Economic Quarterly—Volume 101, Number 4—Fourth Quarter 2015—Pages 275–301

Monitoring Economic Activity in Real Time Using Diffusion Indices: Evidence from the Fifth District

Santiago Pinto, Sonya Ravindranath Waddell, Pierre-Daniel G. Sarte

Information on the state of U.S. economic activity compiled by statistical agencies, such as the Bureau of Labor Statistics (BLS), the Bureau of Economic Analysis, or even sections of the Federal Reserve Board, is often released with a one-month lag and is subject to further revisions, typically at the three-month and one-year mark.[1] Moreover, information on economic activity collected by these agencies at a more regional level is limited, so that data on wages, inventories, or shipments at the level of a U.S. state, for example, are not easily obtainable in real time. In part to compensate for this lack of information, several regional Federal Reserve Banks, including Atlanta, Dallas, Kansas City, New York, Philadelphia, and Richmond, produce survey-based diffusion indices that attempt to monitor in real time the direction of change in various regional economic conditions. In this article, we provide an assessment of this effort by the Federal Reserve Bank of Richmond (FRBR) based on ex-post sectoral information related to employment and wages for the Fifth Federal Reserve District. We also provide an assessment of the extent to which more disaggregated regional diffusion indices, not currently constructed but under

We wish to thank R. Andrew Bauer, Jackson Evert, John A. Weinberg, and Zhu Wang for their comments and suggestions. The views expressed in this article are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Richmond or the Federal Reserve System. All errors are our own.

DOI: http://doi.org/10.21144/eq1010401

[1] The work by Croushore and Stark (2001) examines the implications of data revisions on the estimation of macroeconomic models. See Croushore (2011) for a review of the macroeconomic literature on real-time data.

consideration, are likely to inform individual states within the Fifth District.

Diffusion indices of the kind constructed by Federal Reserve Banks as well as many other institutions, such as the widely publicized Institute for Supply Management index (ISM) or the Michigan Survey of Consumers index of consumer sentiment (MSC), aim to measure the breadth of change in a variable of interest, say employment, based on the proportions of its disaggregated component series that move in different directions (increase, decrease, or remain unchanged). This traditional interpretation relying on notions of optimism and pessimism is discussed by Moore (1983) and is distinct from diffusion indices constructed using factor analytic methods in Stock and Watson (2002).

In this article, we build on work by Pinto, Sarte, and Sharp (2015) and highlight the fact that appropriately scaled diffusion indices (defined as the difference between the fractions of sectors that expanded and contracted) capture the contribution of changes in the extensive margin, or the breadth of change, to aggregate changes in a series of interest. For the case of employment in the Fifth Federal Reserve District, we show that changes in this extensive margin, measured by a synthetic diffusion index constructed from observed data, accounts for the bulk of changes in aggregate employment growth. In this context, a synthetic diffusion index is defined as the diffusion index that would be obtained by way of a survey if the sampling were extensive enough to capture the true performance of all sectors making up aggregate employment. Thus, a synthetic diffusion index is a diffusion index that is constructed using disaggregated data that are actually observed ex post.

The finding that a synthetic employment diffusion index for the Fifth District closely follows aggregate employment growth in the District arises in part because aggregate employment growth is well approximated by a formula that uses uniform weights in place of sectoral employment shares in the calculation of the aggregate series. These uniform weights can then naturally be related to the proportion of individual series that move in different directions in a diffusion index. We then show that the actual Fifth District employment diffusion index, produced using firm-level surveys carried out by the FRBR, closely tracks the corresponding ex-post diffusion index constructed using observed data. A key difference is that the survey-based Fifth District index, which proxies closely for aggregate employment growth in the Fifth District (when scaled appropriately), is published in close to real time, whereas the synthetic diffusion index may only be constructed using ex-post data that are subject to revisions up to a year after their initial release.

This article also points to some limitations of using survey-based diffusion indices to track economic changes in real time. In particular, even if a survey-based index were to exactly mimic its "true" synthetic counterpart constructed with data observed ex post, it may perform poorly in tracking the aggregate series of interest. We illustrate this point using a synthetic diffusion index constructed using sectoral data on wages in the Fifth Federal Reserve District. Specifically, we show that such an index fails to effectively track aggregate wage growth in the District. This result follows from the fact that, in the case of wages, changes over time are driven to a greater extent by the intensive margin—the percent change in wages in sectors whose wages are changing in a given month—rather than the extensive margin—the number of sectors whose wages are either increasing or decreasing in a given month. In that sense, the degree to which changes in the extensive margin contribute to changes in an aggregate series is a central consideration in the interpretation of diffusion indices.

Finally, there is a persistent need for timely economic information on U.S. states. Data at the state level are more sparse and less timely than at the national level. At the same time, more granular measures are generally more useful to local economic development practitioners, who tend to be concerned with local information, than measures for the entire Fifth District. Consequently, this article explores some of the implications of producing more localized diffusion indices specific to particular states. To gauge the potential information content of such indices, we examine the behavior of synthetic employment diffusion indices for each of the states within the Fifth Federal Reserve District (District of Columbia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia) constructed using observed data. Our findings suggest that their behavior is far from uniform across indices. Both the volatility of the growth rate in employment and the relative importance of the intensive and extensive margins differ considerably across states. Moreover, the analysis shows that the informational content of the aggregate Fifth District diffusion index would be relevant for states such as Virginia and North Carolina, but much less so for DC and West Virginia. In part, the latter result emerges because economic activity in smaller states such as West Virginia tends to be more concentrated in particular industries or sectors. Thus, in areas where the extensive margin fails to explain a large portion of the overall variation in economic activity, diffusion indices that capture economic information in a larger region do not necessarily provide information that compensates for the lack of real-time economic data on those areas.

This article is organized as follows. Section 1 describes the data used in our analysis. Section 2 reviews key aspects of how aggregate economic performance relates to economic performance at a more granular level. Section 3 then decomposes economic performance at a disaggregated level into extensive and intensive margins and uses these margins to explain the relationship between diffusion indices and aggregate growth rates. This section also highlights an example in which these measures are closely related, thus providing an underpinning for survey-based diffusion indices designed to capture changes in economic activity in real time. Section 4 highlights some limitations of diffusion indices. Section 5 explores the potential usefulness and other aspects of producing diffusion indices at a more localized level, such as an individual state, rather than an entire Federal Reserve District. Section 6 provides some concluding remarks.

1. DATA

Because diffusion indices aim to provide a sense of the direction of change, or breadth of change, in economic activity, these are most often constructed from disaggregated data such as individual survey data. Diffusion indices constructed using factor analytic methods, as in Stock and Watson (2002), in fact also share this reliance on more granular data. To assess the effectiveness of diffusion indices as real-time estimates of changes in economic activity, this article makes use of two sets of disaggregated data related to the Fifth Federal Reserve District. It also makes use of diffusion indices constructed from surveys of firms in the Fifth Federal Reserve District by the FRBR.

The first set of data is state employment by industry from the Quarterly Census of Employment and Wages (QCEW) program at the BLS. The QCEW data are derived from the quarterly tax reports submitted to state workforce agencies by employers subject to state unemployment insurance laws. Employment covered by the unemployment insurance (UI) programs represents about 97 percent of all wage and salary civilian employment in the country. The employment data are monthly, but are subject to a six-month lag in availability. For most industries, the QCEW data is available from January 1990; therefore the sample used in this analysis covers January 1990 through December 2014. The QCEW data are available at the state level for industries as granular as the six-digit North American Industry Classification (NAICS) code. We include data on the six jurisdictions covered by the Fifth Federal Reserve District (District of Columbia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia) starting at the four-digit NAICS level, subject to data availability for the full time period. To the extent that the data are not available for any industry at the four-digit level, the industry is aggregated to the three-digit NAICS code, along with all of the other industries covered in that three-digit code. This process is repeated, when necessary, by aggregating the three-digit NAICS into two-digit NAICS codes. Therefore, the final data set is a balanced panel that combines employment by state by industry at the four-, three-, and two-digit NAICS classification levels. Included are data on six regions and 868 industry/state series that are broken down as follows: Washington, D.C. (30 industries), Maryland (157 industries), North Carolina (207 industries), South Carolina (163 industries), Virginia (190 industries), and West Virginia (121 industries).

