) \tag{53}$$

where $\phi(*)$ is the density function of the t distribution with 5 degrees of freedom (used as the kernel density), and $\frac{\phi(K^{dist}(g',g)N^{dist}(g'))}{\sum_{g''}\phi(K^{dist}(g'',g)N^{dist}(g''))}$ is the weight given to nearby match group g'.[53]

Next, consider the estimation of $P(wchar(l(g)), z(g)|f)$ and the conditional probabilities that either a job stayer, industry stayer, or industry mover with particular non-location characteristics will be hired to fill a position of position type f. Let $K^{wchar}(g, g')$ represent the metric capturing how similar alternative groups g' are to g for the purpose of estimating the propensity for firms of type f to hire (or retain) workers with particular non-location characteristics.

$K^{wchar}(g, g')$ assigns infinite distance (i.e. zero weight) to groups g' featuring different combos of establishment size, average pay, industry, and match characteristic $z(g)$ than the target group g. $K^{wchar}(g, g')$ assigns small distances to the conditional probabilities for groups g' representing hiring new (retaining) workers with the same non-location characteristics $wchar(l(g)) = wchar(l(g'))$ among firms from the same position type $f(g) = f(g')$ who are hiring workers from nearby locations. The distance metric increases in the tract pathlength between $loc(l(g'))$ and $loc(l(g'))$, but flattens beyond a threshold distance, so that groups featuring all origin locations (but shared values of other characteristics) contribute to the estimate.

Larger (but finite) distance values for $K^{wchar}(g, g')$ are assigned to conditional probabilities from groups g' that feature different (but nearby) destination locations (so $f(g) \neq f(g')$ but has the

[53]The degrees-of-freedom choice is effectively a bandwidth choice, since a larger number of degrees of freedom generates less smoothing (smaller weight in the tails). 5 is used as the bandwidth for both this and the kernel densities presented below. The results are insensitive to moderate changes in bandwidth choice, but choosing a very large bandwidth results in very volatile simulation estimates across target tracts, highlighting the need for smoothing.

same combination of non-location position characteristics). Again, the distance metric increases in the pathlength between $loc(f(g))$ and $loc(f(g'))$, but eventually flattens at a large but non-infinite value. As before, the weight given to a group g' also depends on the precision of its corresponding number of total hires made by firms of the position type $f(g')$, which is proportional to $h(f(g'))$.

Again, the motivation here is that targeted worker characteristics and the retention/new hire decision (conditional on the utility bids required by workers in different locations) is likely to be driven more by an establishment's production process (proxied by size, mean pay, and industry) than by its location. Since there still may be spatially correlated unobserved heterogeneity in production processes conditional on the other establishment observables, we place greater weight on the worker composition/retention decisions of proximate firms. More distant firms receive non-negligible weight only when too few local observations exist to form reliable estimates. The estimator for $P(wchar(l(g))|f)$ can be expressed via:

$$P(wchar(l(g)), z(g))|f(g)) \approx$$
$$\sum_{g'}\left(\frac{\phi(K^{wchar}(g', g)h(f(g')))}{\sum_{g''}\phi(K^{wchar}(g'', g)h(f(g'')))}\hat{P}(wchar(l(g')), z(g'))|f(g')) \right. \tag{54}$$

Bringing the pieces together, this customized smoothing procedure has a number of desirable properties. First, by requiring the same origin and destination locations as a necessary condition for non-zero weight when estimating the propensity for particular position types to hire workers from each location, one can generate considerable precision in estimated CCPs without imposing assumptions about the spatial links between locations. Second, at the same time, one can still use information contained in the hiring and retention choices of more distant establishments to learn about the propensity for establishments of different sizes, pay levels, and industries to retain and hire workers at different skill levels and from unemployment. Third, the procedure places non-trivial weight on match groups featuring less similar worker and establishment characteristics only when there are too few observed hires/retentions made by establishments associated with groups featuring very similar characteristics to yield reliable estimates. Fourth, overall the estimated probabilities $P(g|f)$ place weight on many groups, so that no element of the resulting smoothed distribution contains identifying worker or establishment information, eliminating disclosure risk.

A8 Assessing the Duration of the Shock

The static assignment model we use to assess the distribution of welfare changes across worker types following local labor demand shocks compares the labor market equilibrium at baseline with a new one reached after the market has re-equilibrated following the shock. Since we use year-to-year changes in job allocations to identify the surplus parameters that govern the counterfactual simulations we consider, we are implicitly assuming that the market takes roughly a year to re-equilibrate following the small, 250 job stimuli we consider. One way to define the duration of the

re-equilibration period is the length of time it takes for all or nearly all of the vacancy chains created by the new 250 positions to end with hires from unemployment (as opposed to further employment-to-employment transitions, which replace one vacancy with another). In this appendix, we attempt to estimate the duration of the shock by using data on each supersector's mean vacancy durations and shares of new hires from each other supersector and from unemployment to calibrate simulations of vacancy chain durations.

Specifically, we first collect 2012 estimates of mean vacancy duration by supersector from the Federal Reserve Economic Data's DHI-DFH series, which is based on the methodology of (Davis et al., 2013).[54] We then take each supersector's shares of hires from unemployment, from other firms in the same supersector, and from other firms in other supersectors from Table 1. For each hiring supersector, we distribute its share of hires from other industries across origin supersectors using the supersector origin-destination matrix of employment-to-employment transitions from the LEHD's Quarterly Workforce Indicators series, aggregated across all quarters of 2012.

We assume that the composition of new hires in each supersector is not meaningfully affected by the concentration of many jobs created in one location (arguably reasonable given the small size of the shocks we consider), so that each vacancy creates an independent chain of E-to-E and U-to-E hires, all of which sample from the hiring distribution just described.

We then simulate 1,000 shocks for each target supersector consisting of 250 newly created jobs as follows.[55] In the first period ($t = 1$), each newly created job posts an associated vacancy, which remains open for the mean duration of the chosen supersector. At the end of the period, the vacancy is filled either by a new hire from unemployment or from each other supersector, based on comaring a random uniform draw with cutoffs chosen to match the supersector-specific hiring distribution described above. If the draw implies that the newly hired worker was previously unemployed, the vacancy chain ends. Otherwise, the new worker is poached from a firm in the appropriate supersector. In the second period, this origin supersector then posts a vacancy to replace the lost worker, which remains open for its mean vacancy duration, and then is filled by comparing a new random draw with cutoffs corresponding to its hiring distribution. This process continues until all of the vacancy chains created by the original shock have been filled with hires from unemployment. Adding up all of the vacancy durations along each vacancy chain provides the distribution of vacancy chain durations among all the created vacancies for a given shock, and averaging across 1,000 shocks provides the distribution of vacancy chain durations for a typical shock.

Table A1 displays the results of this exercise. Columns 1-6 of the first row displays various quantiles of the average distribution of vacancy chain durations across all supersectors, weighting each by employment share. Overall, 99% of chains end with a U-to-E hire within 351 days, and 99.5% end within 403 days, consistent with an expected time until re-equilibration of around one year. Columns 7-10 show that 90% of chains are completed within 6 months, 99% within a year,

[54] We multiply their estimates measured in business days by 7/5 to translate the values to total days.

[55] Given the assumed independence of each vacancy, this is equivalent to simulating 250,000 separate vacancy chains for each supersector.

99.9% within 18 months, and essentially all are completed within 2 years.

Furthermore, the remaining columns show that the share of vacancy chains completed within a year is extremely consistent across targeted supersector, ranging only from 98.4% (Information) to 99.6% (Leisure & Hospitality). This is despite the fact that mean vacancy duration varies widely across supersectors, ranging from 10 days for construction to 55 days for the information sector. The explanation dovetails with one of the main findings of the paper: job creation shocks become more generic in their firm composition as they ripple out via employment-to-employment transitions. Thus, the right tail of the vacancy chain duration distribution is primarily determined by the national average of vacancy chain durations and the national share of hires from unemployment. In contrast, lower percentiles of the distribution are quite sensitive to the shock's industry composition: the median vacancy chain duration varies from 30 days for construction to 107 days for information.

Overall, these simulations provide clear support for using year-to-year employment reallocations to identify the joint surplus parameters $\{\theta_g\}$ and a one year horizon for evaluating shock incidence.

A9 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 A13 and A14) 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 ex-

hibit 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 A15 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 A13 and A14 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 A16 mimics Table A15 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 A15, 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%.

A10 Model Validation

The simulations consider relatively large, locally focused labor demand shocks, but the estimated surplus parameters $\hat{\Theta}^{D-in-D}$ that underlie them are identified from millions of quotidian job transitions driven by small firm expansions/contractions and worker retirements and preference or skill changes over the life cycle that generate considerable offsetting churn in the U.S. labor market. Thus, one might reasonably wonder whether parameters governing ordinary worker flows are capable of capturing the response to sizable, locally focused positive or negative shocks. To address this concern, in this section we describe and present results from a model validation exercise in which surplus parameters estimated on pre-shock ordinary worker flows were used to forecast the reallocation of workers after actual local economic shocks observed in the LEHD sample.

Specifically, 421 shocks to employment in a census tract were identified in the LEHD sample that satisfied the following criteria: 1) the shock occurred in a sample state during the years 2003 - 2012; 2) exactly one establishment experienced an employment change of at least 100 workers (usually a closing or opening); 3) at least 100 more or 100 fewer positions were filled in the chosen census tract than the year before; 4) the change in the number of positions constituted at least 10% and at most 200% of the total number of filled positions in the chosen census tract in the prior year; 5) The chosen tract featured at least 200 positions in the year prior to the shock; 6) no other tract in the same PUMA experienced an offsetting shock more than 50% as large as the shock to the chosen tract; and 7) less than 50% of the change in number of positions filled in the year of the shock was offset by a shock to the same tract in the opposite direction the following year.

These criteria ensure that a sufficient number of states report data in both the shock year and the prior year to properly capture any worker reallocation, that the shock was likely to be demand-driven and big enough to represent a meaningful disruption to both the chosen tract and the surrounding area, and that the shock was sufficiently persistent to rule out the possibility of a spuroius "shock" due to a reporting error by a large firm in the unemployment insurance data.

To create a forecast of the worker reallocations that a given shock occurring in year y would engender, the full set of model parameters was estimated based on the nationwide sample of worker transitions between years $y-2$ and $y-1$, using the same procedures for smoothing and aggregating types featuring distant locations described in Section 4.1. A counterfactual allocation was then generated by holding fixed the estimated surplus parameters but imposing the marginal distributions of origin and position types from the pair of years capturing the shock, $f^{y-1}(l)$ and $h^{y}(f)$. Since the exact composition of the shock (as reflected in $h^{y}(f)$) is built into the forecast, the test of the model is the degree to which the particular flows of workers of different worker types to particular

destination position types that resulted from the shock can be predicted.

We assess the accuracy of the forecast using the index of dissimilarity, which measures the percentage of predicted job matches that must be reassigned to a different match group to perfectly match the distribution of actual job matches across groups. It sums the absolute differences in the share of all matches assigned to g in the forecast and the actual data across all match groups g and multiplies by one-half: $\sum_g \frac{1}{2}|\hat{P}(g) - P(g)|$. Since most shock-induced reallocation likely occurs among workers initially near the target tract, we evaluate forecast accuracy only among groups featuring workers who were working or most recently working in the PUMA of the target tract.

To help understand the sources of improvements and shortfalls in model fit, we also compute the index of dissimilarity between the true allocation and several alternative forecasts. The first is a standard parametric conditional logit specification, in which the probability that a random position of type f is filled by a worker whose match would be assigned to group g is given by $P^y(g|f) = \frac{e^{X_g^y \lambda}}{\sum_{g'} e^{X_{g'}^y \lambda}}$, where X_g^y includes a substantial set of regressors constructed for year y that capture the kinds of predictors of joint surplus that researchers often use, and λ is the corresponding vector of parameters estimated from the relationship between the previous year's data, $P^{y-1}(g|f)$ and X_g^{y-1}. The regressors include full sets of dummies for the following categorical variables: origin-destination distance bins using tract pathlength within PUMA, PUMA pathlength within state, and state pathlength between states, initial earnings quartile $\times$ supersector dummies, age category $\times$ supersector dummies, initial earnings $\times$ firm size quartile dummies, age category $\times$ firm size quartile dummies, initial earnings $\times$ firm average pay quartile dummies, and age category $\times$ firm average pay quartile dummies. The regressors also include indicators for whether the group g is associated with job stayers ($1(z(g) = 1)$) or industry stayers among firm movers ($1(z(g) = 2)$), the worker type frequency $n(l(g))$ interacted with the geographic category of the position type associated with g (tract, PUMA, or state), an interaction between $n(l(g))$ and an indicator for whether $f(g)$ represents the "nonemployment" position type, and dummies for whether the origin and position types associated with match group g share a PUMA and share a state.

The second alternative forecast simply imposes that the CCPs that existed between $y-2$ and $y-1$ also hold during the shock year, so that $P^y(g) = \hat{P}^{y-1}(g|f)h^y(f)$. The third forecast mimics the second, except that the smoothing procedure described in Section A7 is applied to the $y-2$ data prior to constructing $\hat{P}^{y-1}(g|f)$. Like much research on either worker job search or firm job filling, all these alternative forecasts ignore the problem's two-sided nature, and thus do not impose that the proposed allocation satisfies the marginal distribution of worker types, $n^{y-1}(l)$. The fourth forecast uses Choo and Siow (2006)'s version of the model, in which the idiosyncratic job match-level surplus component ϵ_{ik} is replaced by two terms capturing surplus interactions between worker and position type and worker type and position rather than between worker and position: $\epsilon^1_{if(k)} + \epsilon^2_{l(i),k}$. This comparison is useful for assessing the importance of assumptions about correlation structure among unobserved components in driving predictions about counterfactual assignments.

