Labor Reallocation and Business Cycles*

Gabriel Chodorow-Reich
Harvard University and NBER

Johannes Wieland
UC San Diego

December 2014

Abstract

*Chodorow-Reich: Harvard University, Department of Economics, Littauer Center, Cambridge, MA 02138 (e-mail: chodorowreich@fas.harvard.edu); Wieland: University of California, San Diego, Department of Economics, 9500 Gilman Dr. #0508, La Jolla, CA 92093 (email: jfwieland@ucsd.edu). This research was conducted with restricted access to Bureau of Labor Statistics (BLS) data. The views expressed here do not necessarily reflect the views of the BLS or the U.S. government. We are grateful to Jessica Helfand and Michael LoBue of the BLS for their help with the Longitudinal Database.

1. Introduction

The effect of sectoral reallocation on business cycle outcomes remains an important and unresolved question. The importance stems from the relevance to policy discussions, most recently around the causes of the Great Recession, and from the implications for the presence of frictions in labor markets. The question remains unresolved due to two empirical challenges. First, the small number of national recessions in periods with high frequency, high quality industry level employment data limit inference based only on national variation. Second, reallocation within a business cycle may reflect cyclical sensitivities that vary across industries (Abraham and Katz, 1986), and business cycles can cause permanent reallocation of inputs (Schumpeter, 1942).

We provide new and robust evidence that labor reallocation affects the duration and severity of recessions, and the recovery. To circumvent the small number of national recessions, we study the effects of labor reallocation in broadly defined local labor markets in the United States. We identify 796 local recessions in 220 local labor markets during the period 1975-2014. Of these, 70% begin within six quarters of a national employment peak.

We address the problem of cyclical reallocation by developing a new methodology to isolate exogenous variation in sectoral shifts. We begin by defining a new measure of reallocation based on the change in industry employment shares between a local employment peak and the first month in which employment in the area surpasses its previous peak. Our measure consists of summing over industries the absolute value of the change in the industry's employment share between the two dates. Intuitively, the measure gives the minimum fraction of total peak employment that changes industries between the employment peak and the period of full recovery.

We then introduce an instrument for our reallocation measure to address issues of causality. Our instrument follows Bartik (1991) in using national industry trends to instrument for local outcomes. Specifically, we sum the absolute value of industry employment share changes *outside* each local area and over the period of the most recent *national* employment cycle, and then re-weight the changes using the local area peak employment shares. The instrument thus gives a measure of predicted reallocation based on an area's initial industry employment distribution and national industry employment trends.

We implement our exercise using confidential employment data by local area and industry from the Bureau of Labor Statistics Longitudinal Database, merged with the public use counterpart of these data, the QCEW. The confidential data permit analysis at a highly disaggregated (NAICS 6 digit) level. We use the public use version to extend the analysis back to 1975.

We find economically large effects of reallocation on business cycle outcomes. On average, an increase in reallocation of one standard deviation results in a recession 8 months longer and 2 percent deeper. The subsequent recovery – the period between the recession trough and the first month in which the area regains its previous employment peak – lasts an additional 18 months. These results hold for reallocation measured across different levels of industry aggregation, for sample subperiods, and for different measures of local labor markets. They do not exhibit sensitivity to the treatment of outliers.

Finally, we use a series of model economies to interpret our findings. We begin with a benchmark frictionless setting in which exogenous sectoral reallocation has no causal effect on business cycle outcomes. We use this framework to validate our identification strategy in economies with varying assumptions about traded and non-traded goods, nominal rigidities, and so forth. We then introduce a friction to switching industries which generates a causal effect of reallocation on the path of aggregate employment.

Our paper relates to literatures on the causes and consequences of input reallocation and business cycles. In an early and influential contribution, Lilien (1982) argued that sectoral shifts caused many of the fluctuations in unemployment in the 1970s, a point subsequently disputed by Abraham and Katz (1986). The Abraham and Katz critique of Lilien (1982) motivates much of our methodological approach. Debate over the importance of sectoral reallocation has renewed in the context of the slow recoveries from the most recent two national recessions (Groshen and Potter, 2003; Koenders and Rogerson, 2005; Aaronson, Rissman, and Sullivan, 2004; Berger, 2014; Sahin, Song, Topa, and Violante, 2014; Mehrotra and Sergeyev, 2012; Garin, Pries, and

Sims, 2013). Methodologically, our paper follows most closely Autor, Dorn, and Hanson (2013) and Charles, Hurst, and Notowidigdo (2014). Autor et al. study the effects of China's export growth on the U.S. commuting zones which had previously produced goods exported by China, while Charles et al. examine outcomes in MSAs experiencing large manufacturing declines. Our paper differs in its focus on business cycles rather than secular trends. As such, we construct a measure that does not rely on a specific source of variation in sectoral reallocation, and also construct a model to interpret our findings. Finally, our paper complements recent work on the consequences of reallocation at the worker level (Jaimovich and Siu, 2014; Fujita and Moscarini, 2013).

Section 2 defines our reallocation measure and instrument and places them in the context of the identification challenge and existing measures of reallocation. Section 3 describes the employment data and our concept of local labor markets. Section 4 presents summary statistics of the reallocation measure and our local business cycles. Section 5 contains the paper's core results on the effects of reallocation on business cycle outcomes. In section 6 we interpret these results through the lens of a series of model economies. Section 7 concludes.

2. Measurement and Identification

Our empirical strategy rests on two innovations in the measurement of reallocation in different areas and periods. First, we define a new measure of reallocation and apply it over a full employment cycle, rather than period by period. Second, we introduce an instrument for reallocation at the local level.

2.1. Measure of Reallocation

Consider an economy consisting of A distinct areas, each with I different industries. Let L_{ait} be employment in area a and industry i at time t, $L_{at} = \sum_{i=1}^{I} L_{ait}$ the total employment in the area, and $s_{ait} = {}^{L_{ait}}/_{L_{at}}$ industry i's employment share.

We define reallocation $R_{at,t+j}$ in area a between t and $t+j$ as the scaled sum of absolute sectoral employment share changes,

$$R_{a,t,t+j} = \frac{1}{2} \sum_{i=1}^{I} \left| s_{a,i,t+j} - s_{a,i,t} \right|. \tag{1}$$

$R_{a,t,t+j} = 0$ if employment grows at an identical rate in every industry, and $R_{a,t,t+j} = 1$ if all industries with positive employment in t disappear by $t + j$. In general, $R_{a,t,t+j} \subseteq [0, 1]$, with higher realizations indicating more reallocation. For the same area and time period, $R_{a,t,t+j}$ is weakly increasing in the level of industry disaggregation. For example, $R_{a,t,t+j}$ constructed over NAICS 4 digit industries will equal or exceed $R_{a,t,t+j}$ constructed over NAICS 3 digit industries. The difference between the $R_{a,t,t+j}$s constructed over broad and narrow industries gives the "within reallocation," the excess reallocation beyond that required for the broad definition.