When the number of establishments in a particular industry in a county or state are too few, the BLS suppresses data in order to preserve confidentiality. Therefore, for certain four- and even three-digit NAICS classification levels, some state data are not available. When the data was combined to create three- or two-digit industries, there were monthly jumps in employment in some of the aggregated industries that represented not an increase or a decrease in employment, but the suppression (or addition) of an industry. For this reason, there were a number of outliers (positive and negative) that created considerable volatility in growth rates. We removed outliers by linearly interpolating growth rates that were above the 90th or below the 10th percentile of the distribution. The data were then seasonally adjusted in SAS using the Census Bureau's X-12 ARIMA program. The adjustment was consistent with the seasonal adjustment that the BLS uses for its Current Employment Statistics payroll employment data.

The second set of data—data on wages—also comes from the QCEW database. The sample period is the same (1990 through 2014); however, the data are available only quarterly. This article uses total wages collected in a state/industry combination over the quarter. The number of industry sectors matches that in the employment data. All manipulation of the data, such as aggregating industries, removing outliers, and controlling for seasonality, is consistent with the employment data.

Finally, the third set of data is collected through the FRBR monthly surveys of manufacturing and service sector activity across the Fifth Federal Reserve District. The survey of manufacturing firms began in 1986 but took its current monthly form in November 1993. The survey asks respondents questions about shipments of finished products, new order volumes, order backlog volumes, capacity utilization (usage of equipment), lead times of suppliers, number of employees, average work week, wages, inventories of finished goods, and expectations of capital expenditures. The services survey began in 1993 and reports on revenues, number of employees, average wages, and prices received. For retailers, the survey also includes questions on current inventory activity, big ticket sales, and shopper traffic. For this analysis, the manufacturing and services surveys are combined, and diffusion indices are developed from the questions on employment and wages. The survey data are seasonally adjusted according to the same process used to adjust the QCEW data.

There is considerable variation in the number of respondents over time in the Richmond surveys. In 1993, the number of respondents started at around 250 but then fell to a low of 82 respondents by the end of 2000. The number then rose to around 150 respondents by the middle of 2001 and stayed between 150 and 200 respondents until a large jump in 2011 that can be attributed to a consolidation of survey contacts (until 2011, separate surveys were run for the North and South Carolina and Maryland/Washington, D.C., regions). For the past few years, the number of respondents has vacillated around 200 businesses. For wages, the number of respondents jumped considerably from April to May of 1997, since May 1997 was the first month that the question on wages was asked in the manufacturing survey. It is also worth noting that over the years, some questions on the surveys were added, changed, or clarified. Finally, in March 2002, survey respondents began to be able to respond online, although many responses were still faxed and mailed. By December 2010, all responses had to be submitted online.

2. ECONOMIC ACTIVITY IN THE SMALL AND THE LARGE

Formally, diffusion indices are summary statistics of the form,

$$\mu D_t + \kappa, \tag{1}$$

where D_t is the difference between the proportion of a set of disaggregated series that increased between two dates, $t-1$ and t, and the proportion that decreased over the same period,

$$D_t = \frac{N_t^u}{N} - \frac{N_t^d}{N}, \tag{2}$$

where N is the total number of series or categories being considered, say sectors, and N_t^u and N_t^d are the number of series that increased and decreased, respectively; μ and κ are normalizing constants. In the case of the FRBR diffusion indices, $\mu = 100$ and $\kappa = 0$. Thus, index values greater than zero are interpreted as an expansion, say in employment, and negative index values are conversely interpreted as a contraction; upper and lower bounds of 100 and -100 are indicative of all sectors expanding and contracting, respectively. Observe that N_t^u/N in equation (2) also has the interpretation of an average over all categories or sectors where each sector is assigned a value of 1 if reporting an increase in activity and zero otherwise, and similarly for N_t^d/N.

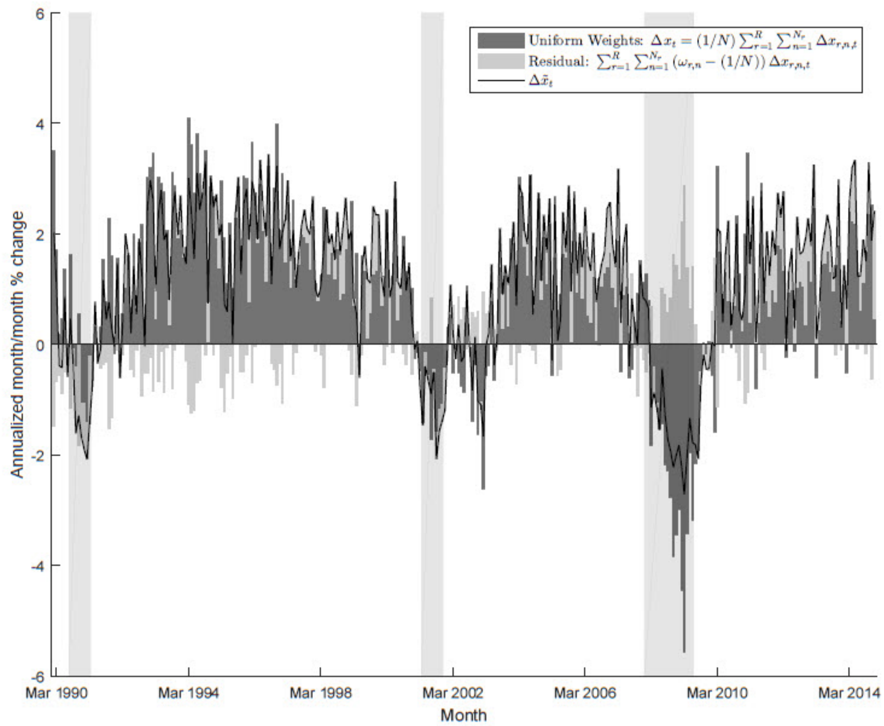
One of the simplest ways in which performance in a given area of the economy is assessed concerns the behavior of the aggregate growth rate in the corresponding variable. Thus, the behavior of aggregate employment growth over a given period, for example, gives us a sense of the performance of the labor market over that period. Aggregate employment growth in turn is a summary of employment growth at a more granular level, say employment growth in the various labor markets across all sectors that make up the aggregate series. In that sense, the estimate of overall growth in a given month hides the details of how this estimate comes about. Put another way, aggregate employment growth may come in moderately high because of a few sectors whose employment grew very rapidly while all other sectors muddled through or even declined or because employment in a wide array of sectors grew at a moderate rate. In contrast, diffusion indices give us a sense of the breadth of economic performance through a summary measure that combines the proportions of sectors whose employment increased relative to those whose employment fell. In this section and the next, we highlight important features of how these different measures of performance relate to each other.

Consider an economy composed of R regions, indexed by $r = 1, ..., R$, in which various sectors operate. There are N_r active sectors in region r, indexed by $n = 1, ..., N_r$. The total number of sectors across all regions is given by $N = \sum_{r=1}^{R} N_r$.[2] Denote employment in a given region r in sector n at date t by $x_{r,n,t}$, and its monthly annualized growth rate by $\Delta x_{n,r,t} = 1200 \times \ln(x_{r,n,t}/x_{r,n,t-1})$. We assume observations over T periods. Because our concern centers around assessing economic conditions in real time, our focus in this paper will be on the highest-frequency data available for a given series, thus monthly in the case of employment. Let $\Delta \widetilde{x}_t$ denote aggregate employment growth across all sectors and regions. Then, it follows by way of an identity that

$$\Delta \widetilde{x}_t = \sum_{r=1}^{R} \sum_{n=1}^{N_r} \omega_{r,n,t} \Delta x_{r,n,t}, \tag{3}$$

[2] For the purpose of our current analysis, the parameter values are given by $R = 6$, $N_{DC} = 30$, $N_{MD} = 157$, $N_{NC} = 207$, $N_{SC} = 163$, $N_{VA} = 190$, $N_{WV} = 121$, and $N = 868$. In the present context, it does not really matter whether sectors are region specific or not. Our analysis relies on aggregate data either at the Fifth District or state level. As we will see later, the aggregation is performed weighting each observation uniformly. Essentially, each sector-region observation is treated as an individual observation.

