Economic Quarterly—Volume 103, Numbers 1-4—First-Fourth Quarter 2017—Pages 1–35

Inequality Across and Within US Cities around the Turn of the Twenty-First Century

Felipe Schwartzman

You have just finished your PhD, and you have excellent offers. One of them is in a major city such as Washington, DC, and the other is in a smaller place like Charlottesville, Virginia, or Durham, North Carolina. In terms of the quality of each department, all offers look like excellent moves, and your advisor would be thrilled to see you in either place. It comes down to where you would rather live. If you move to the larger city, opportunities might look better down the line. There are multiple great universities in and around Washington, increasing the number of people you can interact with and learn from. International organizations in Washington, such as the IMF and the World Bank, are willing to pay high salaries to people with your qualifications, bringing up your market wage. At the same time, a large city offers unique amenities like superb restaurants and great art. On the other hand, rent is expensive: you will probably need to settle for a smaller house or a longer commute.

Suppose instead that you have never pursued a PhD. In fact, you barely graduated from high school. Your choices may look fairly different. Washington, DC, will have few, if any, good jobs for you, since there is no space for the large industrial facilities that are likely to offer good jobs for people without college degrees. You do not have much interest in the elitist art emphasized in posh neighborhoods and cer-

The views expressed in this paper are those of the author and should not necessarily be interpreted as those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

DOI: https://doi.org/10.21144/eq1030101

tainly no disposable income to go to nice restaurants. Smaller towns may offer you better prospects, and you can purchase a better house there with your wage.

In the end, irrespective of your education, all of these factors are also influenced by your own personal preferences. Durham might be closer to relatives, or you might have a special appreciation for monuments and memorials on the National Mall.[1]

Together, all these factors determine a spatial equilibrium, in which people choose where to live and wages and rental prices adjust accordingly. What recent research has shown is that the nature of this spatial equilibrium has changed in the US in the past few decades. Large cities such as Washington, DC, New York, or San Francisco are increasingly places for the skilled elites. Those cities have experienced higher wage growth, and their growth has been unequal; concentrating income among educated professionals. At the same time, much of that wage growth has been offset by increased rents. Not surprisingly, the share of college-educated workers in these cities has increased.

Those facts can be accounted for by a spatial equilibrium framework as consequences from relative increases in the demand for skilled labor by firms in large, skilled cities. An important part of the trends may have to do with adoption of computer technology. The mechanism is also most likely related to greater spillovers among those skilled workers in those cities, but the precise nature of those spillovers is still open to more research.

In what follows, I describe in greater detail the research documenting those facts and the lessons that one can derive about the underlying mechanisms. In Section 1, I lay out the facts, and in Section 2, I lay out the explanations. Section 2 includes the presentation of a canonical urban equilibrium model with two occupations that can be used to think through different mechanisms and discuss the evidence surrounding alternative hypotheses.

1. KEY FACTS

The key facts about inequality across and within US cities can be summarized as follows: if you have a college degree or, more generally, are a more skilled worker, you are more likely to live in a larger city. Wages in those cities are generally higher, although they are also offset by higher rental prices. Those relative wage gains are particularly pronounced

[1] Perhaps the answer is to go for the happy medium and come work in Richmond, Virginia!

among more highly skilled workers, making those large, skilled cities more unequal.

The literature typically identifies skills with individual characteristics correlated with productivity. High-skilled workers can be identified as ones who have high levels of education, work in high-wage occupations such as management or law, or work in high-wage industries such as professional and business services. One alternative to this vertical definition of skill is adopted by Bacolod, Blum, and Strange (2009), who examine how salaries vary with different skill dimensions such as cognitive processing or personal interaction. Alternatively, skills can be estimated from structural models (Baum-Snow and Pavan 2013; Eeckhout, Pinheiro, and Schmidheiny 2014).

In spite of this variety of measures, different papers have recently documented fairly robust facts about spatial inequality in the US around the turn of the twenty-first century:

1. Larger cities have a greater concentration of high-skilled workers.

This fact is true regardless of how one measures skill. Baum-Snow and Pavan (2013), Eeckhout et al. (2014), and Davis and Dingel (2017) provide evidence for measures of skill quality based on education and occupation. For example, Baum-Snow and Pavan (2013) report that in the late 2000s, about 40 percent of the workforce of cities in the top size decile had a college degree, whereas in cities in the bottom size decile only about 20 percent of the workforce had a college degree.

The extent to which these relationships have changed since the 1980s appears to be dependent on the exact definition of skills. Berry and Glaeser (2005) and Diamond (2016) find that cities with a high share of college graduates have experienced a larger increase in that share, but Baum-Snow, Freedman, and Pavan (forthcoming) argue that such a relationship is not apparent if skilled workers are redefined to include those with "some college." This seems to suggest that much of the change over this period has occurred due to more people finishing college. The sensitivity of changes over time to definitions is probably also a reflection of high persistence of the educational composition of cities, a fact emphasized by Beaudry, Doms, and Lewis (2010).

Such a sensitivity of time trends to the boundaries between high- and low-skilled workers appears to call for a more disaggregated view. Disaggregation reveals that in recent periods both the highest- and lowest-skilled workers tend to concentrate in large cities, with smaller cities exhibiting a more concentrated skill distribution. For example, when measuring skills by education, Eeckhout et al. (2014) find that large cities include more college graduates, but also more high school dropouts. Of the latter, many (but not all) appear to be recent

international immigrants.[2] The differences in dispersion were not always present. Davis and Dingel (2017) show evidence to the effect that this phenomenon is relatively new, with larger cities exhibiting uniformly more skilled workers in 1980. This move toward a more extreme distribution of skills in recent decades is consistent with Autor and Dorn's (2013) description of a geographic dimension to job market polarization, with cities that have a high share of workers in occupations with intermediate wage levels ("routine intensive" occupations) seeing large shifts in their labor composition toward low-wage occupations ("service" occupations).

2. Nominal wages are overall higher and increasing in larger and in more skill-intensive cities, but real wages are not necessarily.

A key distinction when interpreting geographic data is between nominal and real wages (or income, more broadly), where the latter incorporates a local price adjustment. Because not all goods consumed by households can be freely traded between cities, the law of one price does not necessarily hold everywhere. In particular, land is the ultimate nontradable good and corresponds to a large fraction of households' consumption baskets.

One of the most important stylized facts of urban economics is that larger cities exhibit higher nominal wages. Most recently, Baum-Snow et al. (forthcoming) have calculated that from 2005–07 nominal wages increase 0.065 percent for each percentage increase in city size. Relatedly, Glaeser and Maré (2001) pin the wage differences between urban and rural areas to around 33 percent in 1990. Those relationships have strengthened over time. The elasticity of wages to city size reported by Baum-Snow et al. (forthcoming) for 2005–07 is about 50 percent larger than what they report for 1980. Also, in the working paper version of their 2013 paper, Baum-Snow and Pavan point out that the wage gap between the largest cities (1.5 million people or more) and rural areas increased from 24 percent in 1980 to 33 percent in 2000. These gaps reflect in part the increasing correlation between city size and skill mix, but after controlling for those they remain sizable at 17 percent and 24 percent, respectively, and the trend remains noticeable. There is, moreover, a strengthening of the relationship between the skill intensity of a city and wages, with more skill-intensive cities exhibiting higher wages for both skilled and unskilled workers (Diamond 2016).

[2] Interestingly, they do not find any differences in average skills between cities even as they confirm the findings by others that those cities feature a higher concentration of college-educated workers. This is because college-educated workers concentrate in the higher quantiles of the skill distribution.

While the relationship between nominal wages and city size is a clear and robust fact of urban economics, this relationship does not necessarily extend itself to real wages. For a recent example, Eeckhout et al. (2014) report that there is no systematic difference in average real wages between cities. Moretti (2013) and Diamond (2016) show, moreover, that cities with a large share of skilled workers are also cities in which rent prices are higher and have increased the most in recent decades. In effect, Moretti (2013) shows that, while inequality of nominal wages across cities has clearly increased, it was met by an increased dispersion in rents, so cross-city inequality in real wages has not increased as strongly.

One important caveat to those findings is that the measurement of local price levels is itself fraught. In a recent paper, Handbury and Weinstein (2015) show that typical price indices measured to compare standards of living across cities are biased because they do not properly account for differences in the quality and variety of goods. They find that after one properly controls for those, there is a negative relationship between the price of tradable goods and city size. Given existing evidence, this would imply real wages that increase with city size.

3. The skill premium is higher and increasing in larger cities or cities with more skilled workers.

Larger cities appear to be more unequal. When comparing rural areas and the three largest metropolitan areas, Baum-Snow and Pavan (2013) report that in 2004–07, the variance of log hourly wages was 0.28 in the former and 0.53 in the latter. This, they show, is a relatively new phenomenon, since in 1979 the variances of log hourly wages for rural areas and the three largest metropolitan areas were 0.19 and 0.24, respectively.

A major focus of the recent literature has been the evolution of the skill premium across and within cities. One robust finding is that wage premia increase with city size. Eeckhout et al. (2014) and Davis and Dingel (2017) report those relationships for recent data using a variety of skill measures based on education, occupation, or observed real wages. Baum-Snow et al. (forthcoming) also point out that the relationship between skill premium and city size has become more pronounced over time, doubling in its strength over that period.

Finally, recent data also show more skilled cities exhibiting larger skill premia (Moretti 2013), although the fact does not appear to be robust to the exact definition of skill. For example, Beaudry et al. (2010) and Hendricks (2011) do not appear to find a robust relationship in recent data. There appears to be more consensus around an increasingly positive relationship between skill composition and skill premia over time. In fact, Beaudry et al. (2010) show evidence that

skilled cities had lower skill premia in 1980 and before, but this negative correlation disappeared in the early 2000s. Together, these facts point to the aggregate inequality trends as having an important geographic component, with large, skill-intensive cities leading the charge.[3]

Implications for Interpretation of Wage Inequality

The facts above suggest a reinterpretation of observed trends in wage inequality. A high wage in New York will sound appealing until one realizes how much one needs to pay for rent. Since both wages and rents are higher in large, skill-intensive cities (Fact 2), this suggests that adjusting for the local cost of living could imply less inequality in standards of living than is implied by wages alone, a point explored by Moretti (2013). He finds that the real wage differential between college and noncollege workers has increased 20 to 30 percent less than the nominal wage differentials, as rents have increased more quickly in skill-intensive cities.

At the same time, New York may offer more than smaller cities in terms of the quality of its restaurants and art scene. Diamond (2016) presents evidence that the increase in rental prices was more than matched by an increase in the amenities provided to residents of more skill-intensive cities. She estimates the effect based on a structural model similar to one we will present in Section 2 below, but in her model, local amenities change endogenously in response to the population composition. She estimates the model using measures of amenities such as quality of public schools, crime rates, and restaurant density, and she finds that once one accounts for those effects, the inequality of standards of living increases by 30 percent more than what is implied by wage inequality trends alone.

2. EXPLAINING THE FACTS

The most natural explanation for the set of facts described above, advanced by Berry and Glaeser (2005), is that the demand for skilled workers has increased more in cities that are larger and more skill intensive, while the demand for unskilled workers has not increased much

[3] Again, the direction and strength of correlations appears to depend on exactly how skill is measured. Moretti (2013) finds a very clear positive correlation between college share and college premium in 2000, whereas Beaudry et al. (2010) and Hendricks (2011) do not find a statistically significant positive correlation between a "college equivalent" share and returns to education. The main difference appears to be again in how skilled labor is defined.

anywhere. In this section, we explain why such a view is a natural fit for the data. We will then examine different theories behind that increase in demand, including the rise of computers and externalities.

The most important alternative to labor demand increases is an increase in endogenous sorting for unobserved worker characteristics, so that, for example, among college-educated workers, it is the most productive ones who choose to live in large cities such as New York and San Francisco. Nevertheless, the most recent literature appears to indicate that such sorting is unlikely to be an important driving force behind the observed facts.

In order to build the argument, we rely on a class of equilibrium models that have been originally proposed by Rosen (1979) and Roback (1982). In those models, cities exist in fixed locations and are characterized by a production technology for a fully tradable good and by their land availability. At a given wage, firms in more productive locations seek to attract more workers. However, as workers move into those cities, their demand for housing pushes rents up. Because workers are free to choose where to live, they will only choose to live in cities with high rents if wages are commensurably high. In spatial equilibrium, rents in more productive cities are just high enough to offset the productivity advantage of firms in those cities. The congestion coming from scarce land ensures that all cities are populated in equilibrium, irrespective of the productivity of their workers.

We now build a variant of such a model with workers of different skills. Similar variants have been used in recent work by Moretti (2013) and Diamond (2016), among others.

Model Setup

There are N cities, indexed $n \in \{1, ..., N\}$. Each of these cities is equipped with a production technology for a tradable good that depends on the number of high- and low-skilled workers in the city. A representative firm in city n can produce quantity Y_n of the final good by employing L_n^H high-skilled workers and L_n^L low-skilled workers according to the constant returns to scale production function:

$$Y_n = F_n(L_n^L, L_n^H).$$

Note that we allow the production function to be city-specific. Differences in the production function may also lead certain cities to produce the final good using more of one or the other type of labor. Those differences can capture "natural advantages" that can make a city more productive than another. The clearest examples of such advantages include proximity to waterways or to fertile terrain, but one could be

naturally skeptical as to whether such natural advantages directly explain the productivity differences between modern cities. In subsequent discussions on the causal mechanisms behind the facts surrounding inequality and geography, we allow for variation in capital stock and for externalities.

Also note that we assume firms are native to individual cities and stay there. This goes counter to a long-standing emphasis of the spatial economics literature on location decisions of firms. The assumption is inoccuous, however, because of our assumptions of constant returns to scale and of city-specific production technology. One could similarly postulate a model where individual firms are free to establish themselves in any city and produce using the local technology. In equilibrium, the zero-profit condition would imply the same spatial distribution of production. A less inoccuous alternative, which we do not explore, would be to allow entrepreneurs with different abilities to choose which city to live in. For an example of a framework with this feature, see Behrens, Duranton, and Robert-Nicoud (2014).

Labor markets are competitive, so firms pay wages equal to the marginal product of labor. For each skill level $k \in \{L, H\}$, this induces the labor demand equation:

$$w_n^k = \frac{\partial F_n(L_n^L, L_n^H)}{\partial L_n^k},$$

where w_n^k is the wage of workers of type k in city n in terms of the final tradable good.

Workers have preferences over the tradable good, housing, and location. For a given worker i, those idiosyncratic experiences are captured by vector worker-specific amenity parameters $\{\varepsilon_1(i), \varepsilon_2(i), ..., \varepsilon_N(i)\}$, where $\varepsilon_1(i)$ parameterizes the worker-specific preference for living in city n. Those capture the extent to which different workers have preferences for different cities. They can incorporate, for example, proximity to family or to the place where the worker grew up. A worker indexed i with skill $k \in \{L, H\}$ living in city n enjoys a utility equal to:

$$U_n^k(i) = G_n(X_n^k, H_n^k, \varepsilon_n(i)),$$

where X_n^k and H_n^k are, respectively, the amount a worker in city n consumes of the tradable good and of housing. Note again that the worker's utility function is allowed to depend on the city n where the household chooses to live. This captures the notion that amenities, such as weather, can make some cities overall more pleasant than others. Those amenity effects operate in addition to the idiosyncratic preference shifts captured by $\varepsilon_n(i)$. Understanding the origin of such

amenities and the extent to which they are endogenous has also been an important topic of research (see Albouy [2012] and Diamond [2016] for recent contributions).

As we will see, preferences for housing play an important role in the model: they introduce a source of congestion at the city level, generating a reason for population to spread across cities in spite of differences in the marginal product of workers. Similar sources of congestion could arise in the presence of other nontradable goods (such as certain traditional services) or trade costs generating home bias in consumption. Other "nonpecuniary" sources of congestion, including pollution and crime, could be captured by endogenizing the amenity parameters.

The workers' purchases of final goods and housing has to satisfy their budget constraints. We assume the only income workers receive is their wage, so that

$$X_n^k + r_n H_n^k \leq w_n^k,$$

where r_n is the rental price of housing, quoted relative to the tradable good. We assume workers can rent houses but not buy them. In a static framework such as the one presented here, this difference is mostly unconsequential. In a dynamic framework, the difference matters since the wealth of workers who purchase housing would become a function of the history of shocks to housing values in the cities where they lived. This has the potential to generate interesting effects over the wealth distribution but would be computationally challenging and has yet to be extensively explored in the literature.

Workers can freely choose in which city to live. We can solve their problem in two stages: first, we solve for the optimal choice of housing and final goods consumption given that the household lives in some city n. This induces a value function $V_n^k(r_n, w_n^k, \varepsilon_n(i))$, satisfying:

$$V_n(r_n, w_n^k, \varepsilon_n(i)) = \max G_n(X_n^k, H_n^k, \varepsilon(i)) \ s.t. \ X_n + r_n H_n \leq w_n^k.$$

Given those value functions, the household then selects as its living location the city where it attains the highest value. This induces a labor supply as a function of wages and rental prices. The number of workers of type k in city n is thus given by the fraction of workers who have a draw of the worker-specific amenity parameter $\{\varepsilon_1(i), \varepsilon_2(i), ..., \varepsilon_N(i)\}$, such that

$$L_n^k = \Pr\left[V_n(r_n, w_n^k, \varepsilon_n(i)) > \max_{n' \neq n} V_n(r_{n'}, w_{n'}^k, \varepsilon_{n'}(i))\right] \bar{L}^k,$$

where $\bar{L}^k$ is the total number of workers of type k.

The model is closed by housing market clearing conditions, implying that housing demand within each city has to be equal to housing supply within each city.

$$L_n^H H_n^H + L_n^L H_n^L = \bar{H}_n, \tag{1}$$

where $\bar{H}_n$ is the supply of housing in city n. Finally, we assume that all rental income is appropriated by absentee landlords who only have preferences for the final good and do not supply labor. This assumption ensures that the market for the tradable good clears.

Model Parameterization

In what follows we argue that Facts 1, 2, and 3 in Section 1 can be largely explained by cross-city variations in the demand for skilled labor. In particular, we specialize the model by assuming that all differences in production functions across locations stem from a skilled-labor augmenting component:

$$F_n\left(L_n^L, L_n^H\right) = \left[\left(L_n^L\right)^{\frac{\theta-1}{\theta}} + \left(\lambda_n L_n^H\right)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}, \ \theta > 1, \tag{2}$$

where θ is the elasticity of substitution between different skill levels, and λ_n is a labor-augmenting parameter specific to the labor of college-educated workers. The assumption $\theta > 1$ is consistent with common estimates of the elasticity of substitution between different types of labor.[4] Under this parameterization, labor demand functions become:

$$L_n^L = \left(w_n^L\right)^{-\theta} Y_n \tag{3}$$

$$L_n^H = (\lambda_n)^{\theta-1} \left(w_n^H\right)^{-\theta} Y_n. \tag{4}$$

We parameterize the household utility of different types of workers k living in different cities n as:

$$G_n^k(X_n^k, H_n^k, \varepsilon_n(i)) = A_n \varepsilon_n(i) \left(\frac{X_n^k}{1-\beta}\right)^{1-\beta} \left(\frac{H_n^k}{\beta}\right)^{\beta}.$$

The parameterization assumes a unit elasticity of substitution between final goods and housing. The term $A_n \varepsilon_n(i)$ captures differences in amenities between cities, with A_n incorporating amenities that affect

[4] See Ciccone and Peri (2005) for a useful summary of estimates of the elasticity of substitution between college- and noncollege-educated workers.

all workers (such as climate or quality of public schools) and $\varepsilon_n(i)$ incorporating worker-specific amenities (such as proximity to family).

Under this parameterization, the value function for a household living in city n can be written as:

$$V_n^k(r_n, w_n^k, \varepsilon_n(i)) = \varepsilon_n(i) A_n \frac{w_n^k}{(r_n)^\beta}.$$

We also assume that $\varepsilon_n(i)$ has a Fréchet distribution with shape parameter ν and is drawn independently for each city. This distribution is commonly used in trade models and has the property that it is stable under the max operator, i.e., the max of two random variables that have a Fréchet distribution is also distributed according to a Fréchet. This property makes it particularly convenient for use in aggregate models where agents make discrete choices. Furthermore, under certain conditions, it emerges naturally as the limiting distribution for the max of a sequence of random variables. It can therefore be motivated by the notion that, when considering a given city, individuals are also choosing the best of several living situations that they have available to them within that city.

As we show in the Appendix, one can then derive the labor supply function as:

$$L_n^k = \frac{\left(A_n \frac{w_n^k}{(r_n)^\beta}\right)^\nu}{\sum_{n'} \left(A_{n'} \frac{w_{n'}^k}{(r_{n'})^\beta}\right)^\nu} \bar{L}^k. \tag{5}$$

The parameter ν controls the degree of heterogeneity in tastes. This in turn governs the supply elasticity of the labor supply. A high value of ν corresponds to low heterogeneity. Thus, small variations in wages received (or rents paid) by workers in some city n imply large changes in the number of workers willing to live in that city. In particular, for the extreme case in which there is no heterogeneity in tastes, ($\nu \to \infty$), all workers have to be indifferent between all locations, so $A_n w_n^k / r_n^\beta$ is the same for all n. Conversely, a low value of ν corresponds to high heterogeneity. In that case, most workers choose where to live based entirely off their idiosyncratic preferences, and variations in the wages (or rents paid) have little bearing on the number of workers living in each city.

In the Appendix, we also show that the average utility of a household that chooses to live in city n is

$$V^k = \Gamma\left(1 - \frac{1}{\nu}\right)\left[\sum_n \left(A_n \frac{w_n^k}{(r_n)^\beta}\right)^\nu\right]^{\frac{1}{\nu}},$$

where Γ is a gamma function. Note that the average utility does not depend on the location, so that we can denote average worker welfare V^k without the city subscript. The reason is that as real wages increase in a city, workers who have lower amenity values for that city decide to live there, so that high real wage cities will include more workers with low idiosyncratic preferences for that city. Given the Fréchet distribution of tastes, the positive impact of the real wage on city welfare is exactly offset by the negative impact of worker selection.

This calculation allows us to write the labor supply condition more compactly as:

$$L_n^k = \left(\Gamma\left(1 - \frac{1}{\nu}\right)\frac{w_n^k}{V^k\,(r_n)^\beta}\right)^\nu \bar{L}^k. \tag{6}$$

Finally, we assume that housing supply does not vary across cities, $\bar{H}_n = 1$. This rules out land endowment as a key determinant of city size in the model and is consistent with the casual observation that some of the largest cities in the US, such as New York or San Francisco, are confined on relatively small land masses.[5]

Explaining the Facts with Variation in Skilled-Labor Augmenting Technology

We now show how, at least qualitatively, one can explain the facts in Section 1 entirely as a function of variations in demand. In order to do this, we make the stark assumption that $A_n = 1$ for all n, so that cities do not differ by intrinsic amenities. The only exogenous difference between cities is thus given in the skilled-labor augmenting productivity parameter λ_n. That serves as a shifter in the demand for skilled labor in different cities.

Algebraically, the way in which variation in λ_n implies Facts 1 and 2 is easiest to see in the case of homogeneous preferences ($\nu \to \infty$).[6] In that case, inspection of equation (6) implies that real wages have to

[5] Hsieh and Moretti (2017) explore the effect of land restriction regulations on city size and find that because of those restrictions, there is less concentration of population in large cities than there should be, leading to substantial output losses.

[6] See the Appendix for a derivation of the results in the general heterogeneous case.

be the same in all locations; that is, for all n where workers choose to live, it must be the case that, for $k \in \{L, H\}$ and $n \in \{1, ..., N\}$,

$$\frac{w_n^k}{r_n^\beta} = V^k. \tag{7}$$

Substituting this expression in the labor demand equations (3) and (4), we have that:

$$L_n^L = \left(V^L r_n^\beta\right)^{-\theta} Y_n,$$
$$L_n^H = (\lambda_n)^{\theta-1} \left(V^H r_n^\beta\right)^{-\theta} Y_n.$$

Taking the ratio of both labor demand functions, we have that $L_n^H / L_n^L \propto (\lambda_n)^{\theta-1}$, so that with $\theta > 1$, cities in which skilled workers are more productive will have a greater proportion of those workers. Substituting the labor demand equations into the CES production function 2, canceling out Y_n, and rearranging yields an expression for rents in each city n,

$$r_n = \left[\left(V^L\right)^{1-\theta} + (\lambda_n)^{\theta-1} \left(V^H\right)^{1-\theta}\right]^{\frac{1}{\beta}\frac{1}{\theta-1}}.$$

It follows that, along the cross-section of cities, rents increase in λ_n. From the indifference condition (7), nominal wages for both worker types must increase with rents. Thus, cities with a higher fraction of skilled workers are also cities with higher nominal wages for all workers, consistent with Fact 2. At the same time, real wages in those cities are not necessarily higher.

Optimal worker demand for housing implies they will spend a fraction β of their income on housing. Using the indifference condition (7), we can write housing demand in terms of rents only:

$$H_n^k = \beta V^k r_n^{\beta-1}.$$

Thus, from housing market clearing condition 1,

$$\left[L_n^H V^H + L_n^L V^L\right] \beta r_n^{\beta-1} = 1.$$

This last expression appears to suggest that population increases with rents, consistent with Fact 1. However, this is not necessarily true, since wages of low-skilled workers are generally smaller than those of high-skilled workers, that is, $V^L < V^H$. In the Appendix, we show that, for small variations in λ_n around a cross-city average, the population

of both types of workers increases with λ_n if and only if $\beta\theta < 1$. To see why this is necessary, note that as λ_n increases, firms will hire fewer low-skilled workers for each unit of output they produce. How much substitution occurs depends on the elasticity parameter θ. However, higher λ_n also implies higher overall output. Thus, the net effect in the demand for low-skilled workers is ambiguous. Output will increase more with productivity if rents are a small share of wage income (low β). Otherwise, workers will demand large wage increases in order to offset small increases in rent, thus limiting output variation. To summarize, L_n^L increases with λ_n if firms are not too ready to substitute between the two types of workers (i.e., if θ is low) and if wages are not too sensitive to rents (i.e., if β is low). Fortunately, those two parameters have been amply estimated. Typically, the housing share of consumption is pinned at $\beta = 1/3$, and the elasticity of substitution between college- and noncollege-educated workers is smaller than $\theta = 2$.[7] Thus, we can safely assume that $\beta\theta < 1$, so that demand for both types of employment rises with productivity of skilled workers. It follows that variations in λ_n can thus also account for Fact 1.[8]

Finally, in order to explain Fact 3, we need to depart from the model with homogeneous preferences. If workers have identical preferences for living in all cities, the wage premium has to be the same in all cities ($w_n^H/w_n^L = V^H/V^L$). Fact 3 emerges once one allows for such labor heterogeneity. Manipulating the labor demand and labor supply equations (3), (4), and (6), we thus have that

$$\left(\frac{w_n^H/V^H}{w_n^L/V^L}\right)^{\nu} = \frac{L_n^H}{L_n^L} = \left(\frac{w_n^H}{w_n^L}\right)^{-\theta} (\lambda_n)^{\theta-1},$$

where the first equation we obtain from labor supply and the second from labor demand. Combining the two equations then yields the wage premium:

$$\frac{w_n^H}{w_n^L} = \left(\frac{V^H}{V^L}\right)^{\frac{\nu}{\nu+\theta}} (\lambda_n)^{\frac{\theta-1}{\nu+\theta}}.$$

It follows that, so long as there is some heterogeneity in preferences ($\nu < \infty$) and $\theta > 1$, the wage premium increases with λ_n.

[7] See Ciccone and Peri (2005) for estimates of θ. A value of β close to 1/3 has been used by Hsieh and Moretti (2017) among others.

[8] One interesting question is what happens if there are multiple tradable industries with different intensities of use for different workers. Then substitution at the city level may be higher than at the firm level or in the aggregate, since it can operate through changes in the industrial composition.

Sources of Variation in Demand for Skilled Labor

The recent literature has identified two main sources of variation in the demand for skilled labor: differential biased technical progress, expressed by differences in adoption of computing technology, and externalities. The first source has been emphasized by Beaudry et al. (2010) and Autor and Dorn (2013), whereas the second has been emphasized by Baum-Snow et al. (forthcoming). We discuss those two in turn.[9] A third source of variation in the demand for skilled labor that has been less explored in the literature comes from the industrial composition of cities and, in particular, the importance of the business services sector (Hendricks 2011). We discuss this third source last.

Computers

One explanation for the increase in the productivity of skilled workers is the adoption of computing technology (Krusell et al. 2000). This is a natural hypothesis, since computers have become cheaper in the same period in which wage inequality has increased. At the same time, in that same period, many of the trends associated with Facts 1, 2, and 3 have been in play. This has motivated Beaudry et al. (2010) and Autor and Dorn (2013) to propose computerization as an explanation for cross-city variation in wage inequality trends.

To see the role of computer adoption, extend the model to allow the production function to include three inputs, the third of which is capital:

$$\begin{aligned} F\left(L_n^L, L_n^H, K_n\right) &= \left[\left(L_n^L\right)^{\frac{\theta-1}{\theta}} + \left(\left(\frac{1}{\alpha}K_n\right)^{\alpha}\left(\frac{1}{1-\alpha}L_n^H\right)^{1-\alpha}\right)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}, \\ \theta &> 1 \end{aligned}$$

Note that capital is complementary to skilled labor but not to unskilled labor. In the spatial setting, Autor and Dorn (2013) have motivated this complementarity through a task-based approach. Computers are particularly adept at enhancing the ability of workers performing abstract tasks involving creativity, coordination, and problem solving, whereas automation can serve as a substitute for workers performing routine tasks such as bookkeeping, clerical work, and repetitive production.

[9] Giannone (2017) is a recent contribution analyzing the implications of biased technical progress and local externalities in a common framework.

Suppose computers are a perfectly tradable good, which can be purchased in any city n at the same price p. The price should be understood to incorporate the "quality-adjusted" cost of computers, so that its time variation would include all the large gains in computational power in the past few decades.[10] Then cost minimization implies that $K_n = \frac{\alpha}{1-\alpha}\frac{w_n^H}{p}L_n^H$, so that substituting into the production function we have:

$$Y_n = \left[\left(L_n^L\right)^{\frac{\theta-1}{\theta}} + \left(\frac{1}{1-\alpha}\left(\frac{w_n^H}{p}\right)^{\alpha} L_n^H\right)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}.$$

The resulting function on the right-hand side is isomorphic to the one in our basic model, with the labor-augmenting technology parameter substituted for an increasing function of the relative cost of skilled labor and computers ($\lambda_n = \frac{1}{1-\alpha}\left(w_n^H/p\right)^{\alpha}$). As the relative price of computers decreases, firms use more of those, increasing the relative productivity of skilled workers.

With homogeneous preferences for location ($\nu \to \infty$), we can rewrite the production function in terms of rents:

$$Y_n = \left[\left(L_n^L\right)^{\frac{\theta-1}{\theta}} + \left(\frac{1}{1-\alpha}\left(\frac{(r_n)^{\beta}}{p}V^H\right)^{\alpha} L_n^H\right)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}.$$

Cities with higher rents are also cities where wages are higher and where firms have the most incentives to adopt computers. This, in turn, generates a correlation between rents and the labor mix of cities. In the presence of heterogeneity in preferences for different cities ($\nu < \infty$), cities with higher rents will also feature more wage inequality. Of course, once we endogenize λ_n, we need to have another source of exogenous variation in order to explain differences in rents across cities. One possibility is to still allow for some exogenous variation in the skill premium, possibly due to the location of universities or to allow for citywide productivity differentials.

In order to explain the observed trends, Beaudry et al. (2010) examine an environment in which they can explain observed trends just from the reduction in the price of computers (p). In their environment, apart from the three factors of production, firms have the option to

[10] Suppose a unit of an input purchased at a price $\tilde{p}$ produces z units of output according to a linear technology. Then one can define the quality-adjusted value of the input by $p = \tilde{p}/z$. In our example, the output would be computational power, and z would capture the increase in quality of computers.

choose between technologies with higher or lower intensity in skilled labor. As computers become cheaper, the incentive to use a technology that is more intensive in skilled labor increases.

Beaudry et al. (2010) also show, however, that this incentive is stronger in cities where the supply of skilled labor is higher to begin with.[11] This generates a correlation between the adoption of computers and the cross-city supply of skilled labor before computers were adopted, which they verify in the data. Intriguingly, they also show that before computers became an important part of production, the cross-city correlation between skill composition and skill premium was negative, as predicted by our model if there are no important differences in λ_n across cities but there are differences in relative labor supply. As time has progressed and computer adoption has increased, they find that correlation flipping and becoming insignificantly different from zero. One limitation of their explanation for the observed trends is that it cannot account for the positive correlation between skill composition and wage premia observed in recent data (or with other definitions of skill). In their framework, as the price of computers declines, all cities adopt the skill-intensive technology.

Externalities

An alternative explanation for the differences in demand for skilled labor in different cities involves the differential impact of local externalities on different types of workers. Moretti (2004) separates relevant theories into those involving learning and those involving labor markets.[12] Learning-based theories emphasize the role of geographic proximity in facilitating the transfer of knowledge.[13] High-skilled workers perform tasks that are more knowledge-intensive, so they would stand to benefit more from those transfers. Matching theories exploit pecuniary externalities that emerge in imperfect labor markets. A deeper

[11] In terms of our model, this would involve extending the utility function to contain a city/skill-specific effect, so that

$$G_n^k(X_n^k, H_n^k, \varepsilon_n(i)) = A_n^k \varepsilon_n(i) \left(\frac{X_n^k}{1-\beta}\right)^{1-\beta} \left(\frac{H_n^k}{\beta}\right)^{\beta}.$$

[12] See Duranton and Puga (2004) for a detailed overview of theories of agglomeration externalities more generally, including the role of sharing, matching, and learning.

[13] In Marshall's (1890) words "Great are the advantages that people following the same trade get from near neighborhood to one another: the mysteries of the trade become no mysteries; but are, as it were, in the air." Lucas (1988) follows up by stating that "Most of what we know we learn from other people (. . .) most of it we get for free. We know that this kind of external effect is common to all arts and sciences – the 'creative professions.'"

pool of skilled workers provides an environment where firms can use them more efficiently.[14]

There is an ample empirical literature pointing to human capital externalities as a source of differences in wages across cities (see Moretti [2004] for a review). This literature focuses on the effects of the skill composition of the workforce. Most recently, the role of externalities in explaining observed trends has been examined by Baum-Snow et al. (forthcoming). In their paper, they focus on what the urban economics literature has called "agglomeration externalities," i.e., external effects associated with overall city size. They allow for city size to affect the productivity of skilled and unskilled labor separately (and also of capital, which they allow for explicitly). They find a robust significant increase over time in the positive agglomeration effect on skilled labor but do not find a robust change in the agglomeration (or congestion) effect on unskilled labor. In terms of our model, they find $\lambda_n = \left(L_n^H + L_n^L\right)^\mu$, where $\mu > 0$.[15] They state that such a change in agglomeration externalities can account for 80 percent of the more rapid increase in wage inequality in large cities, with capital accumulation, the other leading alternative discussed above, accounting for less than 20 percent.

Recent work has also tried to disentangle the two sources of local spillovers. Learning and matching theories have different predictions at the microeconomic level. Learning theories imply that individual productivity is likely to increase with the time that individuals spend in a city and that productivity gains will be embodied in workers who will carry those gains with them if they immigrate. In contrast, matching externalities are unlikely to change with the time a worker has spent in a city and remain specific to the city. The evidence is consistent with learning among the high-skilled workers playing an important role, with Glaeser and Maré (2001), Baum-Snow and Pavan (2013), and De La Roca and Puga (2017) pointing out that high-skilled workers tend to experience faster wage increases when they live in large cities.[16]

Understanding why local externalities have become stronger for skilled workers is challenging. One interesting mechanism is explored by Michaels, Rauch, and Redding (2017). Investigating changes over long periods of time and using a detailed breakdown of occupations into a variety of tasks, they document a rising concentration of occupation-intensive "interactive tasks" in cities. Those tasks are ones "concerned

[14] For example, it could reduce the importance of holdout problems, allowing for more investment in skill-complementary capital (Acemoglu 1996).

[15] In reality, they are measuring changes in μ. All the comparative statics presented before also hold for changes in λ_n.

[16] See also Davis and Dingel (2016) for a microfounded spatial model with knowledge spillovers that captures most of the facts above.

with thought, communication, and inter-social activity" and are typically associated with highly skilled occupations such as nurses, accountants, and statisticians. In their model, those tasks tend to become concentrated as the cost of trading them across cities decreases and large cities are able to exploit their comparative advantage in those tasks.

Industrial and Functional Composition

One potential source of variation in demand for skilled labor is variation in industry composition of cities. One extreme case, which is nested in our model, has each city specialize in a single industry. Since industries may have different skill intensities, such industrial specialization would imply cross-city variation in λ_n. For example, Brinkman (2014) points to the increased concentration of certain skill-intensive industries, such as finance, in large cities as an important factor explaining some of the trends discussed. However, as argued by Hendricks (2011), cross-city variation in industrial composition accounts for only a small fraction of cross-city variation in skill composition. Moreover, he shows that most of the variation in skill composition across cities can be tied to variation in a high-skill-specific productivity component that is common to all industries.

At the same time, Hendricks (2011) also finds that cities with a high fraction of skilled workers are also ones with large business services sectors. He proposes a model where, like computers, the output of those services is complementary to high-skilled labor. When external accountants become cheaper, firms hire those to work with their internal staff rather than as a substitute. In terms of our model, cities where business services are cheaper would thus have higher λ_n. He endogenizes the variation in business services productivity by allowing for increasing returns to scale in that sector. One advantage of this focus on the business services sector over a focus on learning or matching externalities is that it provides a mechanism through which the external effects on the employment of skilled workers spread somewhat uniformly over a fairly heterogeneous set of sectors, in consonance with Hendricks's (2011) data analysis.

A special role for the business services sector is shared by Duranton and Puga (2004). They show that this is associated with an increasing specialization of cities by function, meaning that firms have increasingly concentrated executives and managers in larger cities and production activities in smaller ones. They therefore develop a model in which there are gains to concentrating management in cities with a high concentration of business services. While they do not explicitly tie their model to facts about the skill composition of cities and their wage

differentials, it is natural to map workers active in their management function as relatively high-skilled workers and those active in production as relatively low skilled. Given this mapping, their model could also deliver a concentration of high-skilled workers in large cities, as in the data.

Sorting for Unmeasured Skill

Differences in labor demand need not be the only source of the systematic differences in skill composition, wages, and skill premia across cities. An alternative explanation relies on sorting for unobserved worker characteristics. To see how that works, suppose there are multiple types of workers rather than just two, and assume for simplicity that they are all perfect substitutes in production, so that[17]

$$Y_n = \sum_k \mu^k L_n^k.$$

Note that now the production function is the same for all cities. The parameters μ^k capture the marginal product of workers of type k in any city. Since the different types of workers are perfect substitutes, for any worker type k in any city n, it will be the case that $w_n^k = \mu^k$. Observed wage variation across cities can occur if different types of workers are combined in common bins over the course of empirical analysis. For example, a broad "college-educated" group of workers might include workers with some college, with four-year college degrees, and with postgraduate degrees; within each of these bins, workers may have heterogeneous innate abilities. One can then explain many of the empirical facts described in the literature through such sorting (for a detailed exposition see, for example, Davis and Dingel [2017]).

To see how such an effect of skill sorting could come about, reparameterize the utility function of individual workers to:

$$G_n^k(X_n^k, \varepsilon(i)) = \begin{cases} A_n \varepsilon(i) X_n^k \ if \ H_n^k \geq 1 \\ -\infty \ otherwise \end{cases},$$

that is, a worker requires a minimal amount of housing to survive, which we normalize to 1, but beyond that how much housing they purchase is irrelevant to their utility. This is an extreme form of nonhomotheticity

[17] There are other ways of generating such sorting. See Eeckhout et al. (2014) for an alternative sorting mechanism based on complementarities between different types of workers in nonhomothetic production technology.

in preferences, with housing considered a "necessity." Note that, in order to show that in this setup it is possible to account for many of the key facts as stemming from differences in labor supply, we now allow for an exogenous city-specific variation in amenities, captured by A_n. Under this parameterization, the value for a worker of type k of living in city n is

$$V_n^k\left(w_n^k, r_n, \varepsilon(i)\right) = \varepsilon(i) A_n \left(w_n^k - r_n\right).$$

Given the Fréchet distribution of $\varepsilon(i)$, the fraction of workers of type k living in city n is

$$\frac{L_n^k}{\bar{L}^k} = \frac{\left(A_n\left(w_n^k - r_n\right)\right)^\nu}{\sum_{n'}\left(A_{n'}\left(w_{n'}^k - r_{n'}\right)\right)^\nu}.$$

As before, the fraction increases with wages and decreases with rents. The difference is that now it increases more rapidly with wages when wages are smaller. Substituting in the labor demand condition $w_n^k = \mu^k$ yields:

$$\frac{L_n^k}{\bar{L}^k} = \frac{\left(A_n\left(\mu^k - r_n\right)\right)^\nu}{\sum_{n'}\left(A_{n'}\left(\mu^k - r_{n'}\right)\right)^\nu}.$$

As rents in a city increase, the labor composition shifts toward workers with higher productivity. Finally, note that since all households buy one unit of housing, housing market equilibrium implies that:

$$\sum_k L_n^k = \bar{H}.$$

Since labor supply is increasing in the city-specific amenity and decreasing in rents, it follows that rents have to increase with the city-specific amenity. Thus, cities with higher amenities also have higher rents and a more skilled workforce. If skill differences are not directly observable, this could translate into higher measured wages for each skill level. If skill differences are particularly hard to observe among the most highly skilled workers, this sorting would translate into a higher wage premium for those cities.

The mechanism described here could be criticized on a priori grounds, since it relies on assuming that housing is a necessity. As a matter of fact, while the share of expenditures on shelter does decrease with income, it decreases by relatively little, with higher-income households spending more dollars. Furthermore, one implication of the sorting mechanism is that more-skilled workers are less sensitive to city-specific

prices. This is contrary to much evidence that finds high-skilled workers to be more mobile.[18]

While such a priori criticism could potentially be accommodated with suitable changes to the model without abandoning some of the key insights, more direct empirical assessments have not been favorable to sorting for unobservable characteristics. A first approach, adopted by Combes, Duranton, and Gobillon (2008), is to assess the effects of sorting on city wage differences through worker fixed effects. Those exercises rely on changes in the wages of workers who migrate between cities to separate the effects of sorting from other city characteristics.[19] While this first approach indicates a large role for sorting, it could underestimate or overestimate those effects to the extent that changes in the wages of migrants are not representative of the differences in wages of the overall population of cities. For example, if individuals migrate from small cities to large cities only when they receive particularly good wage offers, this selection effect would lead to an overestimation of the relationship between city size and wages. On the other hand, if workers move to large cities when they are young, take advantage of local learning opportunities to slowly become better workers, and then move back to smaller cities carrying their new abilities with them, such evidence would tend to underestimate the contribution of large cities to worker productivity since wages would change little at the time of moving. Baum-Snow and Pavan (2012) and De La Roca and Puga (2017) specifically control for these effects and find that allowing for learning within cities is especially important. Once that learning is explicitly allowed for, selection for unobservable skills ceases to be an important source of bias in estimating the effect of city size on productivity.

3. CONCLUSION

Inequality in the United States has an important spatial component. More-skilled workers tend to reside in larger cities where they earn higher wages. In the meantime, less-skilled workers make lower wages even when they live in those cities. Those relationships appear to have become more pronounced as inequality has increased. The evidence points to externalities among high-skilled workers as a significant contributor to those patterns. This suggests that policy may face an equity efficiency trade-off. The presence of positive externalities among

[18] See Notowidigdo (2011) for a discussion of how this fact can be rationalized by allowing for transfers that occur to all households irrespective of their income. This is the mirror image of the mechanism discussed here.

[19] See Combes and Gobillon (2015) for a thorough discussion of these and other econometric issues involved in estimating the impact of city size on wages.

high-skilled workers would imply gains to policies that incentivize them to become more concentrated. However, this would tend to increase inequality across and within cities. As pointed out by Fajgelbaum and Gaubert (2018), who discuss these trade-offs in detail, the inability to separate the place one lives from the place one works may lead a utilitarian planner to choose to redistribute income across cities so as to compensate workers who have high amenity value of living in low-productivity cities. While the aforementioned paper provides a substantial first step in that direction, the literature thoroughly exploring these trade-offs from various angles is still in its infancy.

APPENDIX

Appendix 1: Derivation of Labor Supply and Average Utility with Frechet Distributed Preferences

Under a Fréchet distribution, ε_n has a support between zero and infinity with c.d.f. ε_n is $e^{-\varepsilon_n^{-\nu}}$, and the p.d.f. is $\nu\varepsilon_n^{-\nu-1}e^{-\varepsilon_n^{-\nu}}$. Note that in order for any value x to be smaller than the max of several variables, it must be the case that it is smaller than each one of them. Given that individual draws of ε_n are independent, we can calculate this probability by multiplying the individual c.d.f.'s. Specifically, for any given $\hat{\varepsilon}_n$.

$$\begin{aligned}
\Pr\left[\hat{\varepsilon}_n w_n^k/(r_n)^\beta \geq \max_{n'\neq n} \varepsilon_{n'} w_{n'}^k/(r_{n'})^\beta\right] &= \prod_{n'\neq n} \Pr\left[\varepsilon_{n'}(w_{n'})^\beta/r_{n'}^k \leq (w_n)^\beta/r_n^k\hat{\varepsilon}_n\right] \\
&= \prod_{n'\neq n} \Pr\left[\varepsilon_{n'} \leq \frac{(r_{n'})^\beta/w_{n'}^k}{(r_n)^\beta/w_n^k}\hat{\varepsilon}_n\right] \\
&= \prod_{n'\neq n} e^{-\left(\frac{(r_{n'})^\beta/w_{n'}^k}{(r_n)^\beta/w_n^k}\hat{\varepsilon}_n\right)^{-\nu}} \\
&= e^{-\Phi_n\hat{\varepsilon}_n^{-\nu}}
\end{aligned}$$

where $\Phi_n \equiv \frac{1}{\left(w_n^k/(r_n)^\beta\right)^\nu}\sum_{n'\neq n}\left(w_{n'}^k/(r_{n'})^\beta\right)^\nu$. To find Pr $\left[\varepsilon_n w_n^k/(r_n)^\beta \geq \max_{n'\neq n}\varepsilon_{n'}w_{n'}^k/(r_{n'})^\beta\right]$ integrate over all $\hat{\varepsilon}_n$, i.e., calculate

$$\int_0^\infty e^{-\Phi_n\varepsilon^{-\nu}}dF(\varepsilon_n) = \frac{1}{\Phi}\int_0^\infty \Phi\nu\varepsilon_n^{-\nu-1}e^{-\Phi\varepsilon_n^{-\nu}}d\varepsilon_n,$$

where now $\Phi \equiv \frac{\sum_{n'}\left(w_{n'}^k/(r_{n'})^\beta\right)^\nu}{\left(w_n^k/(r_n)^\beta\right)^\nu}$. The integrand is equal to the pdf of a Fréchet with scale parameter Φ and which must therefore integrate to 1. It follows that

$$\int_0^\infty e^{-\Phi_n\varepsilon^{-\nu}}dF(\varepsilon_n) = \Phi^{-1} = \frac{\left(w_n^k/(r_n)^\beta\right)^\nu}{\sum_{n'}\left(w_{n'}^k/(r_{n'})^\beta\right)^\nu}.$$

We can also derive average utility in each city. This is given by $E\left[\varepsilon_n \frac{w_n^k}{r_n^\beta} | \varepsilon_n w_n^k/(r_n)^\beta \geq \max_{n' \neq n} \varepsilon_{n'} w_{n'}^k/(r_{n'})^\beta\right]$. Again, we can write this as

$$\begin{aligned}
\int_0^\infty E\left[\varepsilon_n \frac{w_n^k}{r_n^\beta} | \varepsilon_n \geq \max_{n' \neq n} \varepsilon_{n'} \frac{w_{n'}^k/(r_{n'})^\beta}{w_n^k/(r_n)^\beta}\right] dF(\hat{\varepsilon}_n) &= \int_0^\infty \Phi \varepsilon_n \frac{w_n^k}{(r_n)^\beta} \nu \varepsilon_n^{-\nu-1} e^{-\Phi \varepsilon_n^{-\nu}} d\varepsilon_n \\
&= \frac{w_n^k}{(r_n)^\beta} \int_0^\infty \varepsilon_n \Phi \nu \varepsilon_n^{-\nu-1} e^{-\Phi \varepsilon_n^{-\nu}} d\varepsilon_n \\
&= \frac{w_n^k}{(r_n)^\beta} \Phi^{\frac{1}{\nu}} \Gamma\left(1 - \frac{1}{\nu}\right) \\
&= \left[\sum_{n'} \left(w_{n'}^k / (r_{n'})^\beta\right)^\nu\right]^{\frac{1}{\nu}},
\end{aligned}$$

where the first equality follows from the definition of conditional expectation, and the third equality follows from the fact that the integrand is the expected value of a Fréchet with scale parameter $\Phi^{-\nu}$.

Appendix 2: Comparative Statics with Heterogeneous Preferences

The model can be summarized by the following system of equations:

$$Y_n = \left[\left(L_n^L\right)^{\frac{\theta-1}{\theta}} + \left(\lambda_n L_n^H\right)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}, \ \theta > 1,$$

$$\begin{aligned}
L_n^L &= \left(w_n^L\right)^{-\theta} Y_n, \\
L_n^H &= (\lambda_n)^{\theta-1} \left(w_n^H\right)^{-\theta} Y_n,
\end{aligned}$$

$$L_n^k = \left(\frac{w_n^k}{V^k (r_n)^\beta}\right)^\nu \bar{L}^k$$

$$V^k = \left[\sum_{n'} \left(\frac{w_{n'}^k}{(r_{n'})^\beta}\right)^\nu\right]^{\frac{1}{\nu}},$$

$$L_n^L H_n^L + L_n^H H_n^H = 1,$$

$$r_n H_n^k = \beta w_n^k.$$

Consider an equilibrium where all cities have the same λ_n. Log-linearize around that equilibrium to get:

$$dy_n = \eta dl_n^L + (1-\eta) dl_n^H + (1-\eta) d\lambda_n, \tag{8}$$

$$dl_n^L = -\theta dw_n^L + dy_n, \tag{9}$$

$$dl_n^H = -\theta dw_n^H + dy_n + (\theta - 1) d\lambda_n, \tag{10}$$

$$dl_n^k = \nu dw_n^k - \nu dv^k - \nu\beta dr_n, \tag{11}$$

$$dv^k = \sum_{n'} \frac{L_n}{\bar{L}_n} (dw_{n'}^k - \beta dr_{n'}), \tag{12}$$

$$\sigma(dl_n^L + dh_n^L) + (1-\sigma)(dl_n^H + dh_n^H) = 0, \tag{13}$$

$$dr_n + dh_n^k = dw_n^k, \tag{14}$$

where we use dy_n to denote the log deviation of Y_n from the identical city case and where

$$\sigma = \frac{w_n^L L_n^L}{\sum w_n^k L_n^k},$$

and

$$\eta = \frac{\left(L_n^L\right)^{\frac{\theta-1}{\theta}}}{\left(L_n^L\right)^{\frac{\theta-1}{\theta}} + \left(\lambda_n L_n^H\right)^{\frac{\theta-1}{\theta}}}.$$

We can show that $\sigma = \eta$. To see this, solving the demand equations for w_n^k and multiplying both sides by L_n^k yields

$$w_n^L L_n^L = \left(L_n^L\right)^{\frac{\theta-1}{\theta}} (Y_n)^{\frac{1}{\theta}},$$

$$w_n^H L_n^H = \left(\lambda_n L_n^H\right)^{\frac{\theta-1}{\theta}} (Y_n)^{\frac{1}{\theta}}.$$

Plugging those equations in the formula for σ yields the formulas for η.

Next, use the log-linearized housing demand equation (14) to eliminate dh_n^k from the housing market clearing equation (13) and rearrange to get

$$\sigma(dl_n^L + dw_n^L) + (1-\sigma)(dl_n^H + dw_n^H) = dr_n.$$

Use the labor supply equation (10) to substitute out dw_n^k from the other equations in the system and rearrange. Define $ddy_n = dy_n - d\bar{y}$,

where $d\bar{y}$ denotes the cross-city average log-change in Y_n and analogously for other variables. The system becomes:

$$ddy_n = \eta ddl_n^L + (1-\eta)ddl_n^H + (1-\eta)dd\lambda_n, \tag{15}$$

$$ddl_n^L = -\frac{\theta\nu}{\theta+\nu}\beta ddr_n + \frac{\nu}{\theta+\nu}ddy_n, \tag{16}$$

$$ddl_n^H = -\frac{\theta\nu}{\theta+\nu}\beta ddr_n + \frac{\nu}{\theta+\nu}ddy_n + \frac{\nu(\theta-1)}{\theta+\nu}dd\lambda_n, \tag{17}$$

$$0 = \sum_{n'} \frac{L_n}{\bar{L}_n}(ddw_{n'}^k - \beta ddr_{n'}), \tag{18}$$

$$\sigma\frac{1+\nu}{\nu}ddl_n^L + (1-\sigma)\frac{1+\nu}{\nu}ddl_n^H = (1-\beta)\,ddr_n. \tag{19}$$

Using equations (15), (16), and (17) to substitute out ddl_n^k from equation (19) and using $\sigma = \eta$ yields an expression for rents as a function of productivity.

$$ddr_n/dd\lambda_n = \frac{1+\nu}{1+\nu\beta}(1-\eta) > 0.$$

Manipulating (15), (16), and (17) to obtain an expression ddy_n as a function of ddr_n, we can substitute the expression above to get:

$$ddy_n/dd\lambda_n = (1-\eta)\frac{\nu+1}{1+\nu\beta} > 0.$$

For the limiting case with $\nu \to \infty$, this yields

$$ddy_n = \beta^{-1}(1-\eta)dd\lambda_n.$$

For the other limiting case, with $\nu = 0$

$$ddy_n = (1-\eta)dd\lambda_{n.}$$

We can also obtain expressions for ddl_n^L and ddl_n^H:

$$ddl_n^L/dd\lambda_n = \frac{\nu}{\theta+\nu}(1-\eta)\frac{1+\nu}{1+\nu\beta}\left[1-\theta\beta\right],$$

$$ddl_n^H/dd\lambda_n = \frac{\nu}{\theta+\nu}\left[(1-\eta)\beta\frac{1+\nu}{1+\nu\beta}\left[\beta^{-1}-1\right] + (\theta-1)\left(1-(1-\eta)\frac{\beta+\nu\beta}{1+\nu\beta}\right)\right] > 0.$$

It follows that $ddl_n^L > 0$ if $1 > \theta\beta$, i.e., if land share is sufficiently small or the two types of labor are not too strongly substitutable. Normal calibrations feature $\beta \simeq \frac{1}{3}$ and $\theta \simeq 2$, so that the condition is satisfied and population grows for both types of workers with λ_n, although most strongly for the high type.

Again, with $\nu \to \infty$ the expressions simplify to

$$\begin{aligned} ddl_n^L/dd\lambda_n &= (1-\eta)\left[\beta^{-1}-\theta\right], \\ ddl_n^H/dd\lambda_n &= \left[(1-\eta)\left[\beta^{-1}-1\right]+\eta\left(\theta-1\right)\right] > 0, \end{aligned}$$

and we verify that with $\nu = 0$ employment does not change:

$$\begin{aligned} ddl_n^L/dd\lambda_n &= 0, \\ ddl_n^H/dd\lambda_n &= 0. \end{aligned}$$

Calculating real wages, we get:

$$\begin{aligned} ddw_n^L/dd\lambda_n - \beta ddr_n &= \frac{1}{\theta+\nu}(1-\eta)\frac{1+\nu}{1+\nu\beta}\left[1-\theta\beta\right], \\ ddw_n^H/dd\lambda_n - \beta ddr_n &= \frac{1}{\theta+\nu}\left[(1-\eta)\beta\frac{1+\nu}{1+\nu\beta}\left[\beta^{-1}-1\right]+(\theta-1)\left(1-(1-\eta)\frac{\beta+\nu\beta}{1+\nu\beta}\right)\right]. \end{aligned}$$

Thus, real wages will increase or decrease in relative terms with relative employment. With $\nu \to \infty$, we have that real wages do not change in relative terms at all. With $\nu = 0$,

$$\begin{aligned} ddw_n^H - \beta ddr_n &= \frac{1}{\theta}(1-\eta)\left[1-\theta\beta\right], \\ ddw_n^H - \beta ddr_n &= \frac{1}{\theta}\left[(1-\eta)\left[1-\beta\right]+(\theta-1)\left(1-(1-\eta)\beta\right)\right]. \end{aligned}$$

Finally, we can obtain nominal wages:

$$ddw_n^H/dd\lambda_n = \frac{1+\beta\nu}{\theta+\nu}(1-\eta)\frac{1+\nu}{1+\nu\beta},$$
$$ddw_n^H/dd\lambda_n = \frac{1}{\theta+\nu}\left[(1-\eta)\beta\frac{1+\nu}{1+\nu\beta}\left[\beta^{-1}-1\right]+(\theta-1)\left(1-(1-\eta)\frac{\beta+\nu\beta}{1+\nu\beta}\right)\right] + \beta\frac{1+\nu}{1+\nu\beta}(1-\eta),$$

so that both nominal wages increase unambiguously. Taking limits, for $\nu \to \infty$

$$ddw_n^L/dd\lambda_n = 1-\eta.$$
$$ddw_n^H/dd\lambda_n = 1-\eta,$$

whereas for $\nu = 0$

$$ddw_n^L/dd\lambda_n = \frac{1}{\theta}(1-\eta),$$
$$ddw_n^H/dd\lambda_n = 1-\frac{\eta}{\theta}.$$

Comparative Statics with Housing as a Necessity

With housing as a necessity good as in subsection 2.5, we can derive the labor supply function following the same steps as in Appendix 1 to get:

$$L_n^k = \frac{\left(w_n^k - r_n\right)^\nu}{\sum_{n'}\left(w_{n'}^k - r_{n'}\right)^\nu}\bar{L}^k.$$

It follows that, using the same "deviation" notation as above

$$ddL_n^k = \nu\frac{w^k}{w^k-r}ddw_n^k - \nu\frac{r}{w^k-r}ddr_n.$$

Since all workers buy one unit of housing, the housing market equilibrium equation can be written as:

$$\sum_k L_n^k = \bar{H}_n.$$

Suppose the production function is the same in all cities but housing supply is not. The equations defining the equilibrium of the economy in log-linearized form become:

$$ddy_n = \eta ddl_n^L + (1-\eta) ddl_n^H, \tag{20}$$

$$ddl_n^k = -\theta ddw_n^k + ddy_n \;\; k \in \{L, H\}, \tag{21}$$

$$ddl_n^k = \nu \frac{w^k}{w^k - r} ddw_n^k - \nu \frac{r}{w^k - r} ddr_n, \;\; k \in \{L, H\}, \tag{22}$$

$$0 = \sum_{n'} \frac{L_n}{\bar{L}_n} \nu \left(\frac{w^k}{w^k - r} ddw_n^k - \frac{r}{w^k - r} ddr_n \right), \tag{23}$$

$$\sigma ddl_n^L + (1-\sigma) ddl_n^H = ddh_n, \tag{24}$$

where now $\sigma = \frac{L_n^L}{L_n^L + L_n^H}$. Use the new labor supply equation (22) to eliminate wages from the labor demand equations and the production function to eliminate output:

$$ddl_n^k = -\frac{\theta}{\nu} \left[\frac{w^k - r}{w^k} ddl_n^k + \frac{r}{w^k} ddr_n \right] + \eta ddl_n^L + (1-\eta) ddl_n^H.$$

Combine those with the production function and substitute out the new housing clearing equilibrium (24) to obtain:

$$\left[\eta + \frac{\theta}{\nu} \frac{w^H - r}{w^H} \right] ddl_n^H - \eta ddl_n^L = -\frac{\theta}{\nu} \frac{r}{w^H} ddr_n,$$

$$-(1-\eta) ddl_n^H + \left[1 - \eta + \frac{\theta}{\nu} \frac{w^L - r}{w^L} \right] ddl_n^L = -\frac{\theta}{\nu} \frac{r}{w^L} ddr_n.$$

Solve the system for ddl_n^L and ddl_n^H as functions of ddr_n. From inspection of the structure of the problem, it follows that:

$$ddl_n^H = -\varphi^H ddr_n$$
$$ddl_n^L = -\varphi^L ddr_n,$$

where $\varphi^H \equiv \frac{\theta}{\nu} \frac{\left[\eta + \frac{\theta}{\nu} \frac{w^L - r}{w^L}\right] \frac{r}{w^H} + \eta \frac{r}{w^L}}{D} > 0$ and $\varphi^L \equiv \frac{\theta}{\nu} \frac{\left[1 - \eta + \frac{\theta}{\nu} \frac{w^H - r}{w^H}\right] \frac{r}{w^L} + (1-\eta) \frac{r}{w^H}}{D} > 0$ with $D \equiv \left[\eta + \frac{\theta}{\nu} \frac{w^H - r}{w^H}\right] \left[1 - \eta + \frac{\theta}{\nu} \frac{w^L - r}{w^L}\right] + (1-\eta)\eta$. Note that $\varphi^H > \varphi^L$ if

$$\frac{\theta}{\nu} \left(\frac{w^H}{w^L} - 1 \right) > (1 - 2\eta) \left(1 + \frac{w^H}{w^L} \right).$$

Substituting back into the housing market clearing equation yields:

$$-\left[\sigma\varphi^H + (1-\sigma)\varphi^L\right] ddr_n = ddh_n.$$

It follows that interest rates fall with housing supply, and employment in both sectors rises. At the same time, the labor mix will change. So long as $\frac{w^H}{w^L}$ is sufficiently greater than 1, skilled labor will increase more rapidly with housing supply than unskilled labor. At the same time, because the elasticity of labor supply at the city level is not infinity, wages for skilled labor will have to grow faster.

Consider now an environment with three worker types, L, HL, and HH. The production function is now

$$Y_n = \left[\left(L_n^L\right)^{\frac{\theta-1}{\theta}} + \left(\lambda\left(L_n^{HL} + \mu L_n^{HH}\right)^{\frac{\theta-1}{\theta}}\right)\right]^{\frac{\theta}{\theta-1}}.$$

For simplicity, we will focus on the case with $\theta \to \infty$ so that high- and low-skilled workers are also perfect substitutes. Suppose now there are common city-specific amenities, so that their utility of living in city n is now:

$$G_n^k(X_n^k, \varepsilon(i)) = \begin{cases} A_n\varepsilon(i)X_n^k \ if \ H_n^k \geq 1 \\ -\infty \ otherwise \end{cases},$$

where A_n represents amenities such as good public schools or nice weather. Those factors are common for all types of labor. Labor supply for $k \in \{LH, HH\}$

$$L_n^k = \frac{\left(A_n\left(w_n^k - r_n\right)\right)^\nu}{\sum_{n'}\left(A_{n'}\left(w_{n'}^k - r_{n'}\right)\right)^\nu}\bar{L}^k.$$

Using the fact that different types of labor are perfect substitutes so that their wages do not vary in the cross-section, the log-linearized system in deviation form becomes:

$$ddy_n = (1-\eta^{LH}-\eta^{HH})ddl_n^L + \eta^{LH}ddl_n^{LH} + \eta^{HH}ddl_n^{HH}, \tag{25}$$

$$ddl_n^k = -\nu\frac{r}{w^k - r}ddr_n + \nu dda_n^k, \ k \in \{L, HL, HH\}, \tag{26}$$

$$0 = \sum_{n'}\frac{L_n}{\bar{L}_n}\nu\frac{r}{w^k - r}ddr_n, \tag{27}$$

$$(1-\sigma^{HL}-\sigma^{HH})ddl_n^L + \sigma^{HL}ddl_n^{HL} + \sigma^{HH}ddl_n^{HH} = 0, \tag{28}$$

Substitute the new labor supply equations into the market clearing equation to get:

$$\left[(1-\sigma^{HL}-\sigma^{HH})\frac{r}{w^{L}-r}+\sigma^{HL}\frac{r}{w^{HL}-r}+\sigma^{HH}\frac{r}{w^{HH}-r}\right]ddr_n = dda_n.$$

It follows that rental rates rise with amenities. At the same time, from equation (26) it follows immediately that employment reacts more to interest rates if wages are smaller. Thus, a given increase in amenities will imply larger increases in the population of workers with higher wages. At the same time, even though wages do not change, if one cannot observe HL and HH type workers separately, one will see their average wage rise, as the composition of high-skilled workers shifts toward the ones with a very high skill. Thus, cities with higher amenities will have a more skilled labor force and higher wage premia. One can further generate the correlation with city size by allowing housing supply to be rent elastic.

REFERENCES

Acemoglu, Daron. 1996. "A Microfoundation for Social Increasing Returns in Human Capital Accumulation." *Quarterly Journal of Economics* 111 (August): 779–804.

Albouy, David. 2012. "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas." Working Paper (May).

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (August): 1553–97.

Bacolod, Marigee, Bernardo S. Blum, and William C. Strange. 2009. "Skills in the City." *Journal of Urban Economics* 65 (March): 136–53.

Baum-Snow, Nathaniel, Matthew Freedman, and Ronni Pavan. Forthcoming. "Why Has Urban Inequality Increased?" *American Economic Journal: Applied Economics.*

Baum-Snow, Nathaniel, and Ronni Pavan. 2012. "Understanding the City Size Wage Gap." *Review of Economic Studies* 79 (January): 88–127.

Baum-Snow, Nathaniel, and Ronni Pavan. 2013. "Inequality and City Size." *Review of Economics and Statistics* 95 (December): 1535–48.

Beaudry, Paul, Mark Doms, and Ethan Lewis. 2010. "Should the Personal Computer be Considered a Technological Revolution? Evidence From U.S. Metropolitan Areas." *Journal of Political Economy* 118 (October): 988–1036.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud. 2014. "Productive Cities: Sorting, Selection, and Agglomeration." *Journal of Political Economy* 122 (June): 507–53.

Berry, Christopher R., and Edward L. Glaeser. 2005. "The Divergence of Human Capital Levels Across Cities." *Papers in Regional Science* 84 (August): 407–44.

Brinkman, Jeffrey. 2014. "The Supply and Demand of Skilled Workers in Cities and the Role of Industry Composition." Federal Reserve Bank of Philadelphia Working Paper 14-32 (October).

Ciccone, Antonio, and Giovanni Peri. 2005. "Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950–1990." *Review of Economics and Statistics* 87 (November): 652–63.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. "Spatial Wage Disparities: Sorting matters!" *Journal of Urban Economics* 63 (March): 723–42.

Combes, Pierre-Philippe, and Laurent Gobillon. 2015. "The Empirics of Agglomeration Economies." In *Handbook of Urban and Regional Economics Vol. 5*, edited by Gilles Duranton, J. Vernon Henderson, and William Strange. Amsterdam: North Holland, 247–348.

Davis, Donald R., and Jonathan I. Dingel. 2016. "A Spatial Knowledge Economy." Working Paper (August).

Davis, Donald R., and Jonathan I. Dingel. 2017. "The Comparative Advantage of Cities." Working Paper (March).

De La Roca, Jorge, and Diego Puga. 2017. "Learning by working in big cities." *Review of Economic Studies* 84 (January): 106–42.

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

Duranton, Gilles, and Diego Puga. 2004. "Micro-foundations of Urban Agglomeration Economies." In *Handbook of Urban and Regional Economics Vol. 4*, edited by J. Vernon Henderson and Jacques-François Thisse. Amsterdam: Elsevier, 2063–117.

Duranton, Gilles, and Diego Puga. 2005. "From Sectoral to Functional Urban Specialisation." *Journal of Urban Economics* 57 (March): 343–70.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny. 2014. "Spatial Sorting." *Journal of Political Economy* 122 (June): 554–620.

Fajgelbaum, Pablo D., and Cecile Gaubert. 2018. "Optimal Spatial Policies, Geography and Sorting." Working Paper 24632. Cambridge, Mass.: National Bureau of Economic Research. (May).

Giannone, Elisa. 2017. "Skilled-Biased Technical Change and Regional Convergence." University of Chicago Working Paper (January).

Glaeser, Edward L., and David C. Maré. 2001. "Cities and Skills." *Journal of Labor Economics* 19 (April): 316–42.

Handbury, Jessie, and David E. Weinstein. 2015. "Goods Prices and Availability in Cities." *Review of Economic Studies* 82 (January), 258–96.

Hendricks, Lutz. 2011. "The Skill Composition of U.S. Cities." *International Economic Review* 52 (February): 1–32.

Hsieh, Chang-Tai, and Enrico Moretti. 2017. "Housing Constraints and Spatial Misallocation." Working Paper (May).

Krusell, Per, et al. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* 68 (September): 1029–53.

Lucas Jr., Robert E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (July): 3–42.

Marshall, Alfred. 1890. *Principles of Economics*, 1st ed. London: Macmillan.

Michaels, Guy, Ferdinand Rauch, and Stephen Redding. 2017. "Task Specialization in U.S. Cities from 1880-2000." Working Paper 18715. Cambridge, Mass.: National Bureau of Economic Research. (August).

Moretti, Enrico. 2004. "Human Capital Externalities in Cities." In *Handbook of Urban and Regional Economics Vol. 4*, edited by J. Vernon Henderson and Jacques-François Thisse. Amsterdam: Elsevier, 2243–91.

Moretti, Enrico. 2013. "Real Wage Inequality." *American Economic Journal: Applied Economics* 5 (January): 65–103.

Notowidigdo, Matthew J. 2011. "The Incidence of Local Labor Demand Shocks." Working Paper 17167. Cambridge, Mass.: National Bureau of Economic Research. (December).

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (December): 1257–78.

Rosen, Sherwin. 1979. "Wage-Based Indexes of Urban Quality of Life." In *Current Issues in Urban Economics Vol. 3*, edited by Peter Mieszkowski and Mahlon Straszheim. Baltimore: Johns Hopkins University Press, 324–45.

Economic Quarterly—Volume 103, Numbers 1-4—First-Fourth Quarter 2017—Pages 37–79

The Fed's Discount Window: An Overview of Recent Data

Felix P. Ackon and Huberto M. Ennis

The Dodd-Frank Act of 2010 requires the Federal Reserve to publish data on discount-window transactions with approximately two years' delay, starting from the passage of the Act in July 2010. The availability of these data provides an opportunity, for policymakers and researchers alike, to get a more detailed perspective on the nature of lending in this important and traditional central bank credit facility.

The Fed provides credit through the discount window using three different programs: primary credit, secondary credit, and seasonal credit. Primary credit and secondary credit are emergency credit programs. These programs constitute a backup source of short-term funding for eligible financial institutions. Seasonal credit is aimed at smaller institutions with a predictable (and demonstrable) seasonal pattern in their funding needs. Each loan must be secured by collateral from the borrowing institution.

The primary-credit program is a standing facility in which depository institutions in good financial conditions (in the form of a high score on their examination rating) can access (mainly) overnight funding with "no questions asked" and at an interest rate higher than the target policy short-term rate. Those institutions not eligible for primary credit can receive secondary credit.

Normally, secondary credit is offered at a rate that is fifty basis points above the primary-credit rate. Furthermore, secondary-credit

We would like to thank Chris Mast for helpful discussions and Caroline Davis, Doug Diamond, Borys Grochulski, Beth Klee, Steve McCarther, John Walter, and John Weinberg for comments on a previous draft. The views expressed in this article are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mails: Felix.Ackon@rich.frb.org; Huberto.Ennis@rich.frb.org.

DOI: https://doi.org/10.21144/eq1030102

loans are subject to an extra degree of scrutiny (by the Fed) and cannot be used for certain purposes, such as interest rate arbitrage.

To receive seasonal credit, a financial institution must have a demonstrated seasonal need for funding and an approved seasonal line of credit with its corresponding Reserve Bank. The interest rate is a floating rate that is calculated based on the average of some selected rates in the money market—it is not intended to be a penalty rate like the others (the interest rates at the primary- and secondary-credit programs). A bank that is able to receive seasonal credit can also tap the primary-credit program if the need arises and the bank is in sound financial condition.

There is an extensive theoretical literature exploring issues related to the provision of discount-window credit by a central bank (see Ennis [2016] for a recent review). However, the empirical literature is much more scant.[1] One obvious reason for this situation is that, over the years, disaggregated data on discount window activity have not been made available to the public or researchers on a regular basis. There are (potentially) good reasons for this lack of transparency: central banks are often concerned about the possibility that some degree of stigma is associated with accessing the discount window (for a discussion of this issue see, for example, Courtois and Ennis [2010]). By keeping loan information private, central banks seek to minimize the incidence of stigma.

On the other hand, more information is a constant demand from stakeholders looking to oversee the central bank's activities. The desire for more transparency was duly recognized in the Dodd-Frank Act, and, as a result, a detailed description of the Fed's discount-window actions is now easily accessible on the internet with a two-year lag, approximately. We take advantage of the new data availability in this paper and provide an overview of the main patterns identifiable in such data so far.

The paper is organized as follows. In the next section, we discuss the data used as well as some preliminary statistics. In Section 2, we present information about primary-credit loans. In Section 3, we study

[1] There is some empirical work on emergency lending by government institutions in the US. The early work by Furfine (2003), for example, uses data on market transactions and interest rates and compares that information with the interest rate and volume at the Fed's discount window to infer the attitudes of market participants toward central bank liquidity provision. More recently, several studies have used data on the Term Auction Facility to study different aspects of how emergency credit provision worked during the recent financial crisis (see, for example, Benmelech [2012], Berger et al. [2017], and Armantier et al. [2015]). From a historical perspective, Anbil (2015) and Vossmeyer (2017) study borrowing from the Reconstruction Finance Corporation during the Great Depression. Finally, for a detailed study of the European experience during the recent crisis, see Drechsler et al. (2016).

Table 1 Discount Window Lending–Totals

	Total number of loans	Total number of loans $\leq$ \$10,000	Total amount lent (\$mm)
All	16514	5277	36124.977
Primary credit	11429	4655	26286.218
Secondary credit	650	611	118.976
Seasonal credit	4435	11	9719.781

Note: We use all the data in our sample for these calculations. In a few cases, a given loan entry is composed of multiple loans to the same institution with the same term but, possibly, different amounts, which are consolidated in one entry for the purpose of reporting.

secondary-credit loans, and in Section 4, seasonal credit. Section 5 provides an analysis of the collateral pledged by borrowing banks to their corresponding Reserve Banks. Finally, Section 6 concludes.

1. THE DATA

Our data are a comprehensive list of discount-window loans made by Federal Reserve Banks between July 22, 2010, and June 30, 2015. The Federal Reserve releases the information on a quarterly basis, with approximately a two-year lag. The data for the second quarter of 2015 are the most recent data available at the time of this writing.

There are 16,514 individual loan observations. In some cases, an observation is the result of the consolidation of several loans granted to the same borrower on the same day with the same term to maturity. Furthermore, some loans in the sample are the consequence of a bank rolling over a previous loan. In those cases, loans are counted separately (i.e., a loan taken for one day and rolled over for a second day counts as two independent loans if at origination each of the loans was granted as an overnight loan).

For each loan, the data include: the date the loan was granted, the term to maturity, the date the loan was repaid, the name and location of the borrower and the Reserve Bank playing the role of lender, the type of credit (primary, secondary, or seasonal), the interest rate and loan amount, whether the bank had other loans outstanding, and a description of the collateral pledged by the borrowing bank.

Because the target policy rate (or range) did not change during the sample period, all primary-credit loans in the sample were granted at the same interest rate of 0.75 percent and all secondary-credit loans but one were granted at the interest rate of 1.25 percent (see Section 3 for a detailed discussion of the exceptional loan).

Many of the loans in the sample are likely to be "test" loans. These are loans that depository institutions take in order to test whether the systems involved in processing a discount-window loan are working as expected. Test loans are generally for a very small amount: $1,000 is probably the most common amount, but most of the $10,000 loans in our sample are likely to also be tests, and a few loans of larger amounts may also be tests.

Unfortunately, the data available do not include information about whether a given loan is a test or not. There are a lot of $1,000 loans in the data: 27 percent of the primary-credit loans in the data are $1,000 or less. For secondary-credit loans, the percentage of loans with amounts at or below $1,000 is much higher, 82 percent. If we take an even more conservative approach and consider test loans as all loans for amounts equal to or less than $10,000 in order to identify more accurately actual credit events, then we can see from Table 1 that about a third of the loans were in that category.

Table 1 also reports the number of loans granted under the different programs. Two-thirds of the loans were granted under the primary-credit program. Interestingly, there seem to be a lot fewer test loans happening in the seasonal-credit program. Using our threshold for test loans (amounts less than or equal to $10,000), we calculate that the number of nontest loans in the primary-credit and seasonal-credit programs is comparable. In contrast, the number of secondary-credit nontest loans is much less (around forty).

The total amount lent, instead, is much higher in the primary-credit program. For calculating the total amount lent, excluding the test loans does not make much of a difference. For example, the total value of all primary-credit loans with a face value equal to or below $10,000 is approximately $18 million, which is much less than 1 percent of the total amount lent ($26.3 billion, as reported in the third column of Table 1). It seems clear from these numbers, then, that the primary-credit program is the more active and significant lending program operated by the Fed's discount window.

Since the proportion of test loans is very different across different programs, to get a sense of the size of the typical loans in each program, it is important to take into account the implied composition effect. To this end, Table 2 reports some basic statistics when we exclude from the sample all loans of $10,000 or less. We see in the table that the average size of loans is similar for primary and secondary credit and relatively smaller for seasonal-credit loans. Interestingly, the largest primary-credit loan in the sample is several orders of magnitude larger than the largest loans in the other two programs.

Table 2 Loan Size, Excluding Tests–Basic Statistics

	Number of loans	Total amt. lent ($mm)	Average size of loans ($mm)	Median size of loans ($mm)	Max. loan size ($mm)
All	11237	36105.762	3.213	1.000	1017.018
Primary c.	6774	26268.305	3.818	1.000	1017.018
Secondary c.	39	117.775	3.020	2.000	17.000
Seasonal c.	4424	9719.681	2.197	1.000	26.000

Note: We define as "tests" all loans of size equal or smaller than $10,000. These loans are excluded from the sample used to compute these statistics. In a few cases, a given loan entry is composed of multiple loans to the same institution with the same term but, possibly, different amounts, which are consolidated in one entry for the purpose of reporting.

Table 3 Discount-Window Lending Over Time

Panel A – Total amount lent ($mm)

Year	All	All w/o	Primary c.	Secondary c.	Seasonal c.
2011	6656.703	6653.530	5821.794	50.251	784.659
2012	7859.556	7856.076	6087.402	7.868	1764.286
2013	6176.920	6172.846	3795.127	3.743	2378.050
2014	6849.539	6844.392	3419.413	4.855	3425.271

Panel B – Number of loans

Year	All		Primary credit		Secondary credit		Seasonal c.	
	w/	w/o	w/	w/o	w/	w/o	w/	w/o
2011	2801	1941	2026	1316	169	21	606	604
2012	3021	2049	2115	1290	150	4	756	755
2013	3491	2390	2326	1357	129	1	1036	1032
2014	3996	2554	2679	1333	98	5	1219	1216

Note: We use "w/" to denote the columns where the number of loans includes "test" loans and "w/o" to denote where "test" loans are excluded. We aggregate at the annual level to compare across time the recent evolution of discount-window lending. For these calculations, we do not use the available data for 2010 and 2015 because we do not have full-year information for those years and, as we will see later, the data display seasonal yearly patterns that can make the comparison of full-year and half-year aggregates misleading. We define as "tests" all loans equal to or smaller than $10,000.

Table 3 gives a sense of the evolution over (recent) time of the amount of credit provided through the discount window. We choose to aggregate loans on an annual basis because, as we will discuss later, credit activity exhibits some yearly seasonal patterns.

The total amount lent is relatively stable between the years 2011 and 2014 (see Panel A of Table 3).[2] This relative stability is the combination of two trends: on one side, the amount of lending in the primary-credit facility decreased significantly (a 40 percent drop) after 2012. On the other side, the total amount lent at the seasonal-credit facility has been increasing over time.

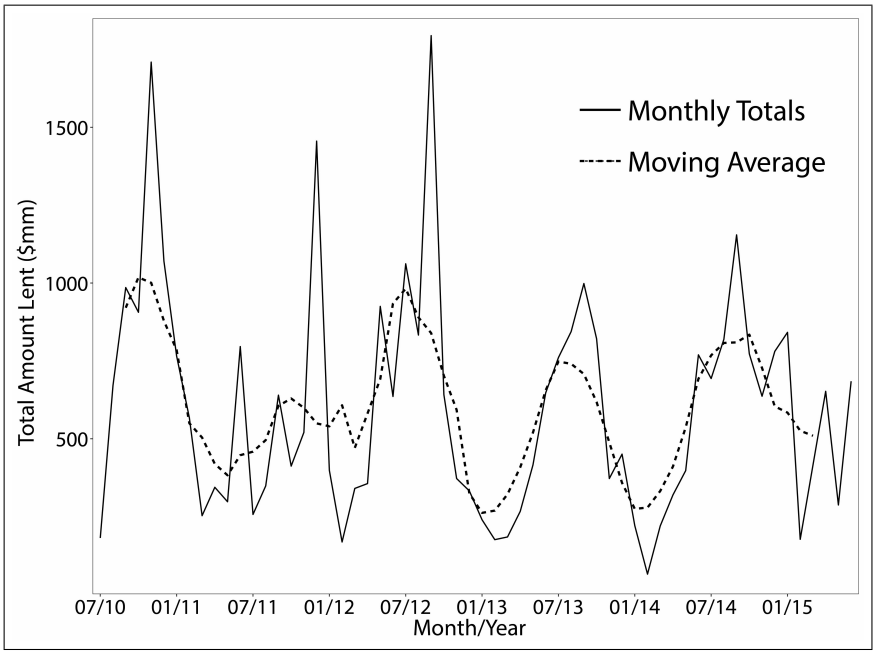
The total amount lent through the secondary-credit facility was relatively large in 2011 but became much smaller in the years after that. It is possible that the repercussions of the financial crisis were still driving banks to borrow secondary credit, as this is the facility available for banks in weaker financial condition. In support of that possibility, note that, for example, the number of "problem banks" in the US (as reported by the Federal Deposit Insurance Corporation [FDIC]) peaked in 2011 at (around) 900 and since then has been experiencing a steady decline, reaching 200 in 2015.

Panel B of Table 3 shows the evolution over time of the total number of loans granted through the different programs. The total number of loans in all three programs is growing over time. This phenomenon is basically explained by two main factors: first, the number of test loans in the primary-credit program has grown over time and more than proportionally (to the total number of loans) in 2014—note that the number of loans in the primary-credit program, net of our measure of the number of test loans, is basically stable over the years at around 1,300 loans. The second factor is that the number of seasonal loans has been increasing consistently over the years under consideration.

Figure 1 plots the monthly total amount lent at the discount window during our sample period. Since the data are volatile, we include in the figure a six-month moving average to highlight any trends present in the data. It is evident in the figure that lending generally increases in the second half of the year. As we will see later, this pattern is present both at the primary-credit and seasonal-credit programs. Aside from those fluctuations, lending seems relatively stable over the five-year period.

[2] Relatively speaking, as a source of funding for banks, the discount window is miniscule during normal times (such as the years covered in our sample). For example, total deposits at FDIC-insured institutions in the last quarter of 2014 were close to $12 trillion.

Figure 1 Total Discount-Window Lending



Note: For each month in our sample, we aggregate all loans in that month to generate the monthly series plotted as the solid line. We also compute a symmetric, six-month moving average, which is the smoother dashed line.

Given that now the expectation in financial markets is that information about activity at the discount window will become public after two years, one natural question to ask is whether banks have used the discount window less as a result of this added transparency. To get a sense of this, it would be good to compare the levels of lending during our sample period with lending in previous periods of similar characteristics—mainly, noncrisis periods. While loan-level data are not publicly available for the period before 2007, the Federal Reserve has been publishing for a much longer time the average daily amount of outstanding loans at the discount window in its H.4.1 weekly data release.

Table 4 uses the data from the H.4.1 release to compare the annual average of daily outstanding loans at the three main discount-window programs for the period 2003–07 and for the period 2011–15. We see in the table that the level of discount-window activity is relatively lower

Table 4 Total Discount-Window Lending Before and After the Crisis

BEFORE THE CRISIS		AFTER THE CRISIS	
Year	Average daily loan amount outstanding ($mm)	Year	Average daily loan amount outstanding ($mm)
2003	98.00	2011	62.058
2004	153.038	2012	72.442
2005	199.346	2013	78.654
2006	224.654	2014	117.585
2007	587.788	2015	124.673

Note: The annual number is the average of the daily numbers reported for each week of the year in the Federal Reserve H.4.1 Data Release.

in the period after the Dodd-Frank Act. This is consistent with the view that disclosure tends to discourage borrowing (Kleymenova 2016). However, other changes in the financial landscape after the crisis may have also contributed to the reduction in discount-window activity. For example, the high levels of interest-paying excess reserves (Ennis and Wolman 2015) tend to reduce the risk for depository institutions to find themselves short of funds due to unexpected payment shocks. Similarly, more recently, heightened regulatory focus on balance sheet liquidity might be inducing banks to hold liquidity buffers that can also protect them from unexpected late-in-the-day payment shocks. Presumably, then, these liquidity enhancements make banks less likely to need to tap the discount window as a backup source of funding.

A look at more disaggregated data shows (see Table A1 in the Appendix) that *both* primary and seasonal credit have decreased after the crisis—secondary credit is very small (and volatile) in both periods. Interestingly, after a significant decrease during the crisis, seasonal credit seems to be converging back to numbers common in the precrisis period. Primary credit, instead, has been consistently smaller and decreasing in time after the crisis (see also Table 3).

Note, finally, in Table 4 that discount-window lending increased significantly in 2007. In fact, the corresponding amounts for 2008 and 2009 (see Table A1) are an order of magnitude larger than the numbers in Table 4 ($32 billion and $40 billion, respectively), and the second half of 2007 was already showing the signs of those forthcoming increases. These "crisis" numbers put in perspective the levels of lending that occur in "normal" times (as in our sample period), which is not zero but relatively moderate.

Going back to the transactions data, it is also possible to get a sense of the variation in activity across Federal Reserve districts. In Panel A of Table 5, we show the total amount lent by each Reserve Bank over our sample period. The largest lenders in terms of total amount lent are the Atlanta, Chicago, and San Francisco Feds. The San Francisco Fed lent mostly through the primary-credit program, and the Atlanta Fed lent more than half of its total through the seasonal-credit program. Indeed, the Atlanta Fed is the largest lender of seasonal credit among all the Reserve Banks, followed by Minneapolis.

Based on the information on the number of loans granted, as provided in Panel B, it seems that the Chicago and Minneapolis Reserve Banks grant a lot more loans at the seasonal-credit facility than the other Banks, and in particular, Atlanta. This, in turn, implies that the average size of the seasonal-credit loans is much smaller in the case of the Chicago than in the case of Atlanta: the first grants a lot of small seasonal loans, and the second grants fewer loans but larger ones.

Philadelphia, San Francisco, and Chicago are the top providers of secondary credit in terms of total amount lent. While the Atlanta Fed provided a large number of secondary-credit loans (244), all of them were $1,000 and, hence, most likely test loans. If we restrict the sample to those loans that are greater than $10,000, we see (in the column labeled "w/o") that San Francisco and New York rank at the top in terms of the number of secondary-credit loans granted during the sample period. It should be said, though, that the small sample size makes it particularly hard in this case to draw strong inferences about general patterns.

The largest lender through the primary-credit program (in terms of total amount lent) is the San Francisco Fed. The Chicago Fed made the largest number of loans through this program, but almost half of them were loans that are very likely to have been test loans (for amounts equal or below $10,000 dollars). For most of the Reserve Banks, between a third and a half of the primary-credit loans can be safely categorized as test loans.

Perhaps surprisingly, when we restrict the sample to loans greater than $10,000 and look at the average size of primary-credit loans across districts (not reported in the table), we see that the Dallas Fed is providing the largest loans on average ($6.4 million), followed by Philadelphia ($5.9 million), and St. Louis ($5.6 million). The New York Fed is fourth (with an average loan amount equal to $5.5 million). In principle, since the highest concentration of large banks is in the New York district, one might have expected that the largest discount-window loans were being originated there. That is not the case in our sample period.

Table 5 Discount-Window Lending Across Federal Reserve Districts

Panel A – Total amount lent ($mm)

District	All	Primary c.	Secondary c.	Seasonal c.
Boston	1656.874	970.536	1.608	684.730
New York	3375.340	3354.844	20.496	0
Philadelphia	1713.236	1678.804	34.432	0
Cleveland	482.539	482.530	0.009	0
Richmond	1550.483	1546.627	3.856	0
Atlanta	5748.636	2176.948	0.244	3571.444
Chicago	4889.066	3892.651	28.196	968.219
St. Louis	2963.001	2105.557	0.420	857.024
Minneapolis	3493.014	1464.924	0.118	2027.972
Kansas City	1761.168	1053.308	0.500	707.360
Dallas	3765.147	3044.297	0	720.850
San Francisco	4726.473	4515.191	29.100	182.182

Panel B – Number of loans

	All		Primary c.		Secondary c.		Seasonal c.	
District	**w/**	**w/o**	**w/**	**w/o**	**w/**	**w/o**	**w/**	**w/o**
Boston	989	542	701	262	10	2	278	278
New York	932	615	909	604	23	11	0	0
Philadelphia	456	284	429	281	27	3	0	0
Cleveland	399	230	390	230	9	0	0	0
Richmond	865	333	825	331	40	2	0	0
Atlanta	1963	1194	1133	608	244	0	586	586
Chicago	3616	2333	2194	1111	202	6	1220	1216
St. Louis	1279	737	871	373	42	0	366	364
Minneapolis	2080	1931	826	715	38	0	1216	1216
Kansas City	1624	1286	1079	745	1	1	544	540
Dallas	645	639	480	474	0	0	165	165
San Francisco	1666	1113	1592	1040	14	14	60	59

Note: We use "w/" to denote the columns where the number of loans includes "test" loans and "w/o" to denote where "test" loans are excluded. There are twelve Federal Reserve districts, each with its corresponding Reserve Bank (which takes the name of the city where it is located). Reserve Banks make loans to institutions headquartered in their district.

Finally, note (Panel B of Table 5) that there are very few loans of $10,000 or less granted by the Dallas Fed (in all three programs). Upon closer inspection of the transactions data, one can see that 264 loans (out of the 645 loans) were for exactly $100,000. This suggests that different Reserve Banks may have different practices when it comes to the size of the loans used for testing, with the Dallas Fed perhaps being a Bank that recommends larger test loans than other Feds.

2. PRIMARY CREDIT

In this section, we explore further the main characteristic of loans granted through the primary-credit program. We look at the distribution of loan sizes and maturity terms. We also investigate the extent to which lending displays seasonal patterns and, lastly, whether some banks are more prone than others to tap the primary-credit facility (that is, the intensity of use across banks).