We next define a full local business cycle for an area a. A local employment peak occurs in period t, $t = p_a$, if local employment in period t both surpasses its previous peak and is higher than employment in any of the next $\bar{J}$ months,

$$L_{a,p_a} > \arg \max_{k=1,2,..,\bar{J}} L_{a,p_a+k}, \quad L_{a,p_a} \geq L_{a,p_a-1}.$$

The cycle lasts T_a periods and ends when the area regains its previous level of employment,

$$T_a = \arg \min_{k>0} \text{ s.t. } L_{a,p_a+k} \geq L_{a,p_a}.$$

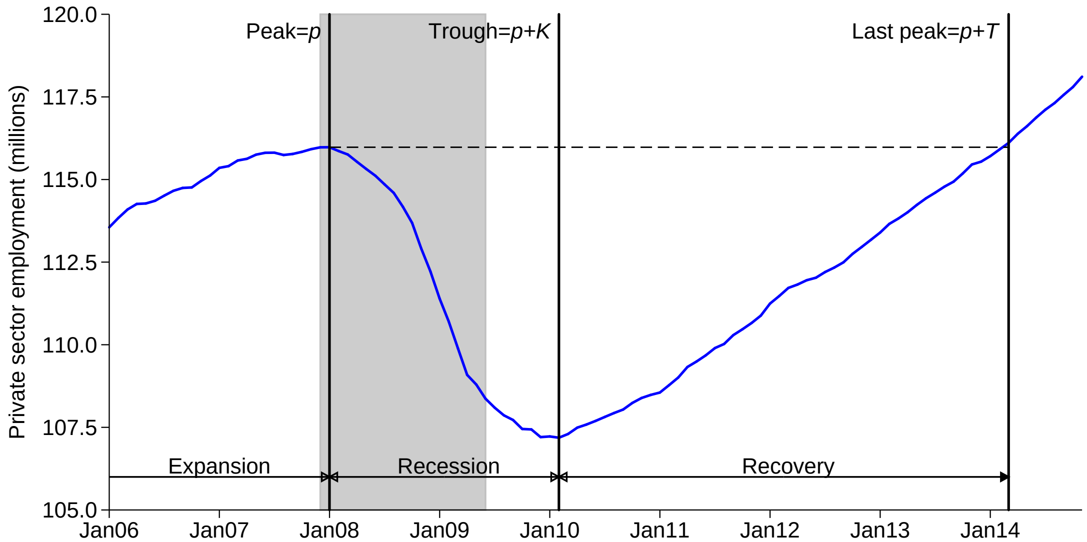
We call $p_a + T_a$ the "last-peak," since it is the date at which the economy regains the employment level from the last employment peak. We further divide the cycle into a recession of duration K_a, defined as the number of months between the peak and the employment trough,

$$K_a = \arg \min_{k \in [0,T_a]} L_{a,p_a+k},$$

and a recovery of length $T_a - K_a$.

Figure 1 illustrates the timing for total private sector employment in the U.S. economy between 2006 and 2014. For comparison, the shaded area shows the NBER recession. Using our timing convention, the peak occurs in January 2008 at 116 million employees, $p_a =$ January

Figure 1 – Recession timing example



2008. The employment trough is in February 2010, giving $K_a = 25$ months. The private sector regains its last peak level of employment in March 2014, $T_a = 74$ months.

Our first major departure from the existing literature consists of calculating our reallocation measure from peak to last-peak,

$$R_{a,p_a,p_a+T_a} = \frac{1}{2} \sum_{i=1}^{I} \left| s_{a,i,p_a+T_a} - s_{a,i,p_a} \right|, \tag{2}$$

which we call full cycle reallocation. The reason to do so stems from the critique of higher frequency measures by Abraham and Katz (1986). Abraham and Katz point out that industries differ in their cyclical sensitivities. For example, durable goods producing industries exhibit higher sensitivity to the cycle than education. As a result, the employment share in durable goods producers falls during recessions and the share in education increases, generating reallocation at the recession frequency. In this case, however, the business cycle causes a temporary reallocation across industries, rather than industry reallocation affecting the business cycle.

By definition, cyclical shifts in employment reverse during the recovery, from $p_a + K_a$ to $p_a + T_a$, and therefore do not contaminate full cycle reallocation. Thus, our full cycle measure isolates long-term trends in employment shares from cyclical reallocation. It also has the advantage, relative to much of the existing literature, of not requiring parametric time series models for either the cyclical component or the trend component of employment shares (see e.g. Brainard and Cutler, 1993; Aaronson et al., 2004; Mehrotra and Sergeyev, 2012, for articles that take the time series approach).

Finally, R_{a,p_a,p_a+T_a} has a natural and intuitive interpretation: it gives the minimum fraction of total peak employment that changes industries between the peak and last-peak.

2.1.1. Comparison to other measures Our full cycle reallocation measure differs from existing metrics mainly in the choice of horizon. In a seminal paper, Lilien (1982) measures sectoral dispersion as a weighted standard deviation of industry employment growth rates,

$$R^{\text{Lilien}}_{at,t+1} = \left[\sum_{i=1}^{I} s_{ai,t}\left(\Delta \ln L_{ai,t+1} - \Delta \ln L_{a,t+1}\right)^2\right]^{\frac{1}{2}}. \tag{3}$$

To illustrate the differences, we rewrite Lilien's measure using an absolute value metric rather than a Euclidean metric,

$$R^{\text{Lilien-absolute}}_{at,t+1} = \sum_{i=1}^{I} s_{ai,t}\left|\Delta \ln L_{ai,t+1} - \Delta \ln L_{a,t+1}\right|, \tag{4}$$

and take a first order approximation of equation (4) around the balanced growth path condition $s_{ai,t+1} = s_{ai,t}\ \forall i$, yielding

$$R^{\text{Lilien-absolute}}_{at,t+1} \approx \sum_{i=1}^{I} \left|s_{ai,t+1} - s_{ai,t}\right| = 2R_{a,t,t+1}. \tag{5}$$

Comparing equations (2), (3) and (5), our measure differs from Lilien's in the timing and the choice of metric. We prefer the absolute value metric over the Euclidean metric because it is less sensitive to outliers.