Figure 1 Employment Growth Rate: Uniform Weights and Residual



where $\omega_{r,n,t} = x_{r,n,t}/x_t$ are weights that, in this case, represent the employment share of a given sector in a given region at time period t. Because the time variation in employment shares is typically small, in the remainder of the paper we consider as a benchmark mean employment shares, $\omega_{r,n}$, independent of time.[3]

Since diffusion indices in (2) implicitly weight individual series uniformly, it is instructive to explore the behavior of a simple aggregate growth rate similar to that in (3) but constructed using uniform weights.[4] In particular, we can write the actual aggregate growth rate,

[3] Foerster, Sarte, and Watson (2011) follow a similar approach.

[4] Observe that N_t^u/N is the sum of series that increase between $t-1$ and t weighted by $1/N$ and similarly for N_t^d/N.

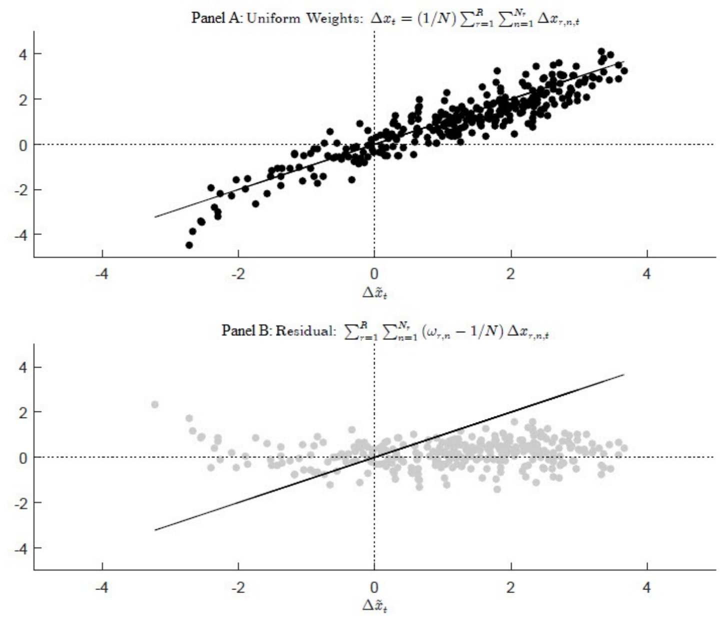
$\Delta\widetilde{x}_t$, as

$$\Delta\widetilde{x}_t = \underbrace{\frac{1}{N}\sum_{r=1}^{R}\sum_{n=1}^{N_r}\Delta x_{r,n,t}}_{\Delta x_t} + \sum_{r=1}^{R}\sum_{n=1}^{N_r}\left(\omega_{r,n} - \frac{1}{N}\right)\Delta x_{r,n,t}, \tag{4}$$

where $\Delta x_t = \frac{1}{N}\sum_{r=1}^{R}\sum_{n=1}^{N_r}\Delta x_{r,n,t}$ is an approximate growth rate computed using uniform weights, $1/N$, and $\sum_{r=1}^{R}\sum_{n=1}^{N_r}(\omega_{r,n} - \frac{1}{N})\Delta x_{r,n,t}$ is a residual that indicates the importance of deviating from actual shares in using uniform weights to arrive at an aggregate growth rate. Gabaix (2011) refers to the second term on the right-hand side of equation (4) as the granular residual. To provide intuition, in an extreme case where overall performance, $\Delta\widetilde{x}_t$, is mainly determined by a few large sectors in different regions, this second term rather than the first term in equation (4) would tend to dominate the decomposition in (4).

Figure 1 shows the decomposition in equation (4) for employment growth over time across the Fifth Federal Reserve District. By and large, the simple growth rate, Δx_t, represents a good approximation of the actual growth rate, $\Delta\widetilde{x}_t$, throughout the sample period. A notable exception concerns the period covering the Great Recession, when employment fell dramatically and where the simple growth rate and the granular residual moved in opposite directions. However, even in this case, it is the granular residual that moves in a direction opposite the actual growth rate and remains positive throughout the recession while actual aggregate growth is negative. On the whole, Figure 1 suggests that the uniform weighting of the simple growth rate, Δx_t, similar to that of the diffusion indices in (2) whereby each series receives uniform weight $1/N$ conditional on an increase or decrease, has relatively minor implications for measuring overall performance. Figure 2 provides an alternative illustration of the decomposition depicted in equation (4). Specifically, the scatter plot in Figure 2, Panel A, shows that calculations of Δx_t using uniform weights for aggregate employment growth line up closely with actual observations on $\Delta\widetilde{x}_t$ along the 45 degree line. In contrast, the scatter plot in Figure 2, Panel B, depicting the granular residual in (4) is relatively flat with respect to $\Delta\widetilde{x}_t$ around zero. In the Fifth Federal Reserve District, overall employment growth, $\Delta\widetilde{x}_t$, averages to 1.11 percent over our sample period, with the simple aggregate growth rate, Δx_t, averaging 0.94 percent over the same period and the granular residual 0.17 percent. From this point onward, therefore,

Figure 2 Employment Growth Rate: Uniform Weights and Residual



we rely on the simple growth rate, Δx_t, as our benchmark measure of aggregate performance or activity.

3. INTENSIVE AND EXTENSIVE MARGINS OF ECONOMIC ACTIVITY

In order to describe how the diffusion index in (2) for employment, say, and correspondingly aggregate employment growth in (3) are related as summaries of economic activity, let

$$\begin{aligned} \Delta x^u_{r,n,t} &= \begin{cases} \Delta x_{r,n,t} & \text{if } \Delta x_{r,n,t} \geq 0 \\ 0 & \text{otherwise} \end{cases} \\ \text{and } \Delta x^d_{r,n,t} &= \begin{cases} -\Delta x_{r,n,t} & \text{if } \Delta x_{r,n,t} < 0 \\ 0 & \text{otherwise} \end{cases}. \end{aligned} \quad (5)$$

Simply put, equation (5) distinguishes between those sectors in particular regions that contribute positively to aggregate employment growth, $\Delta x^u_{r,n,t}$ (up sectors), and those that reduce aggregate growth,

$\Delta x^d_{r,n,t}$ (down sectors). Then, following Pinto, Sarte, and Sharp (2015), and denoting we may write overall employment growth, Δx_t, in the following way,

$$\Delta x_t = \frac{N^u_t}{N}\mu^u_t - \frac{N^d_t}{N}\mu^d_t, \tag{6}$$

where

$$\mu^a_t = \sum_{r=1}^{R}\sum_{n=1}^{N_r} \Delta x^a_{r,n,t}, \quad a = u, d. \tag{7}$$

In other words, overall growth across all sectors and regions, Δx_t, may be thought of as a weighted sum of average cross-sectional growth rates, where μ^u_t and μ^d_t are the average growth rates of all sectors that add to and subtract from overall growth in a given period, respectively. The weights in (6) are the relative proportions of those sector types.