The final five alternative forecasts all consider simplified versions of the baseline model in which

we eliminate heterogeneity in surplus values among 1) non-location firm characteristics, 2) non-location worker characteristics, 3) non-location worker and firm characteristics, 4) industry stayers vs. movers among job switchers, and 5) job stayers vs. job movers, respectively. Comparisons of these forecasts with the baseline specification reveal which dimensions of heterogeneity are important for the accuracy of out-of-sample predictions at different levels of aggregation.

Table 8 contains the results of this exercise. All entries consist of averages across all 421 shocks considered. The two-sided matching model, with parameters estimated from the previous period, 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, predicting the exact joint distribution of origin tract and initial earnings and age categories among workers hired separately for positions defined by tract/size/avg. pay/industry combinations is quite a tall order. Comparing across columns, we see that the parametric logit, despite over 100 regressors, performs considerably worse: 45.8% of transitions starting in the relevant PUMA must be reallocated to a different match group to match the actual post-shock allocation. Holding fixed the full prior year CCP distribution (cols. 3 and 8) performs slightly worse than the two-sided estimator within the target PUMA (35.3% misallocated), while smoothing the CCPs does not help at this level of aggregation (35.6)%. The Choo-Siow model matches the baseline model by this metric, with 35.1% misallocated.

For many purposes, however, forecasting exactly the right origin and destination tracts of transitions may be less important than correctly assessing the degree to which the disruption dissipates farther from the shock. To this end, row 2 reports results in which groups are combined that feature the same worker and establishment characteristics as well as origin and destination locations that belong to the same distance bin (using 14 bins), so that the dissimilarity index is computed over a somewhat coarser set of match groups. Only 11.1% of matches within the target PUMA are now misallocated by the two-sided forecast, with the two CCP forecasts following suit (with smoothing now improving the forecast slightly), suggesting that a substantial share of "incorrect" predictions might nonetheless be sufficiently accurate for most purposes. The parametric logit, by contrast, still performs poorly (36.2%). Furthermore, row 3 shows that combining groups featuring the same distance bins and worker earnings and age categories but different establishment size, average pay, and industry categories reduces the index of dissimilarity to 2.3% for workers originating in the targeted PUMA. This is despite the fact that $P(g)$ still contains 1,500 match groups with only 155 restrictions imposed by $n(l)$ and $h(f)$. Furthermore, the two-sided model significantly outperforms the simpler smoothed and unsmoothed CCP models at this level of aggregation (3.7% and 3.8%, respectively), and slightly outperforms the Choo-Siow model (3.7%). This suggests that the two-sided matching model matches well the locations of job movers and stayers, but is slightly less effective at matching small differences in the establishment characteristics of the jobs to which workers move.

The disaggregated baseline model also generates much more accurate predictions than the five alternative versions from Table A18 that restrict surplus heterogeneity across worker types, firm types, or mover/stayer status. After aggregating to distance bins and across non-location firm characteristics, the baseline model (2.3%) dramatically outperforms the version of the model with no

heterogeneity in firm characteristics (13.0%), despite the fact that the ability to match destination firm characteristics is no longer being assessed. Removing heterogeneity in non-location worker instead of firm characteristics also reduces the goodness of fit (8.6%), while removing both sets causes a required reallocation share of 19.2%. Dropping the distinction between job stayers and movers is inconsequential at this level of aggregation, but causes extremely poor predictions (73.5%) for the full group space that tries to predict which types of workers make job transitions.

For other purposes, the primary goal of a forecast might be to properly predict the geographic and skill incidence of unemployment. To this end, row 4 computes the index of dissimilarity exclusively over the set of groups featuring workers entering or exiting unemployment, so that the exercise is to predict the location and initial earnings and age categories of those losing jobs and the firm composition of those finding jobs (separately by worker initial location). Using the full set of locations, the worker or firm types of only 3.3% of within-PUMA workers entering or exiting unemployment would need to be altered in order for the two-sided prediction to match the allocation that actually occurred. The two-sided estimator easily outperforms the CCP estimators (both estimators are around 9%), and slightly outperforms the Choo-Siow model within the target PUMA (4.2%) Aggregating locations into 14 distance bins (row 5) shows that the two-sided predictions only badly predicts origin and destination locations for 1.0% of unemployment entrants and exiters originating in the PUMA, suggesting that it predicts quite well the geographic and skill incidence of changes in unemployment following the shocks considered. Taken together, the model does quite a good job of predicting the reallocation of workers across job types and particularly across employment/unemployment status that follows major local labor market shocks.

Table A1: Characterizing the Expected Duration of 250 Job Stimulus Packages by Target Supersector

	Key Statistics of the Vacancy Chain Duration Distribution									
Worker	Selected Percentiles						Share Completed by Half-Year			
Category	50th	75th	90th	95th	99th	99.5th	6 months	12 months	18 months	24 months
All	68	112	179	230	351	403	0.8950	0.9906	0.9992	0.9999
Nat. Resources	41	76	140	191	309	362	0.9439	0.9953	0.9995	1
Construction	30	69	135	185	305	357	0.9479	0.9955	0.9996	1
Manufacturing	80	123	188	239	358	409	0.8906	0.9910	0.9992	0.9999
Wholesale/Retail Trade	50	92	152	203	325	376	0.9341	0.9942	0.9995	0.9999
Information	107	159	227	279	403	458	0.8196	0.9837	0.9985	0.9998
Finance/Real Estate	96	151	223	276	397	452	0.8253	0.9844	0.9986	0.9999
Prof. & Bus. Services	57	107	175	224	349	399	0.9126	0.9918	0.9993	0.999
Education/Health	97	146	220	271	392	445	0.8373	0.9851	0.9988	0.9999
Leisure & Hospitality	42	69	127	177	299	351	0.9527	0.9959	0.9997	1
Other Services	51	90	154	206	328	382	0.9325	0.9937	0.9994	1
Government	95	145	215	267	388	441	0.8431	0.9863	0.9988	0.9999

Notes: Columns 1-5 contain the 50th, 75th, 95th, 99th, and 99.5th percentiles of the distribution of simulated vacancy chain durations based on averaging 1,000 shocks featuring 250 new vacancies. A vacancy chain refers to a sequence of vacancies caused by employment-to-employment transitions that replace a vacancy at the destination employer with a vacancy at the origin employer, and is considered completed when a vacancy is eventually filled by an unemployment-to-employment transition. Columns 6-9 capture the expected share of vacancies that end within 6, 12, 18, and 24 months, respectively. Shocks initially consist of 250 new vacancies from a particular supersector, and transition across supersectors based on the matrix of employment-to-employment transition shares within and across supersectors from 2012 Quarterly Workforce Indicators data. Each vacancy along the chain is assumed to last for its supersector's mean vacancy duration, as compiled by the 2012 Federal Reserve Economic Data series using (Davis et al., 2013)'s methodology. Rows 2-12 present distributional statistics for job creation shocks associated with the labeled supersector, while the first row takes an employment weighted average of all the supersector-specific shocks.

Table A2: Summary Statistics from the Smoothed Sample 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
Unemployment	0.120	0.229	0.771					0.288	0.617	0.095	0.314	0.554	0.131
1st Earn. Q.	0.217			0.057	0.703	0.099	0.141	0.295	0.551	0.153	0.313	0.507	0.180
2nd Earn. Q.	0.221			0.032	0.790	0.082	0.097	0.279	0.558	0.163	0.302	0.510	0.188
3rd Earn. Q.	0.221			0.021	0.831	0.075	0.073	0.251	0.563	0.186	0.282	0.504	0.214
4th Earn. Q.	0.221			0.016	0.846	0.073	0.065	0.216	0.551	0.233	0.264	0.456	0.280
Age <30	0.308	0.028	0.181	0.040	0.529	0.091	0.130	0.267	0.581	0.152	0.294	0.521	0.185
Age 31-50	0.427	0.028	0.061	0.024	0.742	0.073	0.071	0.260	0.555	0.185	0.293	0.491	0.217
Age >50	0.264	0.026	0.041	0.018	0.821	0.049	0.045	0.265	0.556	0.179	0.292	0.497	0.211

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		0.683	0.082	0.094	0.290	0.545	0.165	0.301	0.507	0.192
FE Quartile 3	0.241		0.059		0.793	0.069	0.079	0.269	0.556	0.175	0.296	0.505	0.199
FE Quartile 4	0.240		0.045		0.803	0.072	0.081	0.222	0.558	0.221	0.288	0.448	0.264
FS $<$ Median	0.514		0.117		0.700	0.085	0.097	0.308	0.505	0.187	0.332	0.472	0.197
FS $>$ Median	0.486		0.077		0.780	0.067	0.076	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		0.131		0.693	0.076	0.101	0.386	0.391	0.224	0.192	0.561	0.248
Construction	0.049		0.113		0.690	0.091	0.106	0.242	0.535	0.223	0.247	0.531	0.222
Manufacturing	0.089		0.054		0.829	0.035	0.081	0.339	0.490	0.172	0.296	0.518	0.187
Wholesale/Retail	0.204		0.107		0.733	0.077	0.083	0.234	0.570	0.196	0.251	0.522	0.228
Information	0.023		0.068		0.752	0.062	0.118	0.226	0.585	0.190	0.320	0.434	0.246
Financial Activities	0.059		0.061		0.761	0.074	0.104	0.237	0.601	0.162	0.297	0.493	0.211
Prof. Bus. Services	0.143		0.119		0.661	0.091	0.129	0.228	0.584	0.189	0.281	0.478	0.242
Ed. & Health	0.224		0.069		0.796	0.078	0.057	0.308	0.537	0.155	0.344	0.487	0.169
Leis. & Hosp.	0.113		0.179		0.621	0.116	0.084	0.298	0.525	0.177	0.336	0.468	0.196
Oth. Serv.	0.031		0.122		0.722	0.038	0.118	0.301	0.531	0.168	0.353	0.458	0.190
Government	0.047		0.036		0.880	0.025	0.059	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 A3: Specifications for the Baseline Set of Counterfactual Labor Demand Shocks

Spec. No.	Number of Jobs	Firm Avg. Earn. Quartile	Firm Size Quartile	Industry Supersector	Shock Type
1	250	2	1	Information	Stimulus
2	250	2	4	Information	Stimulus
3	250	4	1	Information	Stimulus
4	250	4	4	Information	Stimulus
5	250	2	1	Manufacturing	Stimulus
6	250	2	4	Manufacturing	Stimulus
7	250	4	1	Manufacturing	Stimulus
8	250	4	4	Manufacturing	Stimulus
9	250	2	1	Trade/Trans./Utilities	Stimulus
10	250	2	4	Trade/Trans./Utilities	Stimulus
11	250	4	1	Trade/Trans./Utilities	Stimulus
12	250	4	4	Trade/Trans./Utilities	Stimulus
13	250	2	1	Prof. & Bus. Services	Stimulus
14	250	2	4	Prof. & Bus. Services	Stimulus
15	250	4	1	Prof. & Bus. Services	Stimulus
16	250	4	4	Prof. & Bus. Services	Stimulus
17	250	2	1	Education & Health	Stimulus
18	250	2	4	Education & Health	Stimulus
19	250	4	1	Education & Health	Stimulus
20	250	4	4	Education & Health	Stimulus
21	250	2	1	Leisure & Hospitality	Stimulus
22	250	2	4	Leisure & Hospitality	Stimulus
23	250	4	1	Leisure & Hospitality	Stimulus
24	250	4	4	Leisure & Hospitality	Stimulus
25	250	2	1	Government	Stimulus
26	250	2	4	Government	Stimulus
27	250	4	1	Government	Stimulus
28	250	4	4	Government	Stimulus
29	250	2	1	Other Services	Stimulus
30	250	2	4	Other Services	Stimulus
31	250	4	1	Other Services	Stimulus
32	250	4	4	Other Services	Stimulus

Table A4: Assessing the Impact of Stimulus Packages That Create 250 Jobs by Pathlength Distance from Focal Tract Across Several Outcomes (Averages Across All Stimulus Compositions)

Distance from Target Tract	Share of JtJ Dest.	Initial Locations	Prob. of Stim. Job	Share of Stim Jobs	Change in P(Employed)	Share of Emp. Gains	Avg. Welfare Change ($)	Share of Wel. Gains
Target Tract	0.032	2.0E-05	0.005	0.034	0.001	0.006	322	0.009
			(3.0E-05)	(1.6E-04)	(4.7E-06)	(2.5E-05)	(10)	(4.1E-05)
1 Tct Away	0.057	1.1E-04	0.002	0.051	3.2E-04	0.009	105	0.015
			(9.1E-06)	(1.9E-04)	(1.9E-06)	(3.2E-05)	(3)	(5.5E-05)
2 Tcts Away	0.061	2.4E-04	0.001	0.053	1.4E-04	0.012	51	0.019
			(3.0E-06)	(1.7E-04)	(5.8E-07)	(3.6E-05)	(1)	(5.8E-05)
3+ Tcts w/in PUMA	0.122	0.001	2.8E-04	0.123	7.0E-05	0.032	26	0.055
			(7.1E-07)	(2.5E-04)	(1.7E-07)	(6.5E-05)	(0.4)	(1.3E-04)
1 PUMA Away	0.082	0.001	1.6E-04	0.094	4.6E-05	0.028	17	0.051
			(6.4E-07)	(2.5E-04)	(1.4E-07)	(6.6E-05)	(0.3)	(1.4E-04)
2 PUMAs Away	0.137	0.004	8.1E-05	0.132	3.1E-05	0.051	11	0.092
			(2.3E-07)	(2.6E-04)	(7.0E-08)	(8.4E-05)	(0.2)	(1.8E-04)
3+ PUMAs w/in State	0.328	0.055	2.4E-05	0.302	1.6E-05	0.243	7	0.399
			(7.2E-08)	(0.001)	(3.4E-08)	(0.001)	(0.1)	(0.001)
1 State Away	0.028	0.053	1.3E-06	0.035	2.6E-06	0.072	0.8	0.099
			(5.1E-09)	(1.4E-04)	(3.4E-09)	(1.7E-04)	(0.0)	(2.5E-04)
2+ States Away	0.036	0.262	2.1E-07	0.028	1.0E-06	0.141	0.1	0.070
			(5.3E-10)	(7.1E-05)	(4.4E-10)	(1.3E-04)	(0.0)	(8.2E-05)
Out of Sample	0.117	0.622	4.7E-07	0.148	1.3E-06	0.407	0.1	0.191
			(7.3E-10)	(2.3E-04)	(6.5E-10)	(2.1E-04)	(0.0)	(2.6E-04)

Notes: The column labeled "Share of JtJ Dest." displays the observed share of all job-to-job transitions among 2012 and 2013 dominant jobs whose origin-destination distance fell into the distance bins given by the row labels. The column labeled "Initial Locations" captures the share of workers for whom the distance between their origin position and the targeted census tract fell into the chosen bin (averaged over 300 simulations featuring different target census tracts). The column labeled "Prob. of Stim. Job" indicates the probability that a randomly chosen worker in the row distance bin will receive one of the 250 new positions generated by the simulated stimulus package. The column labeled "Change in P(Employed)" indicates the change in the probability that a randomly chosen worker in the row distance bin will be employed in the destination year as a consequence of the simulated stimulus package. The column labeled "Avg. Welfare Change" indicates the change in job-related welfare (scaled to be equivalent to $ of 2023 annual earnings) that a randomly chosen worker in the distance bin indicated by the row label will experience as a consequence of the simulated stimulus package. The columns labeled "Share of Stim. Jobs", "Share of Emp. Gains" and "Share of Wel. Gains" indicate the share of all stimulus jobs and total employment and welfare gains, respectively, generated by the simulated stimulus package that accrue to workers in the distance bin indicated by the row label.