Our measure also has a close connection to the job reallocation rate defined by Davis and Haltiwanger (1992, p. 828),[1]

$$R^{\text{D-H}}_{at,t+1} = \frac{1}{0.5\left(L_{at+1} + L_{at}\right)} \sum_{i=1}^{I} \left|L_{ai,t+1} - L_{ait}\right| \tag{6}$$

$$= \sum_{a=1}^{I} \bar{s}_{ait,t+1}\left|g_{ai,t+1}\right|, \tag{7}$$

[1]Davis and Haltiwanger call this term SUM_t. In their application a corresponds to a sector, i to an establishment and I to the total number of establishments in that sector.

where $\bar{s}_{ait,t+1} \equiv \frac{(L_{ait+1}+L_{ait})}{(L_{at+1}+L_{at})}$ is the two period average employment share, and $g_{ai,t+1} \equiv \frac{(L_{ait+1}-L_{ait})}{0.5(L_{ait+1}+L_{ait})}$ is the symmetric growth rate of employment of industry i in area a. To illustrate the relationship between $R^{\text{D-H}}_{at,t+1}$ and R_{a,p_a,p_a+T_a}, we rewrite our full cycle reallocation measure in the case where employment at peak and at last-peak are exactly equal, $L_{a,p_a+T_a} = L_{a,p_a}$, as,

$$R_{a,p_a,p_a+T_a} = \frac{1}{2}\sum_{i=1}^{I} \bar{s}_{aip_a} \left|g_{ai,p_a+T_a}\right|.$$

Thus, up to the scale normalization, our measure coincides exactly with the Davis and Haltiwanger (1992) measure evaluated over a full cycle rather than period-by-period.

2.2. Instrument

Full cycle reallocation circumvents the Abraham and Katz (1986) critique by measuring reallocation between two periods of full employment. However, it does not address the possibility of causality running from business cycle outcomes to reallocation. For example, variation in cycle length T may directly affect our reallocation measure. To see this, suppose employment shares exhibit deterministic industry trends from p_a to $p_a + T_a$, $s_{a,i,t+j} = s_{a,i,t} + \Delta_i j$, where $\sum_{i=1}^{I} \Delta_i = 0$. Then

$$R_{a,p_a,p_a+T_a} = \frac{1}{2}\sum_{i=1}^{I} |\Delta_i|\, T_a$$

is mechanically increasing in T_a. By implication, any variable which increases T_a, such as local demand shocks, may also generate an increase in measured reallocation.

For these reasons, we introduce a Bartik (1991) type instrument for local reallocation. The construction of our instrument starts with matching every local business cycle from p_a to p_a+T_a to an aggregate business cycle from $\tilde{p}$ to $\tilde{p} + \tilde{T}$, based on the following rule:

1. If the local business cycle overlaps with an aggregate business cycle, then those two cycles are matched.

2. Otherwise, we match the local business cycle to the previous aggregate business cycle.

We then define our Bartik instrument as

$$R^{\text{Bartik}}_{a,p_a,\tilde{p},\tilde{p}+\tilde{T}} = \frac{1}{2}\sum_{i=1}^{I} \underbrace{\frac{s_{a,i,p_a}}{\tilde{s}_{-a,i,\tilde{p}}}}_{\text{Relative local exposure}} \underbrace{\left|\tilde{s}_{-a,i,\tilde{p}+\tilde{T}} - \tilde{s}_{-a,i,\tilde{p}}\right|}_{\text{Reallocation over national cycle}},$$

where $\tilde{s}_{-a,i,t}$ is the employment share of industry i at time t in all areas *excluding* area a.

Variation in the Bartik measure comes from differential exposure to national employment share trends. For the same reasons discussed in the previous subsection, we compute the national trends from national peak to national last-peak to avoid cyclical influence. The Bartik instrument thus provides a measure of predicted employment reallocation based on an area's initial industry distribution and the trend component of national employment shares during the current or most recent national business cycle. In our empirical implementation, we include fixed effects for the matched national recession to remove the confounding influence of national recession length on national employment share changes.

In sum, we define a new measure of local area reallocation based on the change in employment shares between an employment peak and the month in which the area regains its last peak in employment. We construct an instrument for this measure using an area's pre-recession industry distribution and national employment share trends. These two innovations form the core of an empirical strategy to estimate the effect of reallocation on business cycle outcomes. In section 6, we validate our strategy by re-estimating our empirical specifications in a series of model economies featuring differential employment trends and components such as traded and non-traded goods, nominal rigidities, and so forth.

3. Data

We implement our exercise by constructing cross industry employment reallocation in broadly defined local labor markets in the United States.

Our measures of employment by county and industry come from the Bureau of Labor Statistics Longitudinal Database (LDB) and Quarterly Census of Employment and Wages (QCEW). The LDB reports data by establishment and month and covers the period 1990-

2013. The source data come from quarterly reports filed by employers with state employment security agencies; as a result, the LDB contains essentially universal coverage of private sector employment. Each establishment in the LDB has a 6 digit NAICS code associated with its primary activity. Our LDB sample contains 42 states which allow access to their data through the BLS visiting researcher confidential data access program.

The QCEW is the public use version of the LDB. It contains data at the industry-county level for all 50 states from 1975-2014, subject to disclosure limititations to prevent the release of identifying information regarding single establishments. At the NAICS 2 level there are few such occurrences, but at the NAICS 6 level nearly three-quarters of county-industry cells are suppressed. Thus, our analysis of reallocation across narrow industry definitions necessitates the use of the confidential data.

We combine our LDB sample with NAICS 2 and 3 digit employment in the counties in states not in the LDB, and with 2 digit SIC data for 1975-2000. We also construct NAICS 4 digit and 6 digit employment for counties in our LDB sample from 1990-2013. The result is a data set with county by industry employment at multiple industry aggregation levels and at monthly frequency over the period 1975-2014.[2] Relative to other data sets with employment by geography and industry, such as the Census Bureau's County Business Patterns or Longitudinal Business Database (LBD), the BLS data have the important advantage of providing monthly rather than annual frequency, a requirement for the timing procedure described in section 2.

We seasonally adjust all series at the industry-county level using the multi-step moving average approach contained in the Census Bureau's X-11 algorithm. For NAICS industries with definition changes across the 2002, 2007, and 2012 revisions, we combine industries into the narrowest possible subset such that the subset remains unchanged across the revisions. We also define a new SIC classification, "SIC 1.5," which groups 2 digit SIC industries into 2 digit NAICS industries using the modal employment for 2 digit SIC industries which split into

[2]The QCEW reports employment by county and SIC 3 and 4 digit industry for 1984-2000. Because no national recessions occur between 1984 and 1990, when our LDB NAICS sample begins, we only make use of the 2 digit SIC data to extend our sample before 1990. While the QCEW data currently extend through March 2014, the LDB data end in December 2013.

multiple 2 digit NAICS industries.[3]

We aggregate county-level data into Core Based Statistical Areas (CBSAs). The Office of Management and Budget (OMB) defines CBSAs as areas "containing a large population nucleus and adjacent communities that have a high degree of integration with that nucleus," and distinguishes between Metropolitan (MSA) and Micropolitan (MiSA) areas depending on whether the urban core contains at least 50,000 inhabitants. We further aggregate CBSAs into Combined Statistical Areas (CSAs), again using OMB definitions.[4] CSAs consist of adjacent CBSAs that have "substantial employment interchange," and thus better capture the local labor market. Not all CBSAs belong to a CSA. For example, the San Diego MSA is not part of a CSA, but the Boston-Cambridge-Newton MSA is one of five MSAs in the Boston-Worcester-Providence CSA. In general, CSAs provide a more inclusive definition of a local labor market than do the widely-used commuting zones.