We can further express each component, $\frac{N^a_t}{N}\mu^a_t$, $a = u, d$, of Δx_t in equation (6) as

$$\frac{N^a_t}{N}\mu^a_t = \mu^a\left(\frac{N^a_t}{N} - \varphi^a\right) + \varphi^a(\mu^a_t - \mu^a) + \left(\frac{N^a_t}{N} - \varphi^a\right)(\mu^a_t - \mu^a) + \mu^a\varphi^a, \quad a = u, d, \tag{8}$$

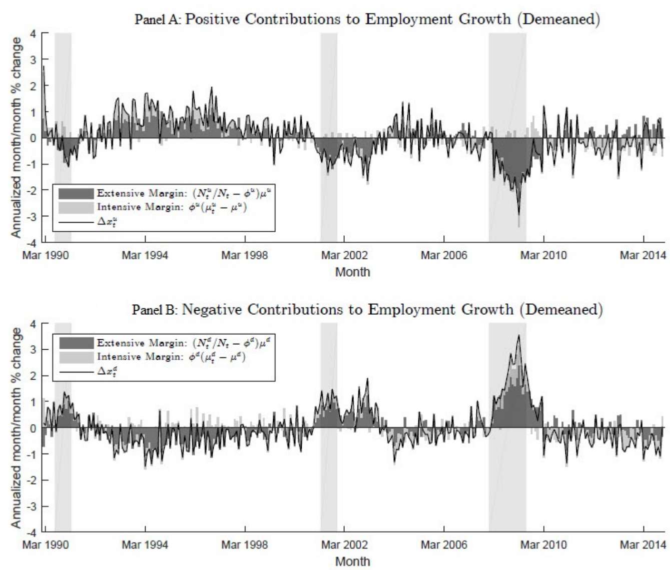
where $\mu^a = \frac{1}{T}\sum_{t=1}^{T}\mu^a_t$, $a = u, d$, are time averages, or long-run cross-sectional averages, of those sectors that contribute positively and negatively to overall growth, and $\varphi^a = \frac{1}{T}\sum_{t=1}^{T}\frac{N^a_t}{N}$, $a = u, d$ are the long-run proportions of those sectors. Thus, equation (8) tells us that, at a point in time, a large increase in overall employment growth by way of $\frac{N^u_t}{N}\mu^u_t$ may come about from the proportion of expanding sectors being higher than usual given their contribution, $\mu^u(\frac{N^u_t}{N} - \varphi^u) > 0$ corresponding to an increasing extensive margin; the cross-sectional average growth rate from those expanding sectors being higher than usual given the typical proportion of those sectors, $\varphi^u(\mu^u_t - \mu^u) > 0$ corresponding to an increasing intensive margin, or both when both are true, $(\frac{N^u_t}{N} - \varphi^u)(\mu^u_t - \mu^u) > 0$. The decline in overall growth by way of $\frac{N^d_t}{N}\mu^d_t$ may be described similarly.[5]

Combining equations (6) and (8), it follows that

$$\Delta x_t \cong \underbrace{\varphi^u(\mu^u_t - \mu^u) - \varphi^d(\mu^d_t - \mu^d)}_{\text{Change in intensive margin}} + \underbrace{\mu^u D_t,}_{\text{Change in extensive margin}} \tag{9}$$

[5] Since the total number of series or sectors is fixed in this context, we should perhaps be referring to a notion of quasi-extensive margin in the sense that entry and exit are not operative.

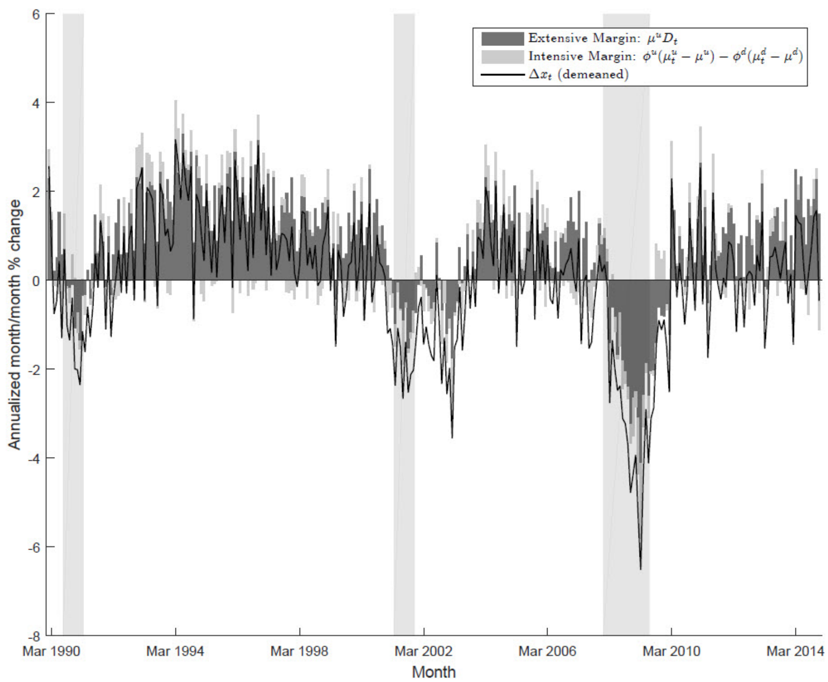
Figure 3 Employment Growth Rate Decomposition: Positive and Negative Contributions



where $D_t = (\frac{N_t^u}{N} - \frac{N_t^d}{N})$ is the difference in the proportions of sectors experiencing positive growth and negative growth respectively defined earlier. In other words, overall economic performance, as measured by an aggregate growth rate, may be interpreted as approximately arising from changes in an intensive margin, the difference between the intensity with which expanding sectors grew and that with which contracting sectors declined, and changes in an extensive margin, the difference between the fractions of sectors that expanded and contracted or the breadth of change in economic activity. The relationship in equation (9) is only approximate in the sense that μ^u and μ^d are close in practice but not necessarily identical so that (9) makes use of the fact that $\mu^d = \mu^u - \delta$ for some small δ. Moreover, the difference in the interaction between intensive and extensive margins in equation (8), $(\frac{N_t^a}{N} - \varphi^a)(\mu_t^a - \mu^a), \quad a = u, d$, may also matter at times.[6] On the whole, however, the question at hand at this point is: which of the two

[6] Refer to Pinto, Sarte, and Sharp (2015) for a complete derivation of expression (9).

Figure 4 Employment Growth Rate Decomposition: Intensive and Extensive Margins



margins in equation (9) tends to explain variations in Δx_t, if any, as a summary of overall economic performance?

Figure 3, Panel A, shows the decomposition of the positive contributions to aggregate employment growth in the Fifth Federal Reserve District, $\frac{N^u_t}{N}\mu^u_t$, in terms of intensive and extensive margins. In this case, $\mu^u = 9.23$, so that expanding sectors contribute about 9.2 percent to employment growth on average, and $\varphi^u = 0.54$, so that expanding sectors represent about 54 percent of all sectors on average. Figure 3, Panel A, makes it clear that variations in $\frac{N^u_t}{N}\mu^u_t$ are to a large degree influenced by variations in the extensive margin. Figure 3, Panel B, shows the decomposition of the negative contributions to aggregate employment growth, $\frac{N^d_t}{N}\mu^d$. As in Figure 3, Panel A, the extensive margin dominates variations in $\frac{N^d_t}{N}\mu^d$. Declining sectors represent about 46 percent of all sectors on average, $\varphi^d = 0.46$, while these sectors reduce aggregate employment growth by 9 percent on average, $\mu^d = 8.95$.

Figure 4 combines Panels A and B of Figure 3 in the manner suggested by equation (9). As expected from the behavior of the individual components $\frac{N_t^u}{N}\mu_t^u$ and $\frac{N_t^d}{N}\mu_t^d$ of Δx_t, changes in the extensive margin explain most of the variations in aggregate employment growth in the Fifth Federal Reserve District. It is interesting to note that in the period following the Great Recession, even though the intensive margin becomes positive, the employment growth rate is still negative. Our analysis reveals that this outcome arises because a large number of sectors are still experiencing a decline in employment (in other words, the extensive margin is still negative), and this effect more than compensates for the positive effect of the intensive margin on the employment growth rate. As a result, the expansion in aggregate employment in the Fifth District since 2009 is largely influenced by the behavior of the extensive margin, or the percentage of sectors experiencing an increase in employment.