"Target Tract" indicates that the worker's origin establishment was in the tract receiving the stimulus package. 1/2/3+ Tct(s) Away" indicates that the origin establishment was one, two, or 3 or more tracts away (by tract pathlength) within the same PUMA. "1/2/3+ PUMAs Away" and "1/2+ States Away" indicate the PUMA pathlength (if within the same state) and state pathlength (if in different states), respectively. "Out of Sample" indicates that the worker's origin establishment was not among the 19 states providing data in the sample.

Standard errors are provided in parentheses, and are based on the sampling distribution among the sample of 300 target tracts simulated for each stimulus package specification.

Table A5: Assessing the Impact of Stimulus Packages That Create 250 Jobs by Distance in Miles from Focal Tract Across Several Outcomes (Averages Across All Stimulus Compositions)

Distance from Centroid of Target Tract	Share of JtoJ Dest.	Initial Locations	Prob. of Stim. Job	Share of Stim. Jobs	Change in P(Employed)	Share of Emp. Gains	Avg. Welfare Gain ($)	Share of Wel. Gains
Within 1 Mile	0.032	8.1E-05	0.003	0.040	4.6E-04	0.007	164	0.013
1-2 Miles Away	0.053	2.1E-04	0.001	0.031	1.9E-04	0.006	55	0.010
3-5 Miles Away	0.093	0.001	0.001	0.095	1.9E-04	0.022	69	0.037
6-11 Miles Away	0.120	0.002	0.001	0.113	2.1E-04	0.030	74	0.053
11-26 Miles Away	0.160	0.003	0.001	0.151	1.5E-04	0.046	54	0.081
26-50 Miles Away	0.070	0.002	2.1E-04	0.064	6.1E-05	0.024	24	0.043
51-100 Miles Away	0.063	0.002	9.0E-05	0.056	3.7E-05	0.027	14	0.051
101-250 Miles Away	0.202	0.026	1.2E-05	0.100	9.9E-06	0.094	4	0.168
>250 Miles Away	0.092	0.342	1.1E-06	0.202	1.9E-06	0.336	0.4	0.354
Out of Sample	0.117	0.622	4.7E-07	0.148	1.3E-06	0.407	0.1	0.191

Notes: See Table A4 for expanded definitions of the outcomes in the column labels. The row labels define sets of workers for whom the distance between the establishment associated with their origin dominant jobs and the census tract receiving the simulated stimulus package fell in the listed distance bin.

Table A6: Assessing the Value of Restricting Stimulus Jobs to Workers Within the Target PUMA: Spatial Employment and Welfare Incidence for Restricted and Unrestricted Stimulus Packages (Each Featuring 250 Positions at a Large High-Paying Manufacturing Firm)

Distance from Target Tract	Change in P(Employed)		Share of Emp. Gains		Avg. Welfare Change ($)		Share of Wel. Gains	
	Unres.	Res.	Unres.	Res.	Unres.	Res.	Unres.	Res.
Target Tract	0.001	0.005	0.004	0.028	296	2076	0.009	0.050
1 Tct Away	2.5E-04	0.001	0.008	0.039	98	474	0.015	0.067
2 Tcts Away	1.2E-04	4.9E-04	0.010	0.039	49	176	0.019	0.061
3+ Tcts w/in PUMA	6.1E-05	1.6E-04	0.028	0.069	27	81	0.056	0.114
1 PUMA Away	3.9E-05	3.7E-05	0.025	0.024	17	16	0.048	0.044
2 PUMAs Away	2.8E-05	2.6E-05	0.046	0.043	12	11	0.087	0.078
3+ PUMAs w/in State	1.6E-05	1.5E-05	0.236	0.215	7	6	0.385	0.334
1 State Away	2.5E-06	2.5E-06	0.070	0.068	0.8	0.8	0.092	0.086
2+ States Away	1.1E-06	9.4E-07	0.144	0.128	0.1	0.1	0.069	0.055
Out of Sample	1.3E-06	1.1E-06	0.429	0.346	0.2	0.1	0.220	0.113

Notes: See Table A4 for expanded definitions of the row labels and the outcomes in the column labels. Table entries consist of various measures of incidence by worker initial distance from the target census tract from a stimulus package consisting of 250 new jobs at large (employment above the worker-weighted median), high-paying (4th quartile of avg. worker pay) manufacturing firms. Columns labeled "Res." report results from specifications in which the new positions are constrained to be filled by workers initially working (or most recently working) in the same PUMA as the targeted tract, while columns labeled "Unres." report results from specifications in which the new positions may be filled by any worker in the nation.

Table A7: Shares of Nationwide Employment and Utility Gains Induced by Job Stimuli among Worker Initial Earnings, Age, and Industry Categories: Stimuli Consist of 250 Jobs at Firms in Different Firm Size/Firm Average Earnings Quartiles (Averaged across Different Firm Industries)

Worker	Share of Employment Gains					Share of Welfare Gains				
Category	Avg.	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi	Avg.	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi
Unemployment	0.438	0.450	0.447	0.431	0.423	0.120	0.131	0.134	0.109	0.105
1st Earn Q.	0.237	0.237	0.238	0.236	0.239	0.198	0.208	0.207	0.187	0.188
2nd Earn Q.	0.141	0.138	0.139	0.142	0.144	0.217	0.219	0.218	0.216	0.217
3rd Earn Q.	0.096	0.093	0.093	0.099	0.100	0.225	0.219	0.219	0.230	0.232
4th Earn Q.	0.088	0.083	0.083	0.093	0.094	0.241	0.223	0.223	0.258	0.258
Age ≤ 30	0.403	0.405	0.417	0.392	0.398	0.323	0.331	0.341	0.307	0.312
Age 31-50	0.398	0.396	0.389	0.405	0.402	0.425	0.417	0.412	0.436	0.434
Age ≥ 51	0.199	0.200	0.195	0.204	0.200	0.253	0.251	0.248	0.257	0.254
Diff. Ind.	0.934	0.931	0.935	0.933	0.935	0.885	0.878	0.891	0.881	0.892
Same Ind.	0.067	0.069	0.065	0.067	0.065	0.115	0.122	0.109	0.119	0.108

Notes: See Table 2 for expanded definitions of worker subpopulations captured by the row labels. See Table 3 for expanded definitions of the establishment size/avg. pay combinations captured by the column labels. The first five columns capture the share of employment gains in the destination year attributable to a 250 job stimulus package accruing to workers whose employment status or earnings in the origin year places them in the earnings/age/industry category listed by the row label. The last five columns capture the share of all stimulus-driven welfare gains (scaled to be equivalent to $ of 2023 annual earnings) accruing to workers in each earnings/age/industry category. Columns 1 and 6 average across all 32 stimulus package specifications. Each of columns 2-5 and 7-10 averages results across 8 stimulus packages featuring jobs with establishments in the same firm size quartile/firm average pay quartile combination but in different industry supersectors (as well as simulated 300 simulations for each stimulus package specification featuring different target census tracts)

Table A8: Cumulative Shares of Employment and Welfare Gains due to a Job Stimulus Accruing to Workers within Different Distances from Focal Tract: Stimuli Consist of 250 New Jobs at Firms in Alternative Industries (Averaged Across Firm Size and Average Earnings Combos)

Panel A: Cumulative Shares of Unemployment Gains

Distance from		Industry							
Focal Tract	Avg.	Info.	Manu.	Trd./Tns.	Prof. Bus.	Ed./Hlth	Lei/Hosp.	Gov.	Oth. Serv.
Focal Tract	0.006	0.005	0.006	0.005	0.005	0.007	0.006	0.006	0.006
1 Tct Away	0.015	0.013	0.015	0.014	0.013	0.018	0.016	0.015	0.016
2 Tcts Away	0.026	0.023	0.027	0.025	0.024	0.031	0.027	0.027	0.027
3+ Tcts w/in PUMA	0.058	0.053	0.059	0.054	0.055	0.066	0.059	0.061	0.059
1 PUMA Away	0.087	0.080	0.087	0.080	0.083	0.097	0.087	0.090	0.088
2 PUMAs Away	0.138	0.131	0.138	0.129	0.133	0.149	0.138	0.143	0.139
3+ PUMAs w/in State	0.380	0.372	0.380	0.372	0.372	0.394	0.377	0.394	0.380
1 State Away	0.452	0.446	0.450	0.444	0.444	0.465	0.449	0.466	0.451
2+ States Away	0.593	0.588	0.591	0.587	0.587	0.605	0.591	0.605	0.592
Out of Sample	1	1	1	1	1	1	1	1	1
Within 10 Miles	0.066	0.062	0.066	0.060	0.063	0.074	0.066	0.068	0.067
Within 250 Miles	0.257	0.250	0.256	0.249	0.252	0.271	0.255	0.266	0.258

Panel B: Cumulative Shares of Welfare Gains

Distance from		Industry							
Focal Tract	Avg.	Info.	Manu.	Trd./Tns.	Prof. Bus.	Ed./Hlth	Lei/Hosp.	Gov.	Oth. Serv.
Focal Tract	0.009	0.007	0.010	0.009	0.008	0.013	0.010	0.009	0.009
1 Tct Away	0.025	0.021	0.025	0.023	0.022	0.031	0.025	0.026	0.026
2 Tcts Away	0.044	0.039	0.045	0.041	0.039	0.053	0.044	0.046	0.045
3+ Tcts w/in PUMA	0.099	0.091	0.102	0.093	0.091	0.115	0.098	0.103	0.099
1 PUMA Away	0.150	0.140	0.152	0.142	0.142	0.170	0.149	0.155	0.149
2 PUMAs Away	0.242	0.233	0.243	0.232	0.232	0.266	0.241	0.249	0.240
3+ PUMAs w/in State	0.641	0.634	0.639	0.635	0.626	0.667	0.638	0.659	0.630
1 State Away	0.740	0.736	0.735	0.734	0.726	0.764	0.739	0.757	0.724
2+ States Away	0.809	0.807	0.804	0.803	0.797	0.831	0.810	0.826	0.797
Out of Sample	1	1	1	1	1	1	1	1	1
Within 10 Miles	0.112	0.106	0.113	0.105	0.108	0.128	0.113	0.115	0.113
Within 250 Miles	0.455	0.447	0.453	0.447	0.445	0.482	0.453	0.465	0.448

Notes: See Tables A4 and 1 for expanded definitions of the row and column labels. Each entry provides the share of net employment gains attributable to a 250 job stimulus package accruing to workers whose distance between their origin jobs and the census tract receiving the stimulus package is less than or within the distance bin indicated in the row label. Different columns consider average employment impacts from stimuli featuring jobs in different industry supersectors. Each column averages results across four stimulus packages with the same industry supersector but in different categories of the establishment size and average worker earnings.

Table A9: Share of Nationwide Employment and Utility Gains From New Stimulus Positions by Distance from Focal Tract: Stimuli Consist of 250 New Positions in Alternative Combinations of Firm Size Quartile/Firm Average Pay Quartile (Averaged Across Industry Supersectors)

Distance from	Share of Employment Gains				Share of Utility Gains			
Focal Tract	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi
Target Tract	0.006	0.006	0.005	0.005	0.010	0.009	0.010	0.009
1 Tct Away	0.017	0.016	0.014	0.013	0.026	0.025	0.025	0.024
2 Tcts Away	0.030	0.028	0.025	0.023	0.046	0.044	0.043	0.042
3+ Tcts w/in PUMA	0.065	0.061	0.055	0.052	0.103	0.100	0.097	0.096
1 PUMA Away	0.095	0.091	0.081	0.079	0.156	0.152	0.146	0.146
2 PUMAs Away	0.149	0.144	0.130	0.127	0.251	0.247	0.234	0.236
3+ PUMAs w/in State	0.396	0.394	0.365	0.365	0.658	0.660	0.620	0.626
1 State Away	0.468	0.467	0.436	0.436	0.760	0.762	0.716	0.721
2+ States Away	0.608	0.607	0.579	0.579	0.830	0.833	0.786	0.789
Out of Sample	1	1	1	1	1	1	1	1
Within 10 Miles	0.072	0.069	0.062	0.060	0.116	0.113	0.111	0.110
Within 250 Miles	0.272	0.268	0.245	0.243	0.471	0.468	0.439	0.442

Notes: See Table A4 for expanded definitions of the row labels. See Table 3 for expanded definitions of the establishment size/avg. pay combinations captured by the column labels. The first four columns capture the share of all net employment gains attributable to a 250 job stimulus package for workers whose distance between their origin jobs and the census tract receiving the stimulus package is below or within in the distance bin indicated in the row label. The last four columns capture the share of all stimulus-driven utility gains accruing to workers below or within in each distance bin. Different columns consider average employment impacts from stimuli featuring jobs with establishments from different combinations of firm size and firm average worker categories. Each column averages results across 8 stimulus packages featuring jobs with establishments in the same firm size quartile/firm average pay quartile combination but in different industry supersectors (as well as across 300 simulations for each stimulus package specification featuring different target census tracts).