Our final sample includes all MSAs and CSAs containing at least one MSA, with employment of at least 50,000 in one month, and where we observe at least 95% of employment at the industry level. The last restriction binds because of disclosure limits in CSAs/MSAs that straddle states not in our LDB sample. The sample contains 1,312 of the 3,144 counties in the United States, covering 87% of 2013 employment.

4. Summary Statistics

4.1. Trends in national reallocation

We begin with an overview of our reallocation measure at the national level. Table 1 tabulates national full cycle reallocation for the last five recessions and at six levels of industry aggregation. We measure reallocation using SIC definitions for the March 1980, the August 1981 and the March 1990 recession, and using NAICS definitions for the March 1990, the December 2000 and the January 2008 recession (the dates refer to the private employment peak). As in

[3] Document this in an appendix.

[4] We use the 2013 OMB county classifications of CBSAs and CSAs for our entire sample to avoid discontinuities from counties switching CBSAs.

our LDB-QCEW merge, it helps to group SIC "1.5" with NAICS 2, SIC 2 with NAICS 3, and SIC 4 with NAICS 6.

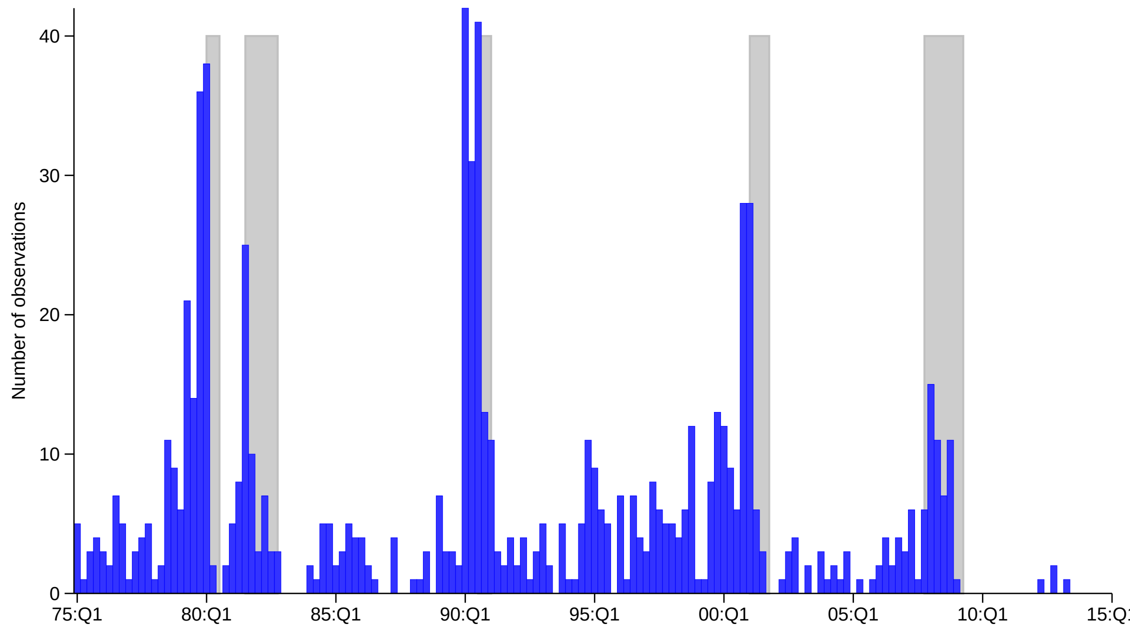
Table 1 – Reallocation by recession and industry detail

			Private sector employment peak				
Class	Detail	Industry count	Mar-1980	Aug-1981	Mar-1990	Dec-2000	Jan-2008
SIC	1.5	20	1.4	2.9	2.6		
NAICS	2	20			3.2	3.6	4.0
SIC	2	76	1.7	3.2	3.0		
NAICS	3	92			3.6	4.3	4.7
SIC	4	958	2.2	4.2	4.1		
NAICS	6	1028			4.8	6.3	6.7

For each SIC/NAICS group, the reallocation measures appear comparable for the overlap recession of March 1990. These similarities help to validate the groupings and facilitate comparison across time and classification. Indeed, table 1 shows a striking secular increase in full cycle reallocation. There is only half as much reallocation during the 1980 cycle as during the 1982 and 1990 cycles. From 1990 onwards each successive recession has more full cycle reallocation. Reallocation across NAICS 2 digit industries is 25% higher in the 2008 recession than in the 1990 recession; reallocation across NAICS 6 digit industries is 40% larger.[5] Comparing reallocation measures using the same industry classification and for the same recession reveals the monotonicity property in aggregation level discussed in section 2.1. For example, of the 6.7% of employment changing 6 digit NAICS industry between the January 2008 peak and the March 2014 last-peak, 4% constituted movement across 2 digit industries, 0.7% movement within 2 digit but across 3 digit industries, and 2% movement within 3 digit but across 6 digit industries.

[5] As noted in section 2, these patterns could partly reflect increasing cycle length. Beginning with the 1981 recession, each subsequent recession has a longer full employment cycle than the previous recession.

Figure 2 – Local recessions per quarter



Notes: the figure shows the timing of the 796 local employment peaks in the sample. 545 of the peaks occur in the six quarters preceding or during an NBER recession.

4.2. Local business cycles

We now turn to our local business cycles. Our timing definitions yield 796 local employment peaks in CSA/MSAs between 1975 and 2014.[6] Figure 2 displays their calendar frequency. Local peaks cluster around national business cycle peaks; 545 of the local peaks occur in the six quarters preceding or during an NBER recession.

Despite the clustering of local cycles, there is substantial variation in their length and depth, which we exploit in our estimation. Table 2 reports the mean and standard deviation of local recession length (K), recession depth ($|g_{p,p+K}|$), recovery length ($T-K$) and average recovery growth rate ($\bar{g}_{p+K,p+T}$). Recession length K and recovery length $T-K$ have standard deviations of 15 months and 19 months respectively. Recession depth has a standard deviation of 4.4% of total employment. The third column of table 2 reports the standard deviation based only on within-national-recession variation. More than 90% of the total variation in local cycles reflects

[6]The regressions that follow have fewer than 796 observations because we do not yet have full cycles for all of these peaks.

within cycle variation. Thus, relative to only national data, the large number of highly variable local cycles deliver a substantial increase in variation.