The exercises above suggest that, insofar as variations in the extensive margin explain the bulk of aggregate growth in a variable of interest, diffusion indices measuring the breadth of change in economic activity, when appropriately scaled, may serve as a close indication of aggregate growth. Equation (9) makes use of the mean cross-sectional growth rate of expanding sectors, μ^u, to scale the diffusion index. As explained earlier, without much loss of generality, μ^d could also have been used since the two estimates are close. More generally, the scaling factor might be chosen so as to maximize the explanatory power of changes in the extensive margin, D_t, with respect to Δx_t based on ex-post observations. In particular, let D_t^S denote the "true" synthetic diffusion index capturing actual changes in the proportions of expanding and contracting sectors observed ex post. We might think of D_t^S as arising from a survey of firms with a large enough sample to capture the true performance of all sectors making up aggregate employment. Thus, we might then rewrite equation (9) as

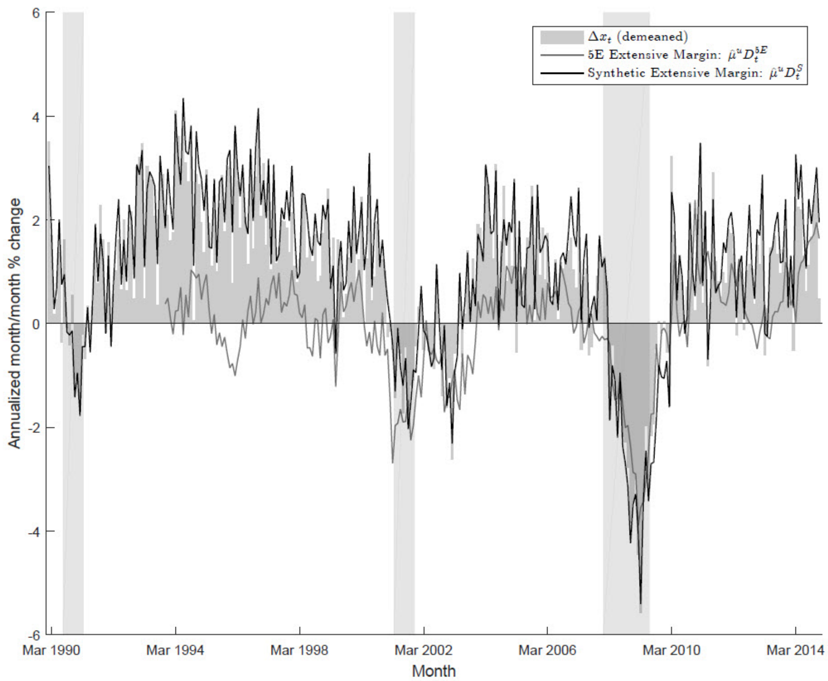
$$\Delta x_t = \underbrace{\alpha + \varepsilon_t}_{\text{Change in intensive margin}} + \underbrace{\mu D_t^S}_{\text{Change in extensive margin}}, \tag{10}$$

and choose α and μ according to a least squares criterion. This yields $\widehat{\mu} = 10.80$ instead of 9.23 used in Figure 4.

Figure 5 illustrates the behavior of the "true" synthetic diffusion index D_t^S, scaled by $\widehat{\mu}$ from (10), against the employment diffusion index produced by the FRBR for the Fifth Federal Reserve District.[7] As discussed earlier, this employment diffusion index represents the

[7] When μ is estimated using the diffusion index calculated by the FRBR instead of D_t^S in equation (10), we obtain $\widehat{\mu} = 12.10$.

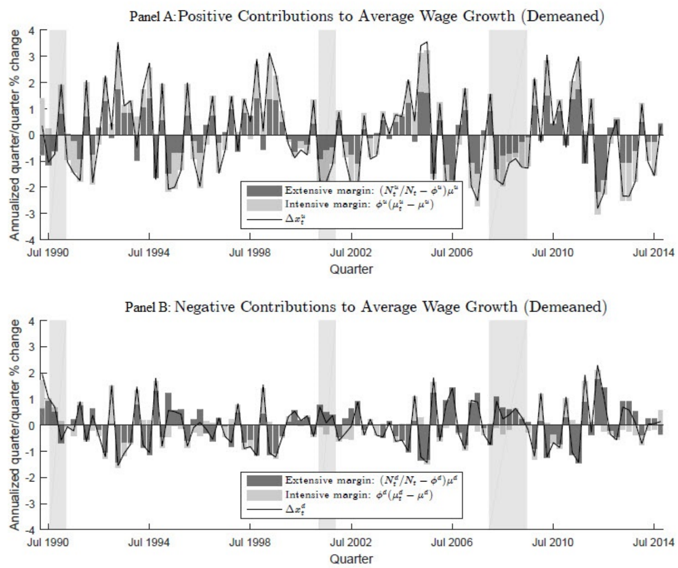
Figure 5 Employment Growth Rate: Synthetic and FRBR Extensive Margins



share of respondents to the FRBR manufacturing and service sector surveys who reported increased employment in the last month minus the share of respondents who reported decreased employment. The data for the month are collected from respondents through the third Wednesday of every month and are available publicly on the fourth Tuesday of every month; thus, they are the timeliest regional data available.

As indicated by Figure 5, the survey-based diffusion index produced in real time by the FRBR lines up remarkably closely with the synthetic diffusion index produced from employment data observed ex post. Figure 5 also shows, however, that the performance of the FRBR employment index improves over time. In the early years of the FRBR's index, from 1993 to 2001, the survey-based diffusion index and the synthetic index are somewhat far apart. There have been a few changes to the surveys over the years. As is clear from the earlier discussion, the number of respondents and the sampling of the respondents changed

Figure 6 Wage Growth Rate Decomposition: Positive and Negative Contributions



over the years due to both economic changes in the region and changes to the survey process. One such process change was that in March 2002, survey respondents began to be able to respond online, although many responses were still faxed and mailed. By December 2010, all responses had to be submitted online. Thus, with those changes, the Richmond Fed's employment diffusion index begins to track its synthetic counterpart much more closely beginning in 2002. Between June 2002 and December 2014, the correlation between the survey-based diffusion index and the synthetic index constructed from observed data is 0.77. In addition to the changes in the survey process explained earlier, other reasons, including variations in the survey composition and sample size, may also explain the shift observed in the survey series starting in 2002.[8]

[8] A more detailed analysis is required to identify such factors. We will revisit this issue in future work.

4. LIMITATIONS OF DIFFUSION INDICES

As the previous section suggests, the importance of the link between diffusion indices and aggregate growth rates hinges crucially on the relative contribution of the extensive margin of activity to overall growth. In the case of changes in employment in the Fifth Federal Reserve District, we saw that changes in the extensive margin contributed significantly to overall employment growth. There is nothing to suggest, however, that this should be the case for all aggregate series of interest. To highlight the potential limitations of diffusion indices, we consider an effort to track wage pressures in real time by way of changes in the extensive margin that keeps track of the proportion of sectors that are seeing increases and decreases in average wages.