Table A10: Cumulative Shares of Employment and Welfare Losses From a Plant Closing Removing 250 Positions at Large High-Paying Manufacturing Firms in the Target Tract Accruing to Workers in Different Distance Bins from the Target Tract by Worker Subpopulation

Panel A: Cumulative Share of Employment Losses

		Earnings Quartile				Age			Ind. Status	
Distance Bin	Unemp	1st Q.	2nd Q.	3rd Q.	4th Q.	$<= 30$	$31 - 50$	> 50	Diff Ind.	Same Ind.
Target Tract	0.005	0.022	0.091	0.217	0.291	0.051	0.098	0.121	0.003	0.612
1 Tct Away	0.010	0.028	0.098	0.224	0.299	0.057	0.105	0.127	0.009	0.623
2 Tcts Away	0.018	0.035	0.106	0.232	0.307	0.066	0.112	0.134	0.016	0.632
Over 3 Tcts within PUMA	0.039	0.056	0.129	0.253	0.323	0.087	0.133	0.153	0.037	0.653
1 PUMA Away	0.058	0.076	0.148	0.269	0.336	0.108	0.151	0.169	0.056	0.668
2 PUMAs Away	0.095	0.110	0.180	0.298	0.359	0.143	0.182	0.198	0.090	0.689
3+ PUMAs w/in State	0.346	0.291	0.346	0.449	0.489	0.354	0.370	0.372	0.301	0.769
1 State Away	0.383	0.324	0.378	0.476	0.509	0.389	0.401	0.401	0.337	0.782
2+ States Away	0.575	0.500	0.535	0.610	0.608	0.572	0.557	0.547	0.517	0.840
Out of Sample	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Within 10 miles	0.044	0.055	0.127	0.254	0.327	0.088	0.136	0.155	0.040	0.648
Within 250 miles	0.220	0.217	0.279	0.383	0.422	0.256	0.283	0.291	0.203	0.733

Panel B: Cumulative Share of Welfare Losses

		Earnings Quartile				Age			Ind. Status	
Distance Bin	Unemp	1st Q.	2nd Q.	3rd Q.	4th Q.	$<= 30$	$31 - 50$	> 50	Diff Ind.	Same Ind.
Target Tract	0.012	0.068	0.203	0.398	0.534	0.198	0.379	0.439	0.006	0.779
1 Tct Away	0.026	0.081	0.216	0.409	0.548	0.211	0.392	0.452	0.019	0.792
2 Tcts Away	0.046	0.098	0.232	0.422	0.560	0.227	0.406	0.465	0.036	0.802
Over 3 Tcts within PUMA	0.099	0.144	0.274	0.454	0.584	0.268	0.439	0.494	0.080	0.823
1 PUMA Away	0.146	0.189	0.311	0.482	0.603	0.304	0.466	0.519	0.121	0.837
2 PUMAs Away	0.237	0.265	0.370	0.526	0.636	0.368	0.513	0.562	0.194	0.857
3+ PUMAs w/in State	0.732	0.643	0.645	0.735	0.810	0.690	0.740	0.766	0.576	0.930
1 State Away	0.787	0.692	0.683	0.763	0.829	0.731	0.768	0.792	0.624	0.939
2+ States Away	0.869	0.784	0.759	0.821	0.872	0.808	0.825	0.846	0.716	0.963
Out of Sample	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Within 10 miles	0.106	0.144	0.271	0.454	0.593	0.271	0.443	0.496	0.088	0.820
Within 250 miles	0.504	0.489	0.540	0.649	0.722	0.548	0.640	0.680	0.409	0.899

Notes: See Table A4 for expanded definitions of the distance bins represented by the row labels. See Table 4 for expanded definitions of the worker subpopulations indicated by the column labels. Each entry gives the cumulative share of employment losses (Panel A) or utility losses (Panel B) among workers whose initial location is closer than or within the distance bin associated with the row label and who belong to the subpopulation associated with the column sublabel due to a simulated plant closing in which 250 positions are removed at large, high-paying manufacturing firms. The average is taken across 200 simulations featuring different target census tracts.

Table A11: Change in Probability of Destination Employment (or Unemployment) at Different Distances from Focal Tract after an Establishment Closing Removing 250 Positions at either Manufacturing or Retail Firms for Workers Initially Employed in the Focal Tract by Worker Subpopulation

Panel A: Large High-Paying Manufacturing

			Earnings Quartile				Age			Ind. Status	
Distance Bin	All	Unemp	1st Q.	2nd Q.	3rd Q.	4th Q.	$<= 30$	$31-50$	> 50	Diff Ind.	Same Ind.
Unemployment	0.007	0.001	0.003	0.006	0.009	0.008	0.006	0.007	0.007	2.3E-04	0.016
Target Tract	-0.046	-0.003	-0.012	-0.030	-0.050	-0.069	-0.041	-0.051	-0.040	-0.001	-0.112
1 Tct Away	0.002	2.0E-05	4.4E-04	0.002	0.002	0.002	0.002	0.002	0.001	-1.3E-05	0.004
2 Tcts Away	0.002	8.1E-05	0.001	0.002	0.003	0.002	0.002	0.002	0.002	-8.1E-06	0.005
3+ Tcts w/in PUMA	0.005	2.8E-04	0.002	0.003	0.006	0.007	0.004	0.005	0.004	3.3E-05	0.011
1 PUMA Away	0.003	2.9E-04	0.001	0.003	0.003	0.004	0.003	0.003	0.002	9.1E-05	0.006
2 PUMAs Away	0.003	3.2E-04	0.001	0.003	0.004	0.004	0.003	0.003	0.003	1.0E-04	0.009
3+ PUMAs w/in State	0.010	0.001	0.003	0.006	0.011	0.016	0.009	0.012	0.009	2.3E-04	0.028
1 State Away	0.002	7.6E-05	2.1E-04	0.001	0.002	0.004	0.002	0.002	0.001	3.7E-05	0.004
2+ States Away	0.003	1.6E-04	2.8E-04	0.001	0.002	0.006	0.003	0.004	0.002	4.7E-05	0.007
Out of Sample	0.011	1.2E-04	0.002	0.005	0.010	0.018	0.007	0.013	0.010	1.0E-04	0.024

Panel B Large Low-Paying Retail

			Earnings Quartile				Age			Ind. Status	
Distance Bin	All	Unemp	1st Q.	2nd Q.	3rd Q.	4th Q.	$<= 30$	$31-50$	> 50	Diff Ind.	Same Ind.
Unemployment	0.007	0.001	0.011	0.008	0.005	0.004	0.008	0.006	0.005	1.8E-04	0.018
Target Tract	-0.036	-0.002	-0.053	-0.045	-0.034	-0.027	-0.046	-0.033	-0.025	-0.001	-0.098
1 Tct Away	0.001	4.9E-05	0.001	0.001	0.001	0.001	0.001	0.001	0.001	5.0E-06	0.003
2 Tcts Away	0.001	4.6E-05	0.002	0.001	0.001	0.001	0.001	0.001	0.001	2.8E-06	0.003
Over 3 Tcts within PUMA	0.003	1.7E-04	0.004	0.004	0.003	0.002	0.004	0.003	0.002	3.4E-05	0.008
1 PUMA Away	0.002	1.3E-04	0.003	0.003	0.002	0.002	0.003	0.002	0.001	4.6E-05	0.006
2 PUMAs Away	0.005	2.6E-04	0.007	0.006	0.004	0.003	0.006	0.004	0.003	8.7E-05	0.012
3+ PUMAs w/in State	0.013	0.001	0.018	0.016	0.013	0.011	0.017	0.011	0.008	2.1E-04	0.034
1 State Away	0.001	9.1E-05	0.001	0.001	0.001	0.001	0.001	0.001	0.001	3.0E-05	0.003
2+ States Away	0.001	1.2E-04	0.002	0.001	0.001	0.001	0.001	0.001	0.001	3.7E-05	0.003
Out of Sample	0.003	6.1E-05	0.005	0.005	0.003	0.003	0.004	0.003	0.003	5.7E-05	0.009

Notes: See Table A4 for expanded definitions of the distance bins represented by the row labels. See Table 4 for expanded definitions of the worker subpopulations indicated by the column sublabels. Each entry gives the change in the probability of employment at a location whose distance falls into the distance bin associated with the row label among workers initially belonging to the subpopulation associated with the column sublabel and working in the previous year (or most recently working) in the focal census tract. The changes in employment probability are due to a simulated plant closing in which 250 positions are removed at either large, high-paying manufacturing firms (Panel A) or large, low-paying wholesale/retail firms (Panel B). Each entry represents an average over 200 simulations featuring different target census tracts. The entries in the row labeled "Unemployment" provides the change in the share of workers who stay or become unemployed due to the plant closing.

Table A12: Comparing the Impact of Plant Closings and Openings at Different Scales and Distances from Focal Tract on Employment and Welfare Outcomes

Panel A: Employment Outcomes

	Change in P(Employed)						Share of Employment Gains or Losses					
Distance from	Plant Opening			Plant Closing			Plant Opening			Plant Closing		
Focal Tract	125	250	500	125	250	500	125	250	500	125	250	500
Target Tract	1.2E-04	2.3E-04	4.3E-04	-.003	-.007	-.016	0.005	0.005	0.004	0.075	0.086	0.099
1 Tct Away	8.7E-05	1.7E-04	3.2E-04	-6.6E-05	-1.2E-04	-2.2E-04	0.013	0.012	0.012	0.082	0.092	0.105
2 Tcts Away	5.5E-05	1.1E-04	2.0E-04	-4.4E-05	-8.0E-05	-1.4E-04	0.023	0.022	0.021	0.090	0.100	0.113
3+ Tcts w/in PUMA	2.8E-05	5.6E-05	1.1E-04	-2.4E-05	-4.6E-05	-8.6E-05	0.047	0.047	0.046	0.111	0.120	0.132
1 PUMA Away	1.8E-05	3.6E-05	7.2E-05	-1.6E-05	-3.0E-05	-5.6E-05	0.069	0.068	0.067	0.130	0.138	0.149
2 PUMAs Away	1.2E-05	2.5E-05	5.0E-05	-1.1E-05	-2.1E-05	-4.1E-05	0.106	0.106	0.104	0.164	0.171	0.180
3+ PUMAs w/in State	6.8E-06	1.4E-05	2.7E-05	-6.0E-06	-1.2E-05	-2.3E-05	0.322	0.321	0.320	0.361	0.365	0.369
1 State Away	1.3E-06	2.7E-06	5.3E-06	-1.2E-06	-2.4E-06	-4.7E-06	0.357	0.356	0.355	0.393	0.397	0.401
2+ States Away	5.7E-07	1.1E-06	2.3E-06	-5.3E-07	-1.1E-06	-2.1E-06	0.531	0.531	0.530	0.558	0.560	0.563
Out of Sample	7.4E-07	1.5E-06	3.0E-06	-6.9E-07	-1.4E-06	-2.7E-06	1	1	1	1	1	1

Panel B: Welfare Outcomes

	Change in E[Welfare] (in 2012 $)						Share of Welfare Gains or Losses					
Distance from	Plant Opening			Plant Closing			Plant Opening			Plant Closing		
Focal Tract	125	250	500	125	250	500	125	250	500	125	250	500
Target Tract	93	176	325	-4645	-8176	-13065	0.018	0.018	0.017	0.367	0.357	0.333
1 Tct Away	55	108	204	-43	-80	-141	0.044	0.043	0.042	0.380	0.370	0.346
2 Tcts Away	32	62	120	-25	-46	-83	0.071	0.070	0.068	0.394	0.384	0.360
Over 3 Tcts within PUMA	14	29	56	-13	-25	-47	0.128	0.127	0.125	0.427	0.418	0.394
1 PUMA Away	9	18	36	-8	-16	-30	0.176	0.175	0.173	0.456	0.446	0.423
2 PUMAs Away	6	12	23	-5	-11	-20	0.254	0.253	0.250	0.505	0.496	0.473
3+ PUMAs w/in State	3	6	12	-3	-5	-11	0.610	0.609	0.607	0.746	0.737	0.719
1 State Away	0	1	2	0	-1	-2	0.653	0.652	0.651	0.775	0.767	0.750
2+ States Away	0	0	0	0	0	0	0.746	0.745	0.744	0.834	0.828	0.815
Out of Sample	0	0	0	0	0	0	1	1	1	1	1	1

Notes: See Table A4 for expanded definitions of the distance bins captured by the row labels, as well as definitions of the outcome measures used in both panels. The column subheadings "125", "250", and "500" indicate the number of jobs in the focal tract that were either added in "plant opening" simulations or removed in "plant closing" simulations whose incidence is summarized in the chosen column. Each "plant opening" or "plant closing" adds positions to or removes positions from large, high-paying manufacturing establishments.