Table 2 – Mean and Standard Deviation of Dependent and Independent Variables

	Mean	St. Dev.	St. Dev. (within national recession)
K	17.46	14.66	13.14
$\lvert g_{p,p+K}\rvert$	5.41	4.41	4.05
$T-K$	20.75	19.47	17.41
$\bar{g}_{p+K,p+T}$	0.43	0.43	0.39
Reallocation	6.18	3.49	3.27
Reallocation Standardized	1.77	1.00	0.94
Bartik Reallocation	3.55	0.83	0.42
Bartik Reallocation Standardized	4.29	1.00	0.51

Notes: K is the number of months from the employment peak to the trough. $\lvert g_{p,p+K}\rvert$ is the cumulative employment decline from the period p peak to the month K trough. $T-K$ is the number of months from the employment trough to the end of the cycle in month T. $\bar{g}_{p+K,p+T}$ is the average growth rate from the trough in month K to the end of the cycle in month T. Reallocation is R_a, Bartik reallocation is R_{-a}. Standardized variables are scaled such that their standard deviation is equal to one.

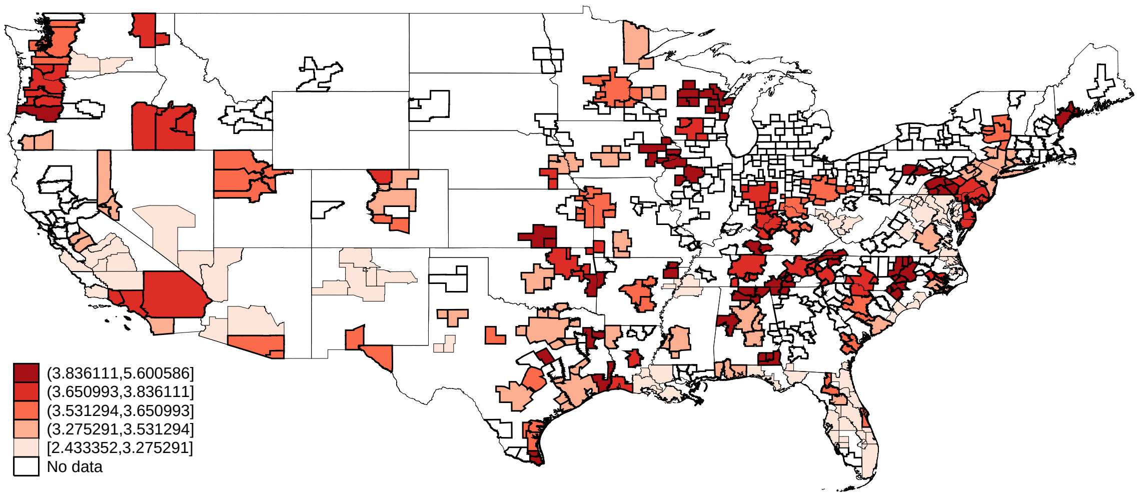
4.3. Instrument and control variables

We next discuss the variation in our instrument. As in our discussion of empirical outcomes, we focus on reallocation across NAICS 3/SIC 2 digit categories over the period 1975-2014.

Figure 3 shows a map of the variation in the instrument around the 2000 recession.[7] We provide the corresponding maps for the other national recessions in appendix ??. We split our MSA/CSAs observations into quintiles based on to their Bartik reallocation, and mark higher reallocation levels with darker shades of red. Note that all CBSAs belonging to an observation have the same color. Uncolored CBSAs belong to a CSA/MSA that does not have full cycle matched to this national recession. The map shows that Bartik reallocation is not easily explained by geographic factors. For instance, the Washington-Baltimore-Arlington CSA is in the bottom quintile of our instrument, but the adjacent Harrisburg-York-Lebanon is in the top quintile. The greatest concentration of high reallocation occurs in Eastern Tennessee and Western North Carolina, but these states also have low reallocation areas.

[7] For disclosure reasons, the map shows reallocation across 2 digit industries.

Figure 3 – Geographic Distribution of Bartik



Notes: the figure shows the geographic distribution of Bartik for the national employment peak in December 2000.

To help assess the validity of the exclusion restriction, table 3 reports partial correlations of the Bartik instrument with a set of MSA/CSA level variables, after partialling out national recession fixed effects. The MSA/CSA level variables include house price growth in the years before the peak;[8] employment growth over the 4, 10, and 15 years before the peak; area size, measured by the log of sample mean employment; the Herfindahl of industry employment concentration at the peak; and two Bartik variables for predicted recession depth and recovery speed.

The Bartik business cycle variables merit additional discussion. With inclusion of national recession fixed effects, variation in our instrument comes only from variation in initial industry distribution at the start of a recession. It follows that an area may have large Bartik real-

[8]We construct area house price indexes using the Freddie Mac MSA house price indexes, available 1975-2014. For CSAs combining multiple MSAs, we construct a CSA index as a geometric weighted average of the MSA indexes, using 1990 employment as weights. Noting that our data start in 1975 and the first national recession begins in 1980, we use a 4 year change to minimize loss of observations while still allowing for business cycle frequency lag length.

location because it began the recession with large employment shares in industries shrinking nationally over the course of the national full cycle, or because it began the recession with large employment shares in industries growing nationally. Concretely, in the former case, an area may experience a longer or deeper recession because it began the recession with employment in industries experiencing negative secular shocks, rather than the reallocation causing a deeper recession. We find both cases interesting, but wish to distinguish them. The Bartik recession depth variable predicts local recession depth using local exposure to national peak-to-trough industry employment growth,

$$R^{\text{Bartik-depth}}_{a,p_a,p_a+K_a} = 2 \sum_{i=1}^{I} s_{a,i,p_a} \underbrace{\left(\frac{L_{-a,i,p_a+K_a} - L_{-a,i,p_a}}{L_{-a,i,p_a+K_a} + L_{-a,i,p_a}} \right)}_{\substack{\text{National industry employment} \\ \text{growth peak-to-trough}}},$$

Thus, the Bartik recession depth control orthogonalizes Bartik reallocation with respect to the direct cyclical implications of an area's initial industry distribution. We construct the Bartik recovery growth rate symmetrically to control for the direct cyclical implications of an area's initial industry distribution during the recovery.

The pairwise partial correlation coefficients reported in table 3 are all less than 0.12 in absolute value. Bartik reallocation does not appear to simply reflect one of these latent factors. Reflecting the small correlations, inclusion or not of the control variables has in general a small effect on the point estimates for reallocation reported in section 5. However, these variables do absorb variation in business cycle outcomes, so we include them in our benchmark specification but also report results without the control variables for comparison.