One natural definition of an overall average wage that takes into account wages in all sectors and regions is given by

$$\widetilde{w}_t = \sum_{r=1}^{R} \sum_{n=1}^{N_r} \frac{x_{r,n}}{x} w_{r,n,t} \tag{11}$$

where $w_{r,n,t}$ is the average wage in region r in sector n at date t, and $\frac{x_{r,n}}{x}$ is the corresponding mean employment share in that region and sector. Average wage growth, $\Delta\widetilde{w}_t$, then follows approximately

$$\Delta\widetilde{w}_t = \sum_{r=1}^{R} \sum_{n=1}^{N_r} \frac{W_{r,n}}{W} \Delta w_{r,n,t}, \tag{12}$$

where $W_{r,n}$ is the (mean) total wage bill in region r in sector n, and W is the (mean) total wage bill across all sectors and regions. As in the previous section, we can decompose average wage growth in equation (12) into a uniformly weighted growth rate and a granular residual,

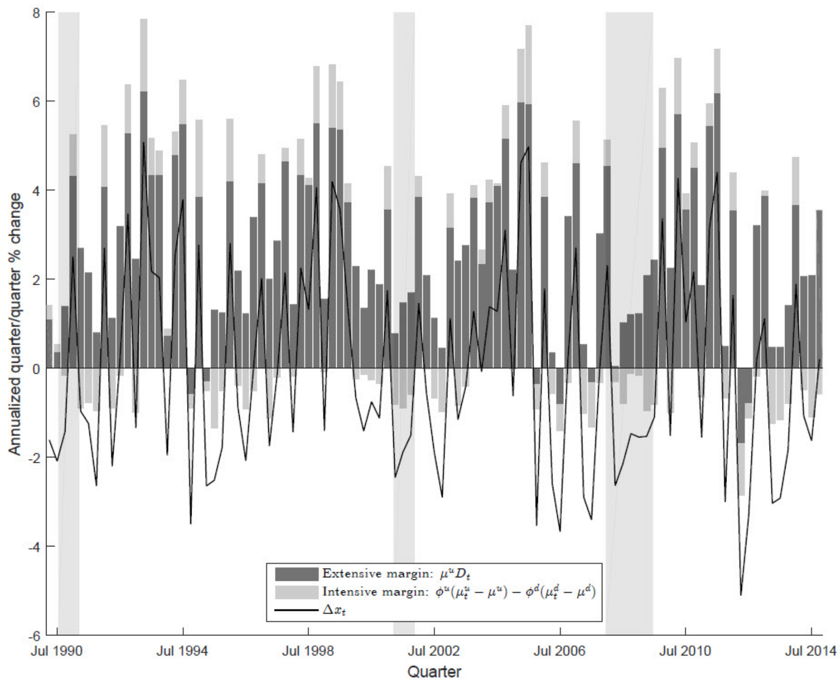
$$\Delta\widetilde{w}_t = \underbrace{\frac{1}{N} \sum_{r=1}^{R} \sum_{n=1}^{N_r} \Delta w_{r,n,t}}_{\Delta w_t} + \sum_{r=1}^{R} \sum_{n=1}^{N_r} \left(\frac{W_{r,n}}{W} - \frac{1}{N} \right) \Delta w_{r,n,t}, \tag{13}$$

and further decompose the simple average growth rate, Δw_t, into intensive and extensive margin changes,

$$\Delta w_t \cong \left[\varphi^u(\mu_t^u - \mu^u) - \varphi^d(\mu_t^d - \mu^d) \right] + \mu^u D_t. \tag{14}$$

Analogously to the decomposition of employment in the previous section, changes in the intensive margin, $[\varphi^u(\mu_t^u - \mu^u) - \varphi^d(\mu_t^d - \mu^d)]$, capture how high increasing wages are rising relative to how badly declining wages are falling in those sectors and regions where wages

Figure 7 Wage Growth Rate Decomposition: Intensive and Extensive Margins

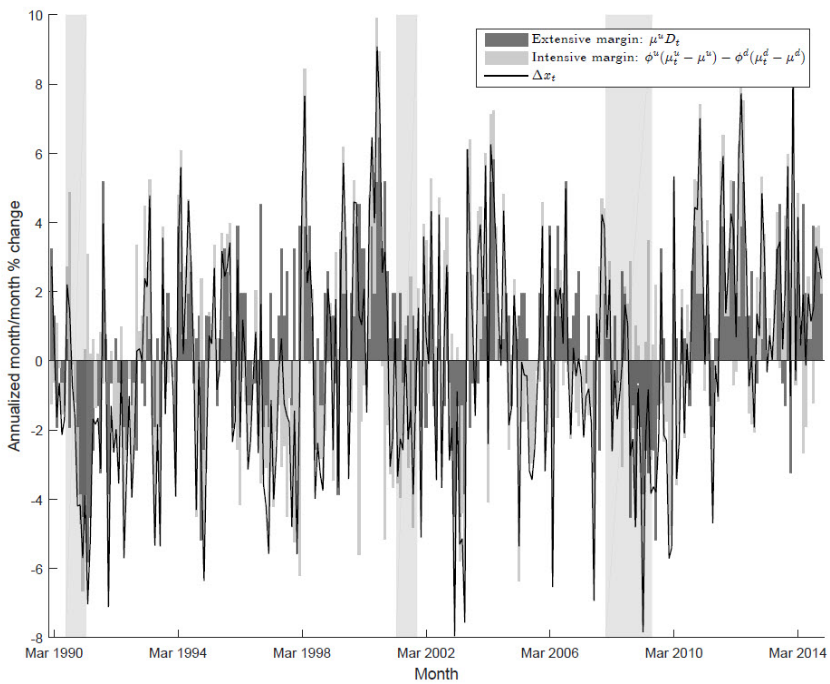


are changing. Changes in the extensive margin, $\mu^u D_t$, reflect the extent to which increasing wages are widespread across sectors relative to decreasing wages.

Figure 6 shows the decomposition of the positive and negative contributions to average wage growth into their respective intensive margins, $\varphi^a(\mu_t^a - \mu^a)$, and extensive margins, $\mu^a\left(\frac{N_t^a}{N} - \varphi^a\right)$, $a = u, d$. A salient feature of the positive contributions is that the intensive margin is at least as important at explaining $\frac{N_t^u}{N}\mu_t^u$ as the extensive margin, with periods in which the former even dominates the latter. The negative contributions to the average wage growth rate are, however, generally much smaller and mostly dominated by the extensive margin.

Figure 7 shows the overall decomposition of the growth rate in average wages into the intensive and extensive margins, as indicated by (14). From the figure, we observe that changes in the extensive margin are, with only a few exceptions, always positive. The (demeaned)

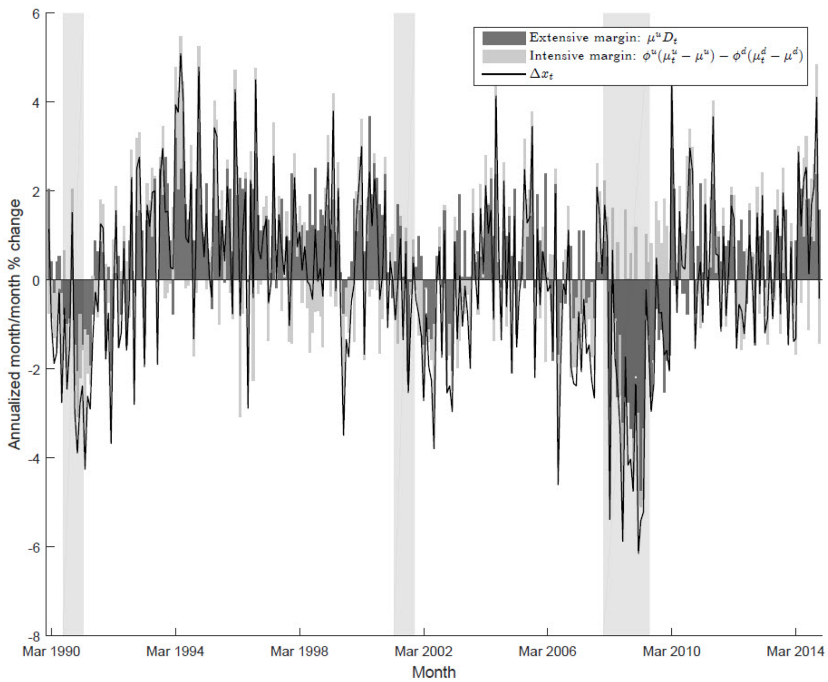
Figure 8 DC - Employment Growth Rate Decomposition: Intensive and Extensive Margins



growth rate of average wages, however, is frequently negative, generally coinciding with periods in which changes in the intensive margin are negative. The decomposition in (14) reveals that the intensive margin will tend to be negative whenever the average wage growth rate of sectors reporting an increase at time period t, μ_t^u, becomes small relative to the corresponding growth rate of those reporting a decline, μ_t^d. In our sample, the changes in μ_t^u dominate, especially in time periods in which the economic activity is low or declining. For instance, average wage growth is below its mean from the first quarter of 2008 until the second quarter of 2009, coinciding with a time period in which μ_t^u was also below its mean. The behavior of μ_t^d is, however, more erratic, with quarters in which μ_t^d was even below its mean during that same period.[9]

[9] Also, note that while the correlation between μ_t^u and Δw_t is 0.80, the correlation between μ_t^d and Δw_t is -0.14.