Table A13: Heterogeneity in the Change in P(Employed) and Cumulative Share of Total Employment Gains by Distance from Focal Tract Across Various Categories of Focal Tracts

Panel A: Urbanicity and # Jobs within 5 Miles

Distance from	Change in P(Employed)					Share of Employment Gains				
Focal Tract	All	Rural	Urban	Low	High	All	Rural	Urban	Low	High
Target Tract	0.001	0.002	0.001	0.002	4.4E-04	0.006	0.012	0.003	0.013	0.003
1 Tct Away	3.2E-04	0.001	8.6E-05	0.001	4.7E-05	0.015	0.030	0.006	0.029	0.006
2 Tcts Away	1.4E-04	2.3E-04	5.5E-05	3.0E-04	4.1E-05	0.026	0.047	0.013	0.047	0.014
3+ Tcts w/in PUMA	7.0E-05	9.3E-05	4.5E-05	1.0E-04	3.1E-05	0.058	0.088	0.037	0.088	0.037
1 PUMA	4.6E-05	4.8E-05	3.2E-05	4.9E-05	2.7E-05	0.087	0.114	0.067	0.117	0.066
2 PUMAs Away	3.1E-05	3.1E-05	2.6E-05	3.1E-05	2.5E-05	0.138	0.167	0.113	0.172	0.114
3+ PUMAs w/in State	1.6E-05	1.8E-05	1.1E-05	1.8E-05	1.3E-05	0.380	0.295	0.480	0.310	0.437
1 State Away	2.6E-06	3.2E-06	2.2E-06	3.0E-06	2.3E-06	0.452	0.392	0.526	0.401	0.497
2+ States Away	1.0E-06	1.1E-06	9.5E-07	1.1E-06	1.0E-06	0.593	0.553	0.642	0.556	0.622
Out of Sample	1.3E-06	1.4E-06	1.1E-06	1.4E-06	1.2E-06	1	1	1	1	1

Panel B: Two-Bedroom Apartment Rent and Poverty Rate

Distance from	Change in P(Employed)					Share of Employment Gains				
Focal Tract	All	Cheap	Expen.	Low	High	All	Cheap	Expen.	Low	High
Target Tract	0.001	0.002	3.3E-04	0.001	0.001	0.006	0.011	0.003	0.004	0.008
1 Tct Away	3.2E-04	0.001	8.0E-05	1.9E-04	4.0E-04	0.015	0.025	0.007	0.010	0.018
2 Tcts Away	1.4E-04	2.7E-04	5.1E-05	1.1E-04	1.7E-04	0.026	0.041	0.015	0.019	0.030
3+ Tcts w/in PUMA	7.0E-05	1.2E-04	3.3E-05	5.2E-05	8.7E-05	0.058	0.084	0.036	0.047	0.067
1 PUMA	4.6E-05	6.3E-05	2.4E-05	4.5E-05	5.3E-05	0.087	0.118	0.057	0.072	0.100
2 PUMAs Away	3.1E-05	4.0E-05	1.9E-05	2.8E-05	3.3E-05	0.138	0.173	0.101	0.125	0.147
3+ PUMAs w/in State	1.6E-05	2.3E-05	9.2E-06	1.4E-05	1.8E-05	0.380	0.288	0.483	0.393	0.384
1 State Away	2.6E-06	3.0E-06	2.1E-06	2.7E-06	2.5E-06	0.452	0.384	0.517	0.460	0.457
2+ States Away	1.0E-06	1.1E-06	9.6E-07	1.0E-06	1.0E-06	0.593	0.547	0.635	0.599	0.596
Out of Sample	1.3E-06	1.4E-06	1.1E-06	1.3E-06	1.3E-06	1	1	1	1	1

Notes: See Table A4 for expanded definitions of the distance bins captured by the row labels. The first five columns of Panel A provide the estimated change in the probability of employment in the destination year caused by a 250 job stimulus package for workers whose distance between their origin jobs and the census tract receiving the stimulus package place them in the distance bin indicated in the row label. The next five columns of Panel A provide the share of total stimulus-driven employment gains that accrue to workers whose distance between their origin jobs and the census tract receiving the stimulus package place them below or within the distance bin indicated in the row label. Each column displays the average employment outcome by distance bin among a subset of simulations featuring focal census tracts whose characteristics align with the column label. "All": An average of all 300 target census tracts chosen as sites of simulated stimulus packages. "Rural"/"Urban": An average over the 60 census tracts featuring the lowest/highest residential density (residents per square mile) among the full 300 target tracts simulated. "Low"/"High": In Panel A (B), an average over the 60 census tracts featuring the smallest/largest number of jobs within 5 miles (poverty rate) among the full 300 target tracts simulated. "Cheap"/Expen.": An average over the 60 census tracts featuring the cheapest/most expensive rent for a two-bedroom apartment among the full 300 target tracts simulated.

Table A14: Heterogeneity in the Average Welfare Gain and Cumulative Share of Total Welfare Gains by Distance from Focal Tract Across Various Categories of Focal Tracts

Panel A: Urbanicity and # Jobs within 5 Miles

Distance from	Avg. Welfare Gain ($)					Share of Welfare Gains				
Focal Tract	All	Rural	Urban	Small	Large	All	Rural	Urban	Small	Large
Target Tract	322	805	216	878	132	0.009	0.020	0.004	0.021	0.004
1 Tct Away	105	239	23	301	16	0.025	0.049	0.009	0.048	0.011
2 Tcts Away	51	90	16	110	14	0.044	0.079	0.018	0.077	0.024
3+ Tcts within PUMA	26	37	14	39	11	0.099	0.152	0.054	0.150	0.064
1 PUMA Away	17	20	10	19	9	0.150	0.201	0.103	0.202	0.115
2 PUMAs Away	12	14	8	13	9	0.242	0.306	0.174	0.308	0.196
3+ PUMAs w/in State	7	8	4	8	5	0.641	0.537	0.736	0.554	0.703
1 State Away	1	1	1	1	1	0.740	0.683	0.793	0.685	0.777
2+ States Away	0	0	0	0	0	0.809	0.759	0.855	0.756	0.845
Out of Sample	0	0	0	0	0	1	1	1	1	1

Panel B: Two-Bedroom Apartment Rent and Poverty Rate

Distance from	Avg. Welfare Gain ($)					Share of Welfare Gains				
Focal Tract	All	Cheap	Expen.	Low	High	All	Cheap	Expen.	Low	High
Target Tract	322	747	96	200	394	0.009	0.019	0.004	0.005	0.013
1 Tct Away	105	264	26	69	127	0.025	0.043	0.012	0.016	0.030
2 Tcts Away	51	114	15	36	60	0.044	0.071	0.022	0.031	0.051
3+ Tcts within PUMA	26	49	10	19	31	0.099	0.149	0.054	0.079	0.114
1 PUMA Away	17	26	8	16	19	0.150	0.213	0.085	0.125	0.171
2 PUMAs Away	12	17	6	10	12	0.242	0.319	0.155	0.217	0.253
3+ PUMAs w/in State	7	11	3	6	7	0.641	0.541	0.748	0.649	0.645
1 State Away	1	1	1	1	1	0.740	0.687	0.789	0.743	0.743
2+ States Away	0	0	0	0	0	0.809	0.754	0.851	0.814	0.812
Out of Sample	0	0	0	0	0	1	1	1	1	1

Notes: See Table A4 for expanded definitions of the distance bins captured by the row labels. The first five columns provide the estimated gain in expected welfare (scaled in $ of 2023 annual earnings) in the destination year caused by a 250 job stimulus package for workers whose distance between their origin jobs and the census tract receiving the stimulus package place them in the distance bin indicated in the row label. The next five columns provide the share of total stimulus-driven welfare gains that accrue to workers whose distance between their origin jobs and the census tract receiving the stimulus package place them below or within the distance bin indicated in the row label. Each column displays the average welfare outcome by distance bin among a subset of simulations featuring focal census tracts whose characteristics align with the column label. "All": An average of all 300 target census tracts chosen as sites of simulated stimulus packages. "Rural"/"Urban": An average over the 60 census tracts featuring the lowest/highest residential density (residents per square mile) among the full 300 target tracts simulated. "Low"/"High": In Panel A (B), an average over the 60 census tracts featuring the smallest/largest number of jobs within 5 miles (poverty rate) among the full 300 target tracts simulated. "Cheap"/Expen.": An average over the 60 census tracts featuring the cheapest/most expensive rent for a two-bedroom apartment among the full 300 target tracts simulated.

Table A15: Regressions Predicting Local Employment and Welfare Incidence Using Standardized Tract Characteristics: Stimuli Adding 250 Positions at Large High-Paying Manufacturing Firms

		All Target PUMA Workers				Low-Paid Target PUMA Workers				All Low-Paid U.S.	
Variable	Mean (S.D.)	Emp. Gain	Emp. Share	Wel. Gain	Wel. Share	Emp. Gain	Emp. Share	Wel. Gain	Wel. Share	Emp. Share	Wel. Share
Pop. Density	4887 (6866)	-6.2E-06 (4.0E-06)	-0.0055 (0.0008)	-2.9 (1.3)	-0.0125 (0.0016)	-6.1E-06 (6.9E-06)	-0.0033 (0.0005)	-2.1 (1.1)	-0.0033 (0.0005)	0.0022 (0.0005)	-0.0011 (0.0006)
Rent (Two-Bed)	1087 (462)	-5.0E-05 (5.0E-06)	-0.0129 (0.0009)	-20.9 (1.6)	-0.0284 (0.0020)	-7.9E-05 (8.5E-06)	-0.0076 (0.0007)	-20.1 (1.4)	-0.0097 (0.0006)	0.0100 (0.0006)	-0.0049 (0.0008)
Poverty Rate	0.155 (0.112)	1.0E-05 (4.0E-06)	0.0009 (0.0007)	2.5 (1.2)	0.0012 (0.0016)	1.9E-05 (6.8E-06)	0.0008 (0.0005)	1.9 (1.1)	0.0000 (0.0005)	0.0001 (0.0004)	-0.0011 (0.0006)
Job Density	2707 (9960)	3.1E-06 (3.2E-06)	0.0007 (0.0006)	1.8 (1.0)	0.0017 (0.0013)	4.2E-06 (5.5E-06)	0.0003 (0.0004)	1.4 (0.9)	0.0003 (0.0004)	-0.0014 (0.0004)	-0.0008 (0.0005)
Median Income	58050 (27190)	-3.2E-05 (5.6E-06)	-0.0009 (0.0010)	-11.2 (1.8)	-0.0037 (0.0023)	-4.5E-05 (9.6E-06)	-0.0004 (0.0007)	-10.5 (1.6)	-0.0028 (0.0007)	0.0009 (0.0006)	-0.0043 (0.0009)
Jobs w/in 5 Mi.	113100 (137300)	-5.4E-05 (4.2E-06)	-0.0016 (0.0008)	-13.8 (1.3)	0.0018 (0.0017)	-8.7E-05 (7.2E-06)	-0.0013 (0.0006)	-12.7 (1.4)	-0.0020 (0.0006)	0.0002 (0.0005)	-0.0019 (0.0006)
% College Grad.	0.279 (0.186)	2.9E-05 (4.5E-06)	0.0032 (0.0008)	13.5 (1.4)	0.0133 (0.0018)	4.7E-05 (7.7E-06)	0.001313. (0.0006)	2 (1.3)	0.0045 (0.0006)	-0.0063 (0.0005)	0.0020 (0.0007)
% PUMA Same Ind.	0.082 (0.044)	1.1E-05 (3.1E-06)	0.0004 (0.0006)	11.2 (1.0)	0.0069 (0.0013)	1.2E-05 (5.3E-06)	-0.0002 (0.0004)	10.0 (0.9)	0.0020 (0.0004)	0.0000 (0.0003)	0.0034 (0.0005)
Outcome Mean	–	1.9E-04	0.0514	58.1	0.0987	3.4E-04	0.0342	53.2	0.0318	0.6579	0.2853
R^2	–	0.262	0.203	0.306	0.177	0.224	0.165	0.323	0.245	0.187	0.118
N	–	3200	3200	3200	3200	3200	3200	3200	3200	3200	3200

Notes: This table reports regression coefficients and their accompanying standard errors (in parentheses) from tract-level regressions based on 3200 simulated stimulus packages creating 250 new positions at large, high-paying manufacturing firms in different randomly chosen focal tracts. Simulated employment and welfare outcomes listed in the column label are regressed on standardized versions of the tract characteristics associated with the focal tract that are listed in the row labels. Tract characteristics were collected by Chetty and Hendren (2018). The first four columns consider as regressands mean outcomes and shares of aggregate gains accruing workers initially in the focal PUMA receiving the stimulus, while the next four display the same regressands computed for the low-paid subset of focal PUMA workers (initially in the bottom two earnings quartiles). The final two columns display shares of employment and welfare gains accruing to low-paid workers nationally (rather than high-paid or initially unemployed workers). "Pop. Density": The focal tract's number of residents per square mile. "Rent (Two-Bed)": The focal tract's average monthly rent for a two-bedroom apartment. "Poverty Rate": The focal tract's share of households below the federal poverty line. "Job Density": The focal tract's employment per square mile. "Median Income": The focal tract's household median income. "Jobs w/in 5 Mi.": The number of jobs within 5 miles of the focal tract. "% College Grad.": The share of the focal tract's adult residents who are college graduates. "% PUMA Same Ind.": The share of the focal PUMA;s residents who were initially employed in firms in the manufacturing industry.