Finally, table 4 presents evidence of the serial correlation in our instrument over successive national business cycles. The table reports the pairwise correlations in CSA Bartik reallocation associated with each national business cycle. A consistent pattern does not emerge. Bartik predicted reallocation has a positive correlation across some national recessions, a negative correlation across others, and in many of the pairs no significant correlation. The absence of strong serial correlation again helps in expanding the variation available to exploit. Still, in what follows we cluster all standard errors by CSA/MSA to account for arbitrary correlation

Table 3 – Instrument correlations

	Dependent variable: Bartik							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Right hand side variables:								
$\Delta_{4\text{ year}} \ln HPI$ at peak	-0.099^{**}							-0.12^{**}
	(0.032)							(0.029)
$\Delta_{4\text{ year}} \ln E$ at peak		-0.033						0.10^{+}
		(0.022)						(0.055)
$\Delta_{10\text{ year}} \ln E$ at peak			-0.062^{*}					-0.065^{*}
			(0.026)					(0.027)
$\Delta_{15\text{ year}} \ln E$ at peak				-0.0034				0.0078
				(0.029)				(0.025)
Log of mean employment					0.012			-0.0071
					(0.021)			(0.022)
Bartik recession depth						0.12^{**}		0.072^{*}
						(0.027)		(0.034)
Bartik recovery growth rate							-0.022	0.053
							(0.022)	(0.038)
Observations	668	665	478	436	684	684	684	436

Notes: All regressions include national recession fixed effects. Standard errors in parentheses and clustered by CSA-MSA.

Table 4 – Correlation of Reallocation across Recessions

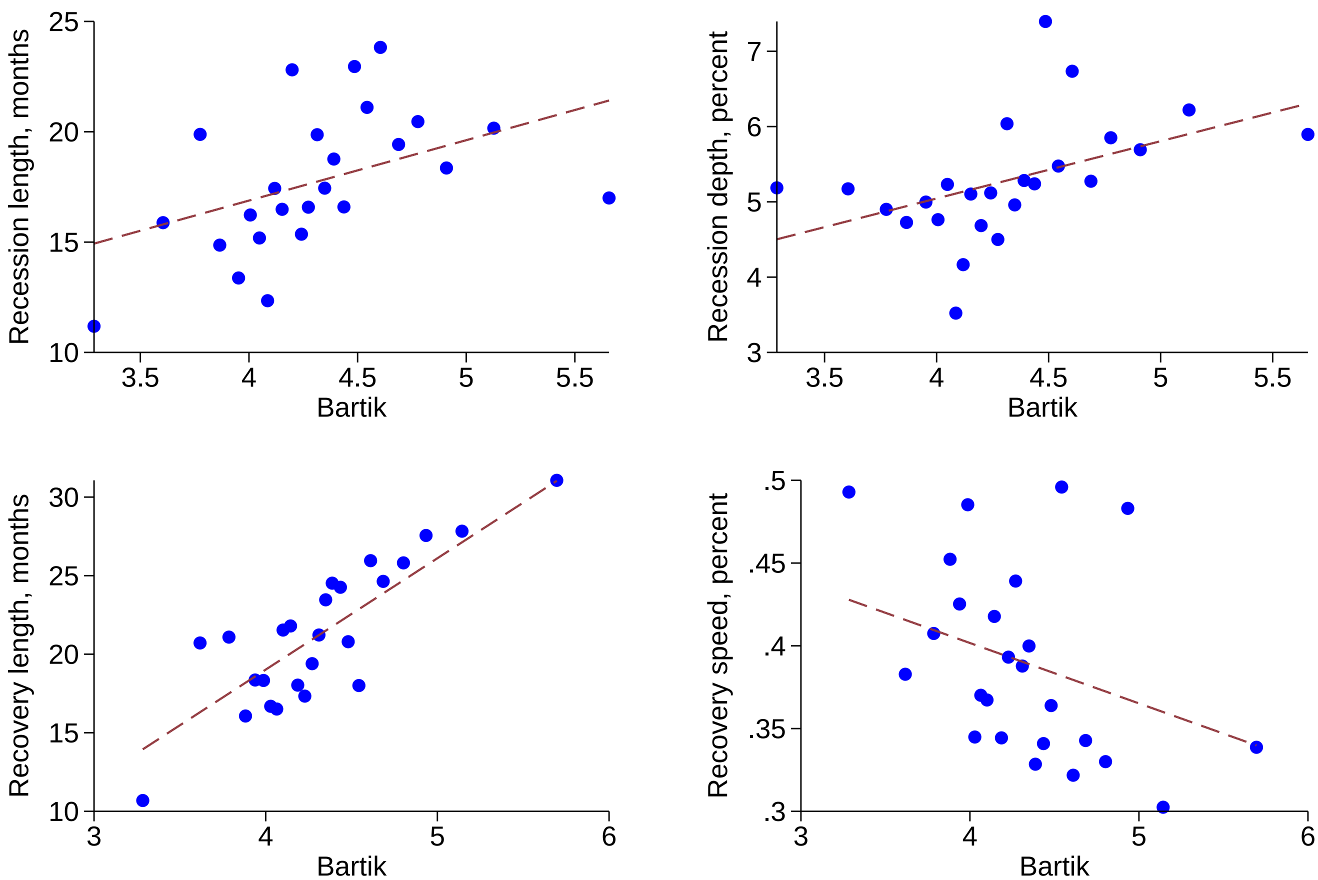
	Bartik:				
	Mar-80	Aug-81	Mar-90	Dec-00	Jan-08
Mar-80	1.00				
Aug-81	0.03	1.00			
Mar-90	0.08	0.27^{**}	1.00		
Dec-00	0.18	0.13^{*}	-0.27^{**}	1.00	
Jan-08	0.14	0.35^{**}	0.45^{**}	0.31^{**}	1.00

Notes: The table lists pairwise correlations of Bartik across recessions.

within a CSA/MSA over time.

5. Results

Figure 4 – Reduced form scatter plots, NAICS 3



Notes: Plotted points are means of 25 quantiles of the full set of observations. The Bartik measure is normalized to have unit variance. The slope of the dotted line equals the regression coefficient using the full set of underlying data.

5.1. Baseline results: 1975-2014, NAICS 3/SIC 2 Reallocation

We begin by showing non-parametric reduced form relationships of the four business cycle outcomes recession length K, recession depth $|g_{p,p+K}|$, recovery length $T-K$, and recovery speed $\bar{g}_{p+K,p+T}$ with the Bartik predicted reallocation instrument. Figure 4 reports these correlations by binning observations into 25 quantiles of the Bartik measure and showing the mean of the Bartik measure and outcome in each quantile.[9] The four panels of figure 4 are our eye test for reallocation mattering to business cycle outcomes.

The scatter plots illustrate statistically strong evidence of exogenous reallocation causing

[9]For comparability with our baseline regressions, we report means after partialling out national recession fixed effects and the ful set of control variables included in column 3 of table 5.

deeper and longer recessions, and especially longer recoveries.[10] The tightness of the fit between reallocation needs and recovery length appears particularly striking. Our reduced form estimates imply that an additional standard deviation in reallocation need results in a 7 month longer recovery from the employment trough to the last-peak.