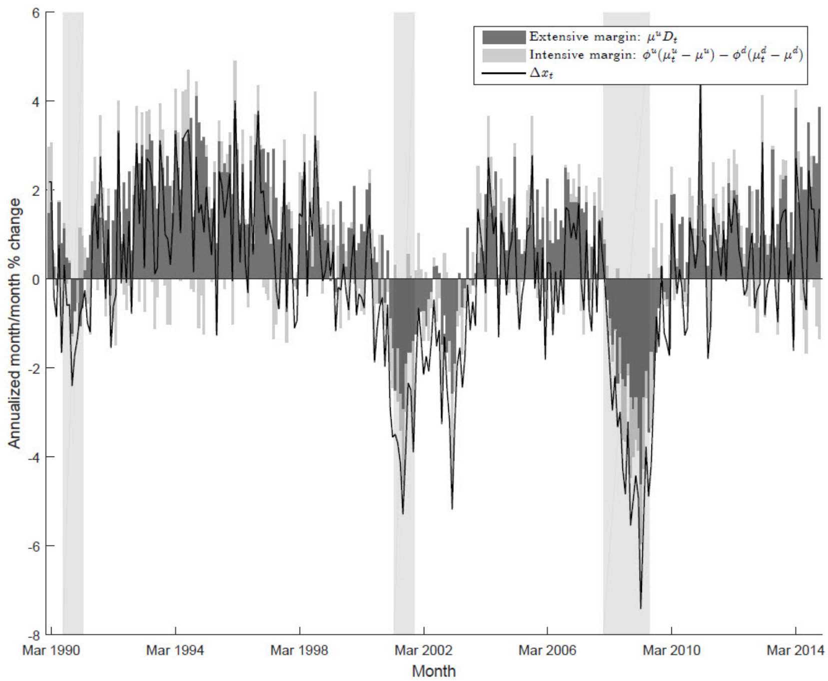
Figure 9 MD - Employment Growth Rate Decomposition: Intensive and Extensive Margins



Put another way, in the case of average wages, changes in the extensive margin are frequently at odds with the behavior of its overall growth rate, highlighting the limitations of diffusion indices as real-time indicators of economic activity. One reason that explains the weak connection between the extensive margin and the overall growth rate in average wages is that wages seldom decline in nominal terms. Other series may certainly show the same kind of pattern.[10]

[10] Further examination of the underlying factors explaining the behavior of different series of interest (specifically those series included in the FRBR survey) would allow us to determine which ones are more like the employment series, where the extensive margin plays a dominant role, and which ones share more closely the characteristics of the wage series.

Figure 10 NC - Employment Growth Rate Decomposition: Intensive and Extensive Margins

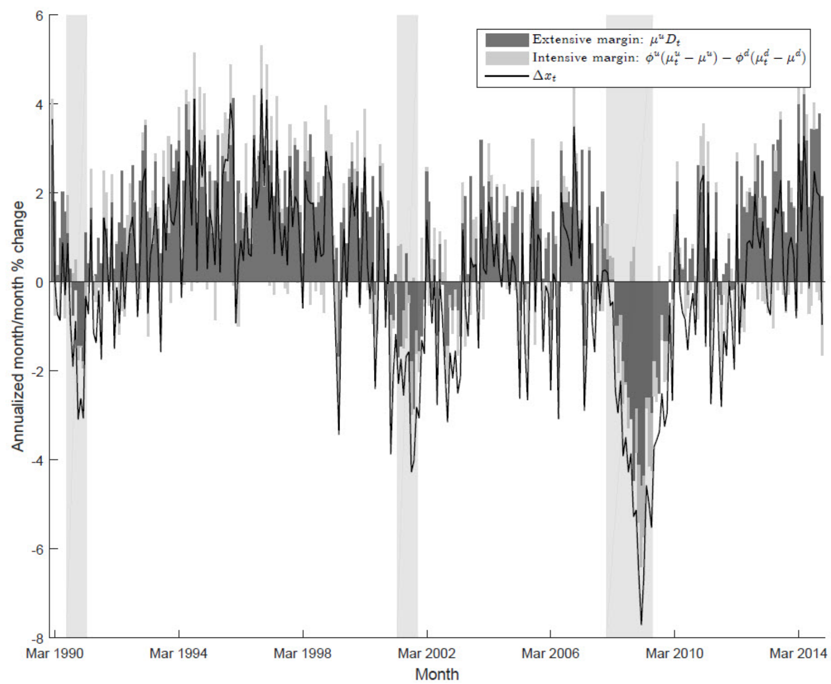


5. STATE-LEVEL DIFFUSION INDICES

Although the survey-based diffusion index for the Fifth District aids in understanding economic activity for the entire region, the dearth of data available for individual states combined with the important role that state boundaries play in economic activity and policymaking, mean that measures of activity at the state level would be more useful to many local policymakers or economic development practitioners than measures related to the Fifth Federal Reserve District. The manufacturing and service sector surveys provide information that is not otherwise available at the state level (such as new orders of manufactured goods or retail shopper traffic) in a timely fashion and include respondents' projections of future activity.

Mainly due to data limitations, the FRBR is unable to construct and report state-level diffusion indices for manufacturing and services

Figure 11 SC - Employment Growth Rate Decomposition: Intensive and Extensive Margins



separately.[11] The FRBR combines survey responses from every state and calculates a Fifth District diffusion index. However, the evolution of this aggregate indicator may not accurately track the performance of each individual state in the District. To understand the implications of conducting an analysis at the level of the District rather than individual states, we use state employment data from QCEW and apply the methodology introduced in Section 3 to each of the states in the Fifth District. Specifically, we decompose state employment growth rates into their intensive and extensive margins and construct synthetic state-level employment diffusion indices.

Figures 8 through 13 show the evolution of the employment growth rate and changes in the intensive and extensive margins for each state in the Fifth District. Table 1 presents their standard deviations, and

[11] However, the FRBR conducts a survey of general business activity for the Carolinas and for Maryland.

Table 1 Employment: Standard Deviation

	$\Delta \mathbf{x}_{r,t}$	**Intensive Margin**	**Extensive Margin**
5E	1.52	0.45	1.24
DC	3.26	2.44	2.23
MD	1.98	1.14	1.42
NC	1.92	0.81	1.55
SC	2.03	0.74	1.67
VA	1.79	0.85	1.36
WV	2.02	1.21	1.36

Table 2 the cross-correlations between the calculated synthetic diffusion indices (or extensive margins). A few remarks are worth making. First, the volatility of the employment growth rate differs considerably across states. The standard deviation of $\Delta x_{r,t}$ throughout the period under consideration is almost twice as high in DC (3.26) as it is in Virginia (1.79). Second, the relative importance of intensive and extensive margins in explaining state-level employment growth also differs considerably across states. While changes in the extensive margin explain the bulk of variations in state employment growth in North Carolina, South Carolina and Virginia, they seem much less relevant to employment growth in DC, Maryland, and West Virginia, where the intensive and extensive margins play essentially similar roles. For the states in the Fifth District, economic activity tends to be concentrated in a lower number of sectors in smaller states, with the extensive margin thus becoming relatively less important. Third, the correlation between the synthetic Fifth District diffusion index D_t^S calculated earlier and the state-level diffusion indices also differs across states, as suggested by Table 2.