Table A16: Regressions Predicting Local Employment and Welfare Incidence Using Standardized Tract Characteristics: Stimuli Adding 250 Positions at Large Low-Paying Retail Firms

		All Target PUMA Workers				Low-Paid Target PUMA Workers				All Low-Paid U.S.	
Variable	Mean (S.D.)	Emp. Gain	Emp. Share	Wel. Gain	Wel. Share	Emp. Gain	Emp. Share	Wel. Gain	Wel. Share	Emp. Share	Wel. Share
Pop. Density	4887	-5.8E-06	-0.0051	-2.8	-0.0127	-4.7E-06	-0.0034	-2.2	-0.0044	0.0028	-0.0004
	(6866)	(4.1E-06)	(0.0008)	(1.0)	(0.0013)	(8.1E-06)	(0.0006)	(1.3)	(0.0006)	(0.0004)	(0.0004)
Rent (Two-Bed)	1087	-5.1E-05	-0.0126	-19.7	-0.0296	-9.1E-05	-0.0083	-24.4	-0.0126	0.0125	-0.0012
	(462)	(5.0E-06)	(0.0010)	(1.3)	(0.0016)	(9.9E-06)	(0.0008)	(1.6)	(0.0007)	(0.0005)	(0.0005)
Poverty Rate	0.155	1.1E-05	0.0009	1.5	-0.0004	2.5E-05	0.0010	2.2	-0.0001	0.0009	0.0001
	(0.112)	(4.0E-06)	(0.0008)	(1.0)	(0.0013)	(7.9E-06)	(0.0006)	(1.3)	(0.0006)	(0.0004)	(0.0004)
Job Density	2707	1.1E-06	0.0004	1.0	0.0013	2.0E-06	0.0001	1.1	0.0003	-0.0010	-0.0003
	(9960)	(3.2E-06)	(0.0006)	(0.8)	(0.0010)	(6.4E-06)	(0.0005)	(1.1)	(0.0005)	(0.0003)	(0.0003)
Median Income	58050	-3.2E-05	-0.0008	-7.4	-0.0034	-5.4E-05	-0.0005	-8.9	-0.0021	0.0022	-0.0006
	(27190)	(5.7E-06)	(0.0011)	(1.4)	(0.0018)	(1.1E-05)	(0.0009)	(1.8)	(0.0008)	(0.0005)	(0.0006)
Jobs w/in 5 Mi.	113100	-5.8E-05	-0.0017	-12.3	0.0020	-1.1E-04	-0.0019	-16.2	-0.0023	0.0001	-0.0011
	(137300)	(4.3E-06)	(0.0008)	(1.1)	(0.0013)	(8.4E-06)	(0.0007)	(1.4)	(0.0006)	(0.0004)	(0.0004)
% College Grad.	0.279	2.1E-05	0.0009	8.5	0.0089	4.1E-05	-0.0002	11.0	0.0035	-0.0083	-0.0003
	(0.186)	(4.6E-06)	(0.0009)	(1.2)	(0.0015)	(9.1E-06)	(0.0007)	(1.5)	(0.0007)	(0.0004)	(0.0005)
% PUMA Same Ind.	0.177	-2.8E-05	-0.0081	-4.5	-0.0098	-5.6E-05	-0.0067	-6.3	-0.0055	-0.0015	-0.0004
	(0.032)	(3.1E-06)	(0.0006)	(0.8)	(0.0010)	(6.2E-06)	(0.0005)	(1.0)	(0.0005)	(0.0003)	(0.0003)
Outcome Mean	–	-2.0E-04	0.0545	48.7	0.0883	4.2E-04	0.0422	63.7	0.0411	0.6871	0.3450
R^2	–	0.280	0.222	0.311	0.272	0.247	0.191	0.299	0.285	0.357	0.027
N	–	3200	3200	3200	3200	3200	3200	3200	3200	3200	3200

Notes: This table reports regression coefficients and their accompanying standard errors (in parentheses) from tract-level regressions based on 2500 simulated stimulus packages creating 250 new positions at large, low-paying retail firms in different randomly chosen focal tracts. Simulated employment and welfare outcomes listed in the column label are regressed on standardized versions of the tract characteristics associated with the focal tract that are listed in the row labels. Tract characteristics were collected by Chetty and Hendren (2018). The first four columns consider as regressands mean outcomes and shares of aggregate gains accruing workers initially in the focal PUMA receiving the stimulus, while the next four display the same regressands computed for the low-paid subset of focal PUMA workers (initially in the bottom two earnings quartiles). The final two columns display shares of employment and welfare gains accruing to low-paid workers nationally (rather than high-paid or initially unemployed workers). "Pop. Density": The focal tract's number of residents per square mile. "Rent (Two-Bed)": The focal tract's average monthly rent for a two-bedroom apartment. "Poverty Rate": The focal tract's share of households below the federal poverty line. "Job Density": The focal tract's employment per square mile. "Median Income": The focal tract's household median income. "Jobs w/in 5 Mi.": The number of jobs within 5 miles of the focal tract. "% College Grad.": The share of the focal tract's adult residents who are college graduates. "% PUMA Same Ind.": The share of the focal PUMA;s residents who were initially employed in firms in the retail/wholesale industry.

Table A17: Assessing Robustness to Model Assumptions: Employment and Welfare Incidence from Plant Opening Simulations for Alternative Models

Panel A: Employment Outcomes

	Change in P(Employed)					Share of Employment Gains				
Distance from Focal Tract	Base Spec.	Job Mult.	Endo. Vac.	Choo Siow	Endo. Surp.	Base Spec.	Job Mult.	Endo. Vac.	Choo Siow	Endo. Surp.
Target Tract	6.7E-04	4.0E-04	6.3E-04	5.3E-04	7.8E-04	0.004	0.002	0.004	0.004	0.006
1 Tct Away	2.5E-04	2.7E-04	2.4E-04	2.3E-04	2.3E-04	0.008	0.005	0.008	0.007	0.007
2 Tcts Away	1.2E-04	2.2E-04	1.2E-04	1.1E-04	1.2E-04	0.010	0.011	0.010	0.009	0.010
3+ Tcts w/in PUMA	6.1E-05	1.7E-04	5.9E-05	5.6E-05	6.7E-05	0.028	0.042	0.028	0.025	0.031
1 PUMA Away	3.9E-05	6.1E-05	3.8E-05	3.9E-05	3.9E-05	0.025	0.022	0.025	0.023	0.025
2 PUMAs Away	2.8E-05	4.2E-05	2.7E-05	2.8E-05	2.8E-05	0.046	0.042	0.046	0.038	0.046
3+ PUMAs w/in State	1.6E-05	2.4E-05	1.6E-05	1.6E-05	1.6E-05	0.236	0.217	0.236	0.219	0.234
1 State Away	2.5E-06	4.2E-06	2.5E-06	2.6E-06	2.5E-06	0.070	0.069	0.070	0.073	0.069
2+ States Away	1.1E-06	1.9E-06	1.0E-06	1.1E-06	1.1E-06	0.144	0.150	0.144	0.154	0.144
Out of Sample	1.3E-06	2.3E-06	1.3E-06	1.3E-06	1.3E-06	0.429	0.441	0.429	0.448	0.428
$<$ 10 miles away						0.056	0.054	0.056	0.052	0.058
$<$ 250 miles away						0.235	0.231	0.235	0.212	0.237

Panel B: Welfare Outcomes

	Avg. Welfare Gain ($)					Share of Welfare Gains				
Distance from Focal Tract	Base Spec.	Job Mult.	Endo. Vac.	Choo Siow	Endo. Surp.	Base Spec.	Job Mult.	Endo. Vac.	Choo Siow	Endo. Surp.
Target Tract	296	342	267	349	417	0.009	0.006	0.009	0.006	0.015
1 Tct Away	98	137	90	150	108	0.015	0.013	0.015	0.014	0.017
2 Tcts Away	49	85	46	74	65	0.019	0.019	0.019	0.018	0.025
3+ Tcts within PUMA	27	60	25	41	33	0.056	0.067	0.055	0.051	0.066
1 PUMA Away	17	33	17	27	17	0.048	0.052	0.048	0.046	0.046
2 PUMAs Away	12	21	11	19	12	0.087	0.092	0.087	0.082	0.084
3+ PUMAs w/in State	7	11	7	11	7	0.385	0.378	0.386	0.365	0.371
1 State Away	1	1	1	1	1	0.092	0.089	0.092	0.092	0.088
2+ States Away	0	0	0	0	0	0.069	0.063	0.069	0.078	0.069
Out of Sample	0	0	0	0	0	0.220	0.221	0.221	0.247	0.219
$<$ 10 miles away						0.108	0.111	0.107	0.100	0.127
$<$ 250 miles away						0.436	0.453	0.434	0.413	0.448

Notes: See Table A4 for expanded definitions of the distance bins captured by the row labels, as well as definitions of the outcome measures used in both panels. The mean outcomes displayed for each of four alternative models are averages over 300 simulations with different focal tracts featuring the creation of 250 positions at large, high-paying manufacturing firms. "Base. Spec.": The baseline assignment model described in Sections 2 and 4; "Job Mult.": the baseline assignment model is augmented with a job multiplier process in which the original 250 manufacturing positions spawn additional service-sector jobs throughout the target PUMA, using a high-tech manufacturing multiplier of 1.71 from Bartik and Sotherland (2019); "Endo. Vac.": the baseline assignment model is augmented by allowing nearby firms to endogenously adjust the number of positions they wish to fill in response to stimulus-induced increases in required pay per efficiency unit of labor. Final equilibrium is determined by the convergence of a fixed point. "Choo Siow": the assignment model uses a Choo-Siow structure in which idiosyncratic part of the surplus consists of a worker-type $\times$ firm component and a worker $\times$ firm type component rather than a worker $\times$ firm component. "Endo. Surp.": The plant opening shock is allowed to change joint surplus values in addition to adding local positions to be filled. Surplus changes for all groups featuring within-PUMA worker and firm types are estimated using the average of revealed surplus changes based on worker reallocations from a set of observed high-paying manufacturing establishment openings between 2003 and 2012.

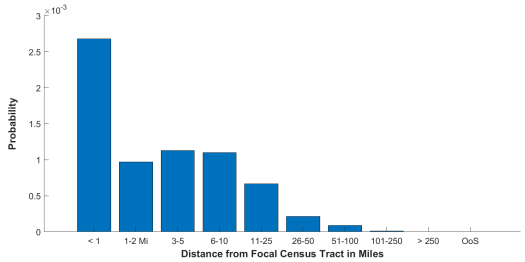
Table A18: Mean Simulated Welfare Gain for Local (Target Tract) Workers by Initial Earnings or Same Industry/Different Industry Category for Various Ways of Restricting Heterogeneity When Modeling the Joint Surplus from Forming Job Matches

Row	Main Spec.	No Firm Char.	No Worker Char.	No Same Ind.	No Same Firm
All	296	252	216	358	8633
Unemployed	267	383	288	302	276
1st Earn. Q.	201	261	201	256	4480
2nd Earn. Q.	274	228	206	320	7373
3rd Earn. Q.	362	219	208	415	13096
4th Earn. Q.	464	206	215	629	21990
Diff. Ind.	244	235	200	350	8503
Same. Ind.	1137	689	529	517	9270

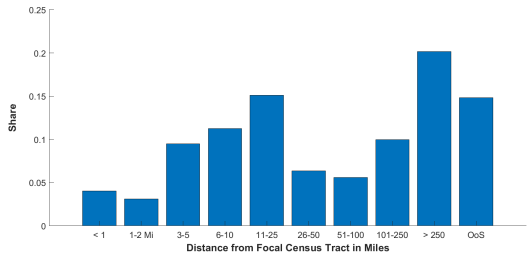
Notes: See Table 4 for expanded definitions of the worker subpopulations captured by the row labels. "Main Spec.": Main specification featuring unrestricted heterogeneity in match surpluses across worker-type/firm-type/job stayer combinations. "No Firm Char.": Removes any heterogeneity in match surpluses by non-location firm characteristics (industry, firm size, firm average pay). "No Worker Char.": Removes any heterogeneity in match surpluses by non-location worker characteristics (initial earnings quartile and age). "No Same Ind.": Removes heterogeneity in match surpluses based on whether the worker is staying within the same industry, conditional on changing firm. "No Same Firm": Removes heterogeneity in match surpluses based on whether the worker is being retained by the same firm.

Figure A1: 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, Distance Measured in Miles

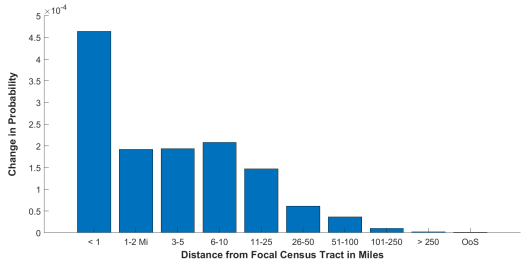
(a) Probability of Obtaining a Stimulus Job



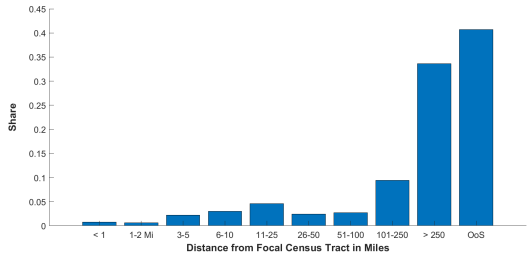
(b) Share of All Stimulus Jobs



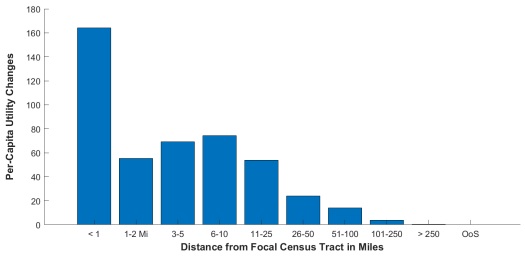
(c) Change in P(Employed)



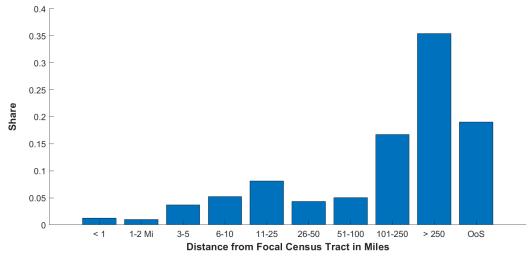
(d) Share of Employment Gains



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



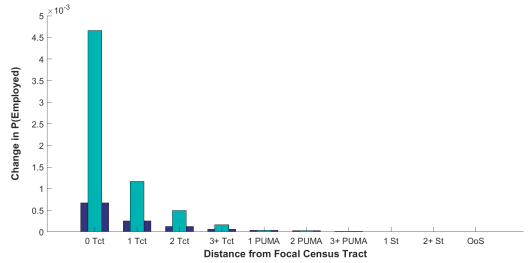
(f) Share of Welfare Gains



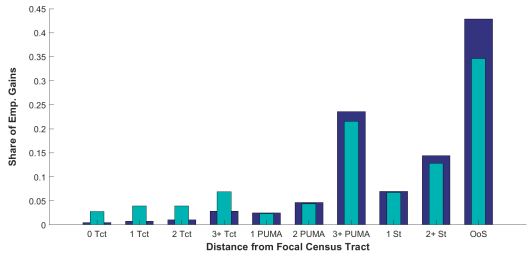
Notes:The bar heights in Figure A1a capture the average probability of obtaining a stimulus job among workers whose number of miles between their initial establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. These probabilities average across different demographic categories and across stimulus packages featuring different firm compositions. Figure A1b displays the share of all stimulus jobs generated by the stimulus that redounds to workers in the chosen distance bin. Figures A1c and A1d display the corresponding gains in employment probability and shares of national employment gains accruing to workers in each distance bin, while Figures A1e and A1f 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). "OoS" indicates that the worker's position was in an out-of-sample state.

Figure A2: Assessing the Value of Restricting Stimulus Jobs to Fill Positions With Workers from the Target PUMA: Spatial Employment and Welfare Incidence for Restricted and Unrestricted 250-Position Stimulus Packages

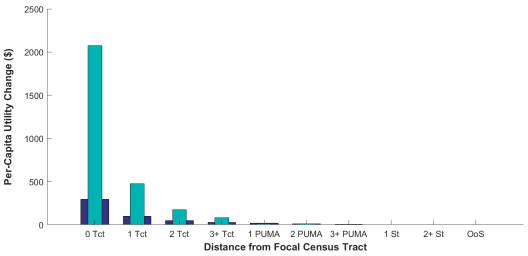
(a) Change in P(Employed)



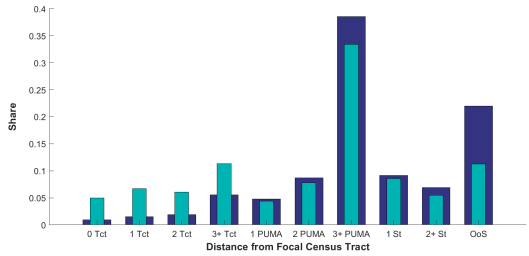
(b) Share of Employment Gains



(c) Avg. Welfare Gain ($)



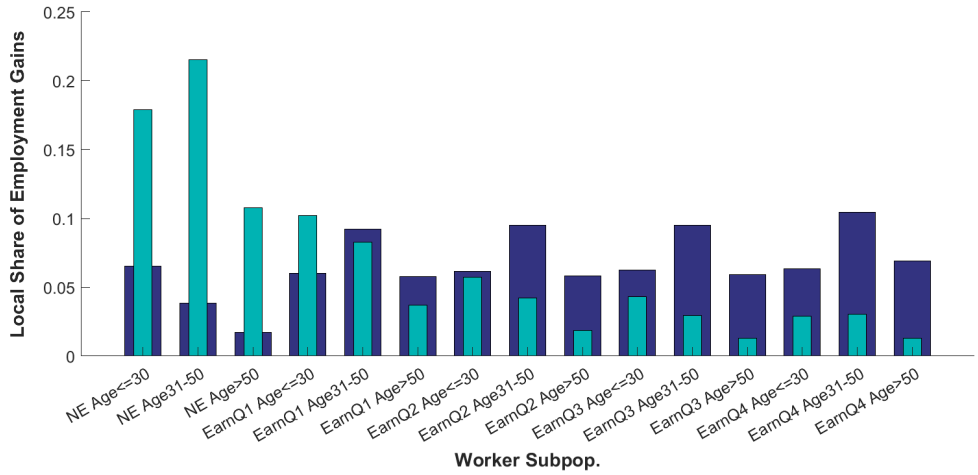
(d) Share of Welfare Gains



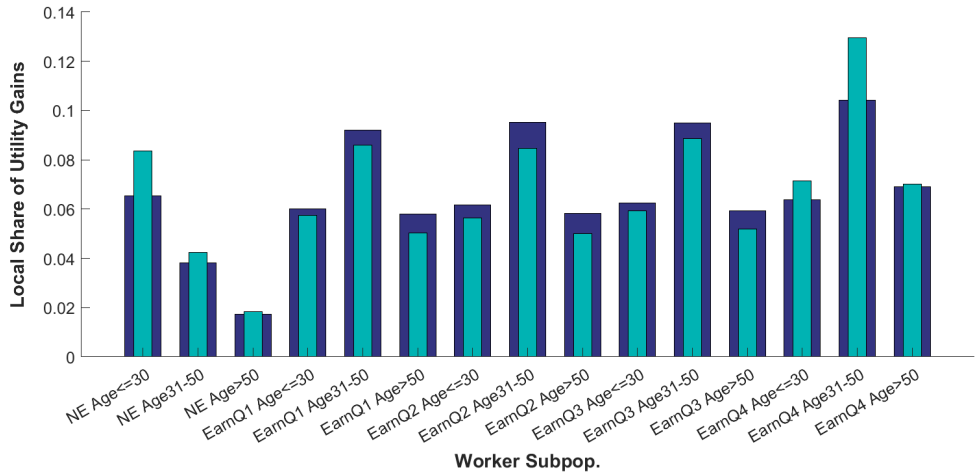
Notes: The bar heights capture the average measure of stimulus incidence associated with the chosen figure from a 250 person stimulus package among workers whose geographic distance between their initial establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the labels. The thin, light blue bars capture the case in which the new positions are restricted to be filled by existing workers within the targeted PUMA, while the wide, dark blue bars capture the case in which new positions can be filled by any worker. 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 A3: Comparing Shares of Focal Tract Employment and Utility Gains with Initial Focal Tract Workforce Shares Among Workers from Different Initial Earnings/Age Combinations: Average across All Simulated Stimuli

(a) Share of Focal Tract Net Employment Gains

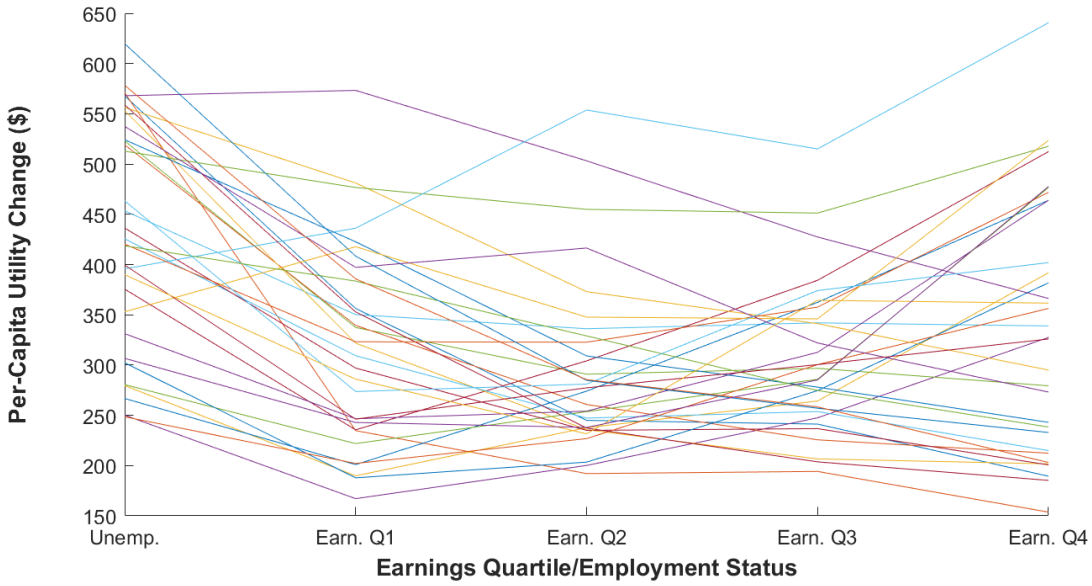


(b) Share of Focal Tract Utility Gains



Notes: The heights of the wider bars within a particular group in Figures A3a and A3b capture the initial share of the focal tract workforce associated with the subpopulation defined by the combination of earnings category and age category given by the label, while the heights of the narrower bars capture this 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.

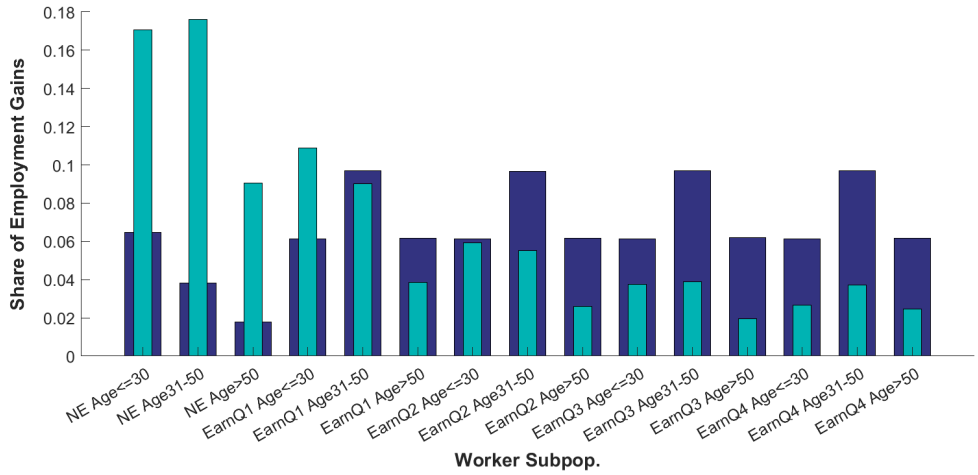
Figure A4: Expected Utility Changes Among Workers from the Targeted Tract by Initial Earnings/Employment Status: All Stimulus Packages



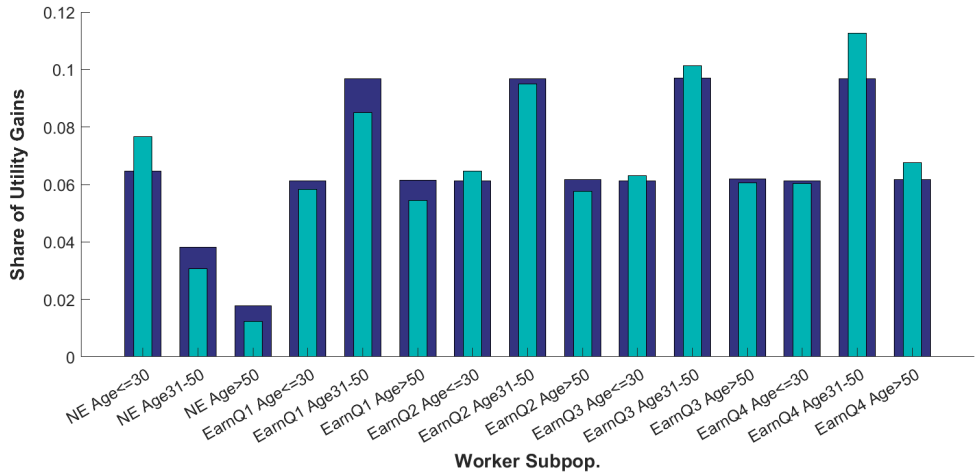
Notes: Each line traces the expected welfare gain among focal tract workers generated by a stimulus package featuring 250 positions among firms with a particular combination of supersector, firm size, and firm average pay categories across alternative unemployment or earnings quartile categories. 32 different lines corresponding to 32 different firm supersector/size/pay level compositions are displayed. Averages are taken across 300 simulations featuring different targeted census tracts for each supersector/firm size/firm avg. pay combo. "Unemp.": Workers who were not employed in the previous year. "Earn Q1/Q2/Q3/Q4": Workers whose pay at their dominant job in the previous year placed them in the 1st/2nd/3rd/4th quartile of the national age-adjusted annualized earnings distribution.

Figure A5: Comparing Shares of National Employment and Utility Gains with National Workforce Shares Among Workers from Different Initial Earnings/Age Combinations: Average across All Simulated Stimuli

(a) Share of Additional Employment



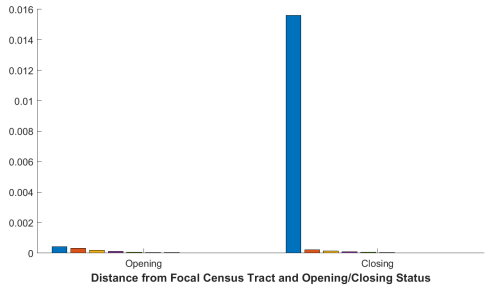
(b) Share of Total Utility Gains



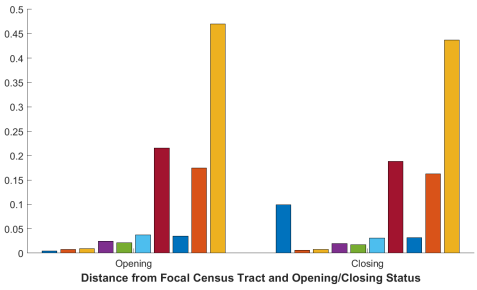
Notes: The heights of the wider bars within a particular group in Figures A5a and A5b capture the initial share of the national workforce associated with the subpopulation defined by the combination of earnings category and age category given by the label, while the heights of the narrower bars capture this subpopulation's share of the 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.