We next turn to the instrumental variables estimates. Table 5 reports results with the recession outcomes, K and $\bar{g}_{p+K,p+T}$, as the dependent variables. Here and elsewhere, we standardize our reallocation measure to have unit variance. The bottom panel of the table reports the first stage F statistics for the excluded Bartik instrument. The F statistics range between 16.8 and 21.0 depending on specification, comfortably above the Stock and Yogo (2005) criteria for weak instruments.

Consistent with the reduced form scatter plots, the IV estimates indicate a statistically strong and economically large effect of reallocation on recession length and depth. Columns (3) and (6) report our preferred specification, with the full set of control variables discussed in section 4.3. A one standard deviation increase in reallocation results in a recession 7.6 months longer, and 2.1% deeper. As previewed in section 4.3, the point estimate changes little with and without the control variables.

Table 6 reports our estimates of the effect of reallocation on recoveries. Columns (1)-(4) indicate a one standard deviation increase in reallocation has a precisely estimated and economically very large effect on recovery length of 17 months. The standard errors are sufficiently tight to reject an increase of less than 13 months at the 95% level. Columns (5)-(8) suggest reallocation may slow the pace of the recovery as well, but here the estimates lack precision.

Together, these results indicate reallocation results in longer and deeper recessions, and longer recoveries.

[10] Statistical significance corresponds to the tightness of the scatter plot around the dashed OLS regression line.

Table 5 – Reallocation recession effects, IV, NAICS 3

	Dependent variable:					
	Recession length K			Recession depth $\lvert g_{p,p+K}\rvert$		
	(1)	(2)	(3)	(4)	(5)	(6)
Right hand side variables:						
Reallocation	8.10*	6.91*	7.60**	1.89*	1.49	2.11**
	(3.18)	(3.13)	(2.52)	(0.84)	(0.93)	(0.71)
Bartik recession depth		0.39	0.25		0.13+	0.15*
		(0.34)	(0.38)		(0.069)	(0.077)
$\Delta_{4\text{ year}} \ln HPI$ at peak			8.39*			3.74**
			(3.88)			(0.97)
$\Delta_{4\text{ year}} \ln E$ at peak			−1.62			2.49
			(3.95)			(2.27)
Log of mean employment			0.93			−0.28
			(0.89)			(0.24)
Herfindahl at peak			−8.11*			5.11**
			(4.03)			(1.44)
National recession FE	Yes	Yes	Yes	Yes	Yes	Yes
Excl. instruments F stat.	19.3	16.8	21.1	19.3	16.8	21.1
R^2	0.54	0.52	0.55	0.55	0.50	0.62
CSA-MSA clusters	220	220	217	220	220	217
Observations	684	684	665	684	684	665

Notes: The dependent variable is indicated in the table header. K is the number of months from the employment peak to the trough. $\lvert g_{p,p+K}\rvert$ is the cumulative employment decline from the period p peak to the month K trough. Standard errors in parentheses and clustered by CSA-MSA.

Table 6 – Reallocation recovery effects, IV, NAICS 3

	Dependent variable:							
	Recovery length $T-K$				Recovery speed $\bar{g}_{p+K,p+T}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Right hand side variables:								
Reallocation	17.2^{**}	17.2^{**}	17.6^{**}	18.1^{**}	-0.12	-0.11	-0.091	-0.16
	(3.03)	(3.05)	(2.20)	(6.45)	(0.13)	(0.13)	(0.086)	(0.24)
Bartik recovery growth rate		0.44	0.54	2.24		0.43^{*}	0.34	0.34
		(2.76)	(2.80)	(7.67)		(0.21)	(0.21)	(0.38)
$\Delta_{4\text{ year}} \ln HPI$ at peak			9.54^{*}				-0.096	
			(4.28)				(0.11)	
$\Delta_{4\text{ year}} \ln E$ at peak			-14.7				0.30	
			(11.5)				(0.29)	
Log of mean employment			4.08^{**}				-0.084^{**}	
			(0.81)				(0.029)	
Herfindahl at peak			2.37				0.51^{*}	
			(3.79)				(0.21)	
Recession length				-0.17				
				(0.21)				
Recession depth				0.29				
				(0.84)				
Recession length ($*10^{-2}$)								-0.42
								(0.73)
Recession depth ($*10^{-3}$)								40.6
								(31.8)
National recession FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excl. instruments F stat.	19.3	19.1	26.2	7.7	19.3	19.1	26.2	7.7
R^2	0.62	0.62	0.67	0.61	0.06	0.10	0.27	0.20
CSA-MSA clusters	220	220	217	220	220	220	217	220
Observations	684	684	665	684	684	684	665	684

Notes: The dependent variable is indicated in the table header. $T-K$ is the number of months from the employment trough to the end of the cycle in month T. $\bar{g}_{p+K,p+T}$ is the average growth rate from the trough in month K to the end of the cycle in month T. Standard errors in parentheses and clustered by CSA-MSA.

5.2. NAICS 2 and 4 digit Reallocation

Reallocation across NAICS 2 and 4 digit industries has effects similar to reallocation across 3 digit industries. Because we do not have SIC detail above the 2 digit level around the 1980 and 1981 recessions, and to ease comparisons, we report results for NAICS 2, 3, and 4 digit industries for the sample beginning in 1990.[11] This timing change also serves as a subsample stability check for the NAICS 3 digit results.

Tables 7 and 8 reports the IV estimates from our preferred specification. Broadly construed, the results are consistent both across industry aggregation levels and with the 1975-2014 sample results. The first stage is strongest at the 3 digit level.[12] The effects of 3 digit reallocation on recession length and depth diminish when moving from the 1975-2014 to the 1990-2013 sample, but remain statistically significant, albeit marginally in the case of recession length. The effect on recovery length is of similar magnitude and precisely estimated across all specifications. The magnitude of the effect on recovery speed appears similar in the full and NAICS-only samples, but the standard error in the NAICS-only sample falls substantially such that we can reject zero effect at the 1% level.

Table 9 helps in interpreting these results. The table reports correlations of the Bartik instrument across different levels of aggregation. For example, column 1 shows the correlation of across NAICS 2 digit reallocation ($R^{\text{across 2}}$) with reallocation measured within 2 digit but across 3 digit ($R^{\text{within 2, across 3}}$), within 3 digit but across 4 digit ($R^{\text{within 3, across 4}}$), and within 4 digit but across 6 digit ($R^{\text{within 4, across 6}}$). The decomposition exploits the adding up identity $R^{\text{across 6}} = R^{\text{across 2}} + R^{\text{within 2, across 3}} + R^{\text{within 3, across 4}} + R^{\text{within 4, across 6}}$.

The positive correlation between $R^{\text{across 2}}$ and $R^{\text{within 2, across 3}}$ and $R^{\text{within 3, across 4}}$ partly explains why reallocation measured using $R^{\text{across 2}}$, $R^{\text{across 3}} = R^{\text{across 2}} + R^{\text{within 2, across 3}}$, and $R^{\text{across 4}} = R^{\text{across 3}} + R^{\text{within 3, across 4}}$ yield similar results. In contrast, the negative correlation

[11] The sample ends in 2013 because the 2014Q1 micro-data was not yet available. Since the 2008 national cycle ended in March 2014, we cannot construct a Bartik measure for local cycles matched to this recession, so that these observations are excluded from the regressions.