Table 2 Correlation Matrix: State and Fifth District Diffusion Indices (Extensive Margin)

	5E	DC	MD	NC	SC	VA	WV
5E	1.0000						
DC	0.4450	1.0000					
MD	0.8174	0.4221	1.0000				
NC	0.8969	0.2924	0.6184	1.0000			
SC	0.8511	0.3197	0.5878	0.7580	1.0000		
VA	0.9095	0.3881	0.7364	0.7400	0.7110	1.0000	
WV	0.5443	0.1884	0.3823	0.4122	0.2816	0.4673	1.0000

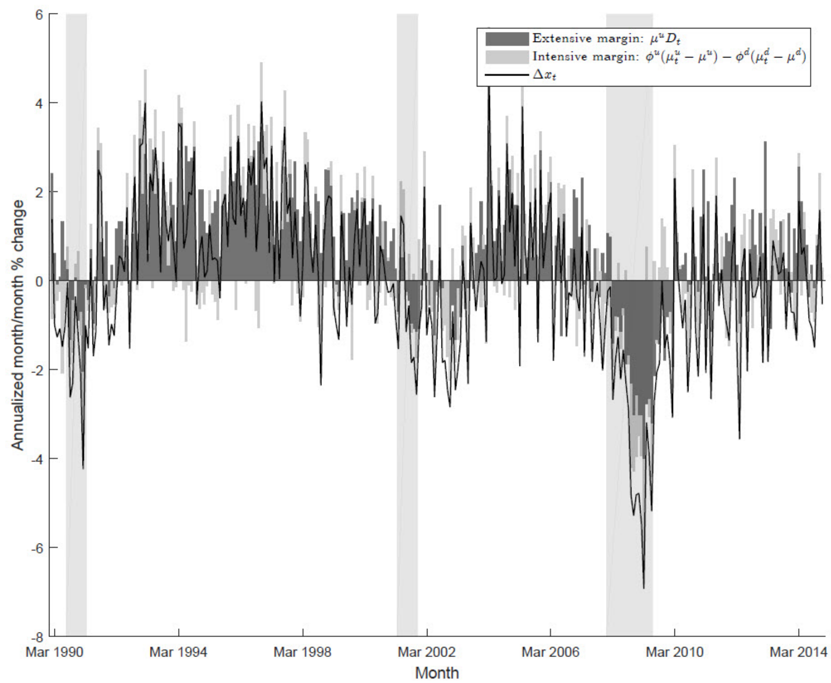
In particular, the diffusion indices for Virginia and North Carolina seem to closely follow the performance of D_t^S, with correlation coefficients of about 0.90. The correlation is also relatively high for South Carolina and Maryland (0.85 and 0.82, respectively). However, the correlations between state and Fifth District indices are much lower for DC and West Virginia (0.45 and 0.54, respectively). Thus, in regions where the extensive margin fails to explain a large component of the overall variation in economic activity, broader-based diffusion indices capturing economic information in surrounding regions do not necessarily make up for the lack of real-time information. Even though the synthetic diffusion index may not accurately represent the behavior of aggregate growth in states where economic activity is concentrated in a few sectors (such as West Virginia), an index based on a large enough sample of survey respondents may perform satisfactorily.

6. CONCLUDING REMARKS

In this article, we provide an analysis of diffusion indices that parses out the conditions under which they are likely to serve as reliable real-time indicators of economic activity. In particular, building on Pinto, Sarte, and Sharp (2015), we highlight the fact that diffusion indices, appropriately scaled, capture the contribution of changes in the extensive margin to aggregate changes in a series of interest. For the case of employment in the Fifth District, we show that changes in this margin in fact account for the bulk of changes in aggregate employment growth.

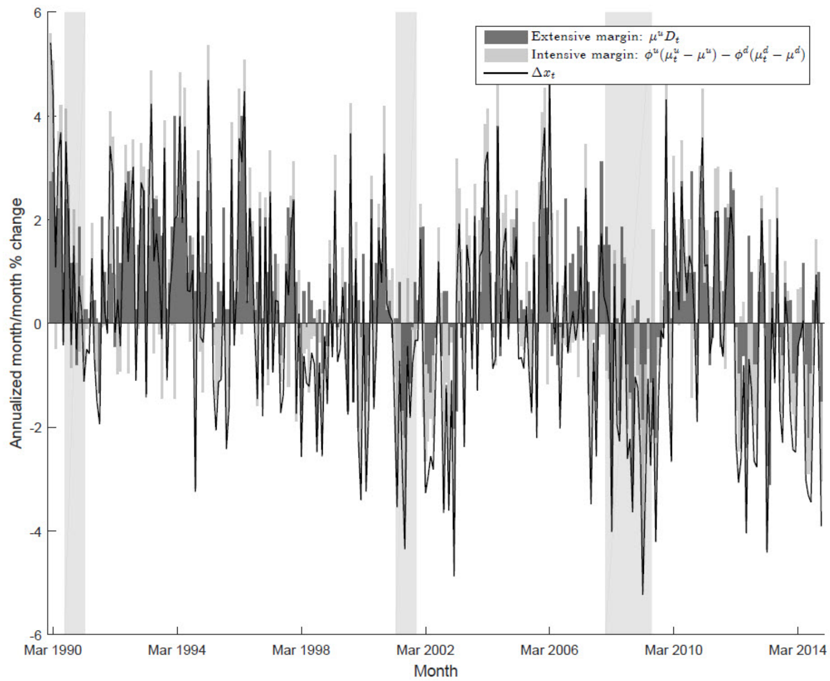
This article also highlights the potential limitations of diffusion indices. Specifically, since diffusion indices capture changes in an extensive margin, these indices are of limited usefulness in cases where aggregate changes are driven by the intensive margin. That is, the intensity with which economic activity increases in particular sectors, for

Figure 12 VA - Employment Growth Rate Decomposition: Intensive and Extensive Margins



example, rather than the number of sectors in which economic activity increases. In the case of average wages, for example, we show that changes in the extensive margin are frequently opposite that its overall growth rate. Finally, we explore the potential usefulness and other aspects of producing diffusion indices at a more localized level, such as an individual state, rather than an entire Federal Reserve District. Given that economic activity is typically more concentrated across sectors in smaller states, and changes in the extensive margin play a smaller role, relying on broader diffusion indices capturing activity in surrounding regions remains of limited use for such states.

Figure 13 WV - Employment Growth Rate Decomposition: Intensive and Extensive Margins



REFERENCES

Croushore, Dean. 2011. "Frontiers of Real-Time Data Analysis." *Journal of Economic Literature* 49 (March): 72–100.

Croushore, Dean, and Tom Stark. 2001. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics* 105 (November): 111–30.

Foerster, Andrew T., Pierre-Daniel G. Sarte, and Mark W. Watson. 2011. "Sectoral versus Aggregate Shocks: A Structural Factor Analysis of Industrial Production." *Journal of Political Economy* 119 (February): 1–38.

Gabaix, Xavier. 2011. "The Granular Origins of Aggregate Fluctuations." *Econometrica* 79 (May): 733–72.

Moore, Geoffrey H. 1983. "Why the Leading Indicators Really Do Lead." In *Business Cycles, Inflation, and Forecasting*, 2nd ed. Cambridge, Mass.: Ballinger, 339–52.

Pinto, Santiago M., Pierre-Daniel G. Sarte, and Robert Sharp. 2015. "Learning About Consumer Uncertainty from Qualitative Surveys: As Uncertain As Ever." FRB Richmond Working Paper 15-09 (August).

Stock, James H., and Mark W.Watson. 2002. "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics* 20 (April): 147–62.