Figure A6: Asymmetry in Employment and Welfare Incidence from Plant Openings and Closings of Equivalent Magnitude

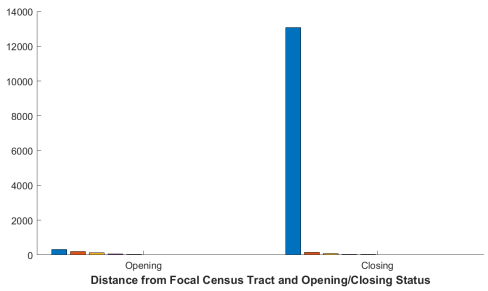
(a) Change in P(Employed)



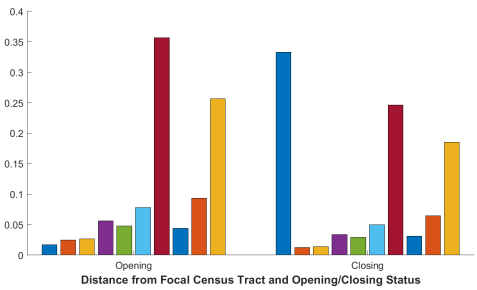
(b) Share of Employment Gains



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



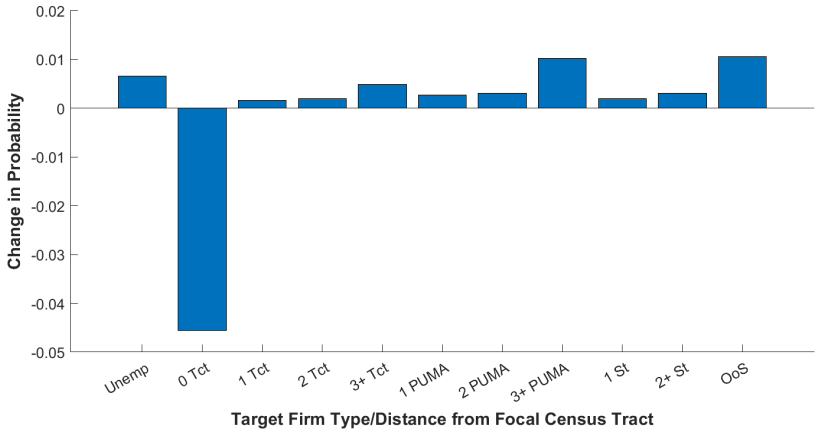
(d) Share of Welfare Gains



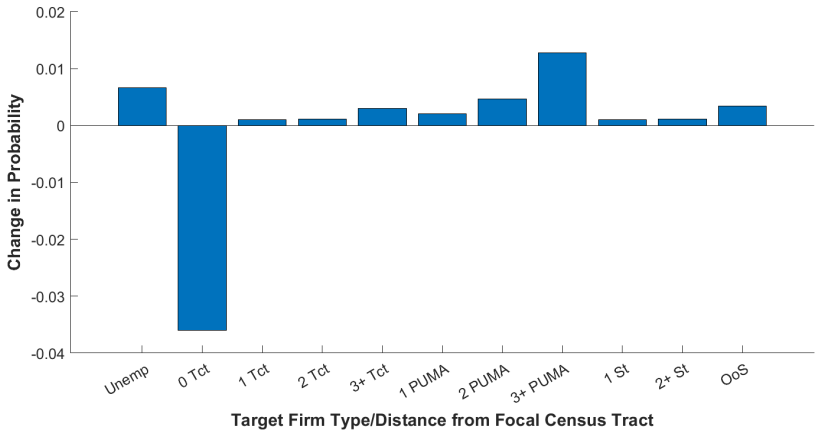
Notes: The bar heights within a particular group in Figures A6a-A6d capture the average value of the incidence measure associated with the figure from pairs of simulated plant openings and closings among workers whose geographic distance between their initial establishments and the census tract experiencing the disaster fell into the distance bins defined in Figure 2. Each opening or closing is associated with the creation or removal of 250 positions at large, high paying manufacturing firms in the focal tract. For each opening or closing, averages are taken across 200 simulations featuring different targeted census tracts.

Figure A7: Comparing Changes in the Distribution of Employment Locations (or Unemployment) for Focal Tract Workers after Plant Closings that Remove 250 Positions from either Large High-Paying Manufacturing Firms or Large Low-Paying Retail Firms

(a) High-Paid Manufacturing



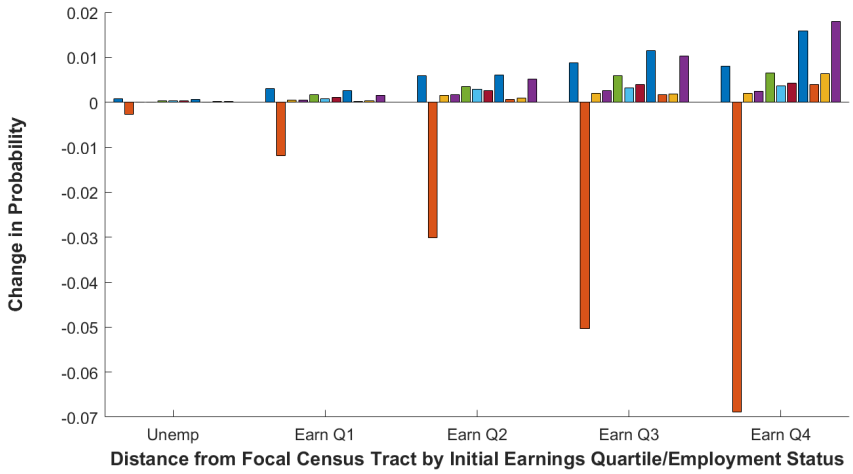
(b) Low-Paid Retail



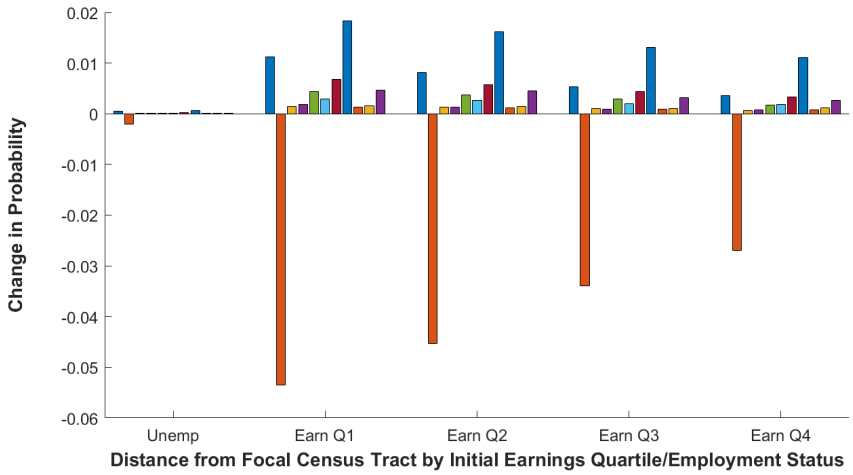
Notes: The bar heights in Figures A7a and A7b capture the impact of experiencing a plant or store closing, respectively, that removes 250 positions on the probability that a worker employed in the previous year (or most recently employed) in the targeted tract would be employed at a position whose distance from the targeted census tract fell into the distance bins defined in Figure 2 (or become/remain unemployed, the leftmost bar in each group). For both plant and store closings, averages are taken across 200 simulations featuring different targeted census tracts.

Figure A8: Sensitivity of the Change in the Distribution of Employment Locations (or Nonemployment) following Plant and Store Closings to the Match between Workers' Initial Earnings/Employment Status and the Closing Firm's Sector and Pay Level

(a) High-Paying Manufacturing: Change in Distribution of Destinations by Initial Earnings/Employment Status



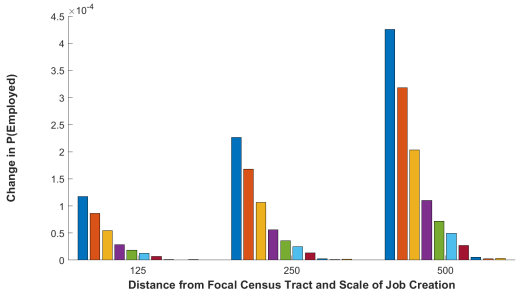
(b) Low-Paying Retail: Change in Distribution of Destinations by Initial Earnings/Employment Status



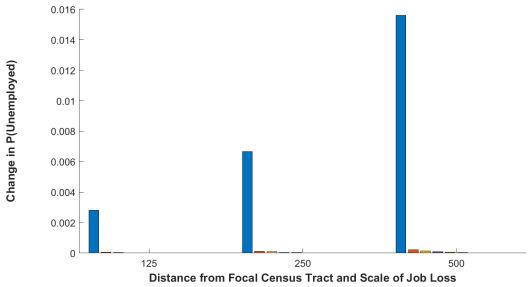
Notes: The bar heights within a particular group in Figures A8a and A8b capture the impact of experiencing a plant closing or store closing that removes 500 jobs in the target tract on the probability that a worker employed in the previous year (or most recently employed) in the targeted tract would be employed at a position whose geographic distance from the target tract fell into the distance bins defined in Figure A7a (or become/remain unemployed, the leftmost bar in each group). Each group of bars captures the change in destination employment probabilities among workers from the initial earnings/employment status given by the label. "Unemp": Workers who were unemployed in the origin year. "Earn Q1/Q2/Q3/Q4": Workers whose pay at their dominant job in the origin year placed them in the 1st/2nd/3rd/4th Quartile of the national age-adjusted annualized earnings distribution. Figure A8a considers a plant closing that removes 250 positions from large, high-paying manufacturing firms, while Figure A8b considers a store or mall closing that removes 250 positions from large, low-paying retail firms. For both plant and store closings, averages are taken across 200 simulations featuring different targeted census tracts.

Figure A9: Employment and Welfare Incidence from Plant Openings and Closings of Different Magnitudes: 125, 250, and 500 Jobs Created or Removed

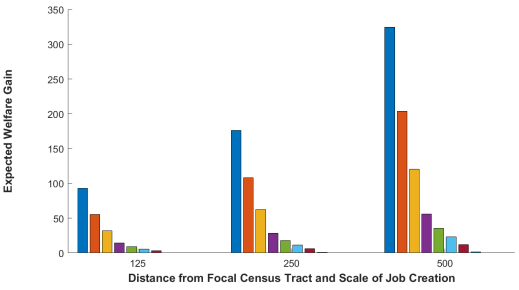
(a) Change in P(Employed): Openings



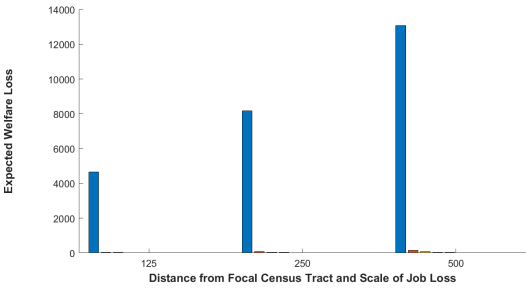
(b) Change in P(Unemployed): Closings



(c) Avg. Welfare Gain (in 2023 $): Openings



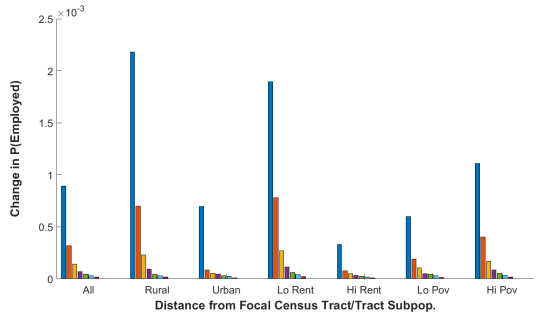
(d) Avg. Welfare Loss (in 2023 $): Closings



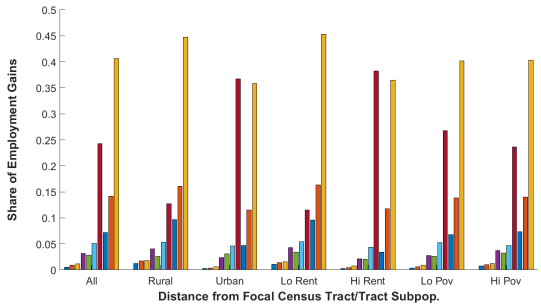
Notes: The bar heights within a particular group in Figures A9a-A9d capture the average value of the incidence measure associated with the figure from pairs of simulated plant openings and closings among workers whose geographic distance between their initial establishments and the census tract experiencing the disaster fell into the distance bins defined in Figure 2. Each opening or closing is associated with the creation or removal of 125, 250, or 500 positions at large, high paying manufacturing firms in the focal tract. For each opening or closing of each scale, averages are taken across 200 simulations featuring different targeted census tracts.

Figure A10: Heterogeneity in the Geographic Concentration of Several Incidence Measures Across Various Subsets of Focal Tracts

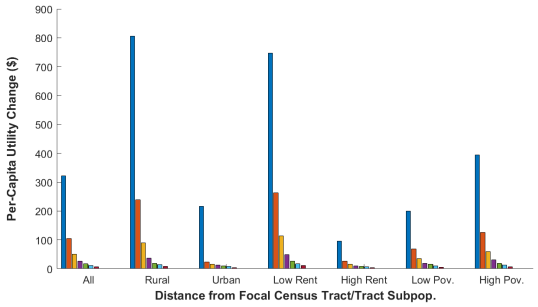
(a) Change in P(Employed)



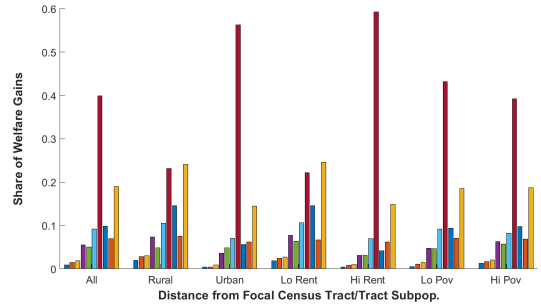
(b) Share of Employment Gains



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



(d) Share of Welfare Gains



Notes: The bar heights within a particular group in Figures A10a-A10d capture the average measure of stimulus incidence associated with the chosen figure from a 250 job stimulus package among workers whose geographic distance between their initial establishments and the census tract receiving the simulated stimulus package fell into the distance bins defined in Figure 2. Each group of bars displays this incidence distribution across distance bins for a particular subset (indicated by the group's label) of the 300 simulations featuring different focal tracts. In addition to averaging over the simulations featuring different target tracts within the chosen subset, the displayed results also average over different stimuli featuring the same target census tract but different firm compositions. "All": Average is taken among all 300 target tracts. "Rural"/"Urban": Average is taken among the 60 target tracts with the smallest/largest residential population density. "Lo Rent"/"Hi Rent": Average is taken among the 60 target tracts with the lowest/highest rent for a two bedroom apartment. "Lo Pov"/"Hi Pov": Average is taken among the 60 target tracts with the lowest/highest household poverty rate.