[12] In fact, the instrument has almost no predictive power for reallocation measured over industries narrower than 4 digit. We omit the results for brevity.

Table 7 – Reallocation IV regressions, NAICS 2, 3, 4

	Dependent variable:					
	K			$\lvert g_{p,p+K} \rvert$		
	NAICS 2	NAICS 3	NAICS 4	NAICS 2	NAICS 3	NAICS 4
	(1)	(2)	(3)	(4)	(5)	(6)
Right hand side variables:						
Reallocation	9.74^{*}	4.69^{+}	3.44	0.51	0.95^{**}	1.01^{*}
	(4.38)	(2.52)	(3.07)	(0.49)	(0.25)	(0.47)
National recession FE	Yes	Yes	Yes	Yes	Yes	Yes
CSA/MSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Excl. instruments F stat.	12.3	26.5	11.5	12.3	26.5	11.5
CSA-MSA clusters	143	143	143	143	143	143
Observations	281	281	281	281	281	281

Notes: The dependent variable is indicated in the table header. K is the number of months from the employment peak to the trough. $\lvert g_{p,p+K} \rvert$ is the cumulative employment decline from the period p peak to the month K trough. Reallocation is R_a, Bartik instrument is R_{-a}. CSA/MSA controls are the Bartik recovery growth rate, $\Delta_{4 \text{ year}} \ln HPI$ at peak, $\Delta_{4 \text{ year}} \ln E$ at peak, log of mean employment, and the Herfindahl at peak. Standard errors in parentheses and clustered by CSA-MSA.

Table 8 – Reallocation IV regressions, NAICS 2, 3, 4

	Dependent variable:					
	$T - K$			$\bar{g}_{p+K,p+T}$		
	NAICS 2	NAICS 3	NAICS 4	NAICS 2	NAICS 3	NAICS 4
	(1)	(2)	(3)	(4)	(5)	(6)
Right hand side variables:						
Reallocation	18.8^{**}	14.0^{**}	16.3^{**}	-0.31^{**}	-0.14^{**}	-0.16^{*}
	(3.98)	(2.15)	(3.03)	(0.11)	(0.046)	(0.081)
National recession FE	Yes	Yes	Yes	Yes	Yes	Yes
CSA/MSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Excl. instruments F stat.	13.0	29.5	13.1	13.0	29.5	13.1
CSA-MSA clusters	143	143	143	143	143	143
Observations	281	281	281	281	281	281

Notes: The dependent variable is indicated in the table header. K is the number of months from the employment peak to the trough. $T - K$ is the number of months from the employment trough to the end of the cycle in month T. $\bar{g}_{p+K,p+T}$ is the average growth rate from the trough in month K to the end of the cycle in month T. Reallocation is R_a, Bartik instrument is R_{-a}. CSA/MSA controls are the Bartik recovery growth rate, $\Delta_{4 \text{ year}} \ln HPI$ at peak, $\Delta_{4 \text{ year}} \ln E$ at peak, log of mean employment, and the Herfindahl at peak. Standard errors in parentheses and clustered by CSA-MSA.

Table 9 – Correlation of Bartik reallocation at different levels

	Reallocation level			
	Across 2	2-3	3-4	4-6
Across 2	1			
2-3	0.196^{**}	1		
3-4	0.255^{**}	0.347^{**}	1	
4-6	-0.338^{**}	-0.0920	0.114^{+}	1

The table reports the correlation coefficients of reallocation across different aggregation levels. 2-3 is reallocation across 3 but within 2 digit industries, and similarly for 3-4 and 4-6.

between $R^{\text{across 2}}$ and $R^{\text{within 4, across 6}}$ accounts for the weakness of the first stage using reallocation across 6 digit industries. We leave investigation of the economics behind this negative correlation to future work.

5.3. Robustness

Here we describe a number of robustness exercises.

- Winsorize dependent and independent variables at quartile ± 5xIQR.
- Drop observations with a peak more than 18 months from the national peak.
- Standardize reallocation to have unit variance within each national recession.
- Commuting zones instead of metropolitan and combined statistical areas.
- Control for population growth during cycle.
- Control for Autor et al. (2013) China exposure.

6. Model

7. Conclusion

References

Aaronson, Daniel, Ellen Rissman, and Daniel Sullivan, "Can Sectoral Reallocation Explain the Jobless Recovery?," *Federal Reserve Bank of Chicago Economic Perspectives*, 2004, (2Q), 36–49.

Abraham, Katharine and Lawrence F Katz, "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?," *Journal of Political Economy*, 1986, *94.*

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–68.

Bartik, Timothy J., *Who Benefits from State and Local Development Policies?*, W.E. Upjohn Institute for Employment Research, 1991.

Berger, David, "Countercyclical Restructuring and Jobless Recoveries," 2014.

Brainard, S. Lael and David M. Cutler, "Sectorial Shifts and Cyclical Unemployment Reconsidered," *The Quarterly Journal of Economics*, 1993, *108* (1), pp. 219–243.

Charles, Kerwin, Erik Hurst, and Matt Notowidigdo, "Masking and Aggregate Employment Changes: Housing Booms and Manufacturing Decline During the 2000s," 2014.

Davis, Steven J. and John Haltiwanger, "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *The Quarterly Journal of Economics*, 1992, *107* (3), pp. 819–863.

Fujita, Shigeru and Giuseppe Moscarini, "Recall Unemployment," 2013.

Garin, Julio, Mike Pries, and Eric Sims, "Reallocation and the Changing Nature of Economic Fluctuations," 2013.

Groshen, Erica and Simon Potter, "Has structural change contributed to a jobless recovery?," *Current Issues in Economics and Finance*, 2003, *9* (8).

Jaimovich, Nir and Henry Siu, "The Trend is the Cycle: Job Polarization and Jobless Recoveries," 2014.

Koenders, Kathryn and Richard Rogerson, "Organizational dynamics over the business cycle:

a view on jobless recoveries," *Review*, 2005, (Jul), 555–580.

Lilien, David M, "Sectoral shifts and cyclical unemployment," *The Journal of Political Economy*, 1982, pp. 777–793.

Mehrotra, Neil R and Dmitriy Sergeyev, "Sectoral Shocks, the Beveridge Curve and Monetary Policy," *mimeo*, 2012.

Sahin, Aysegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante, "Mismatch Unemployment," *American Economic Review*, 2014, *104* (11), 3529–64.

Schumpeter, Joseph A., *Capitalism, Socialism, and Democracy* 1942.

Stock, James and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," in Donald Andrews and James Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press, 2005, chapter 5, pp. 80–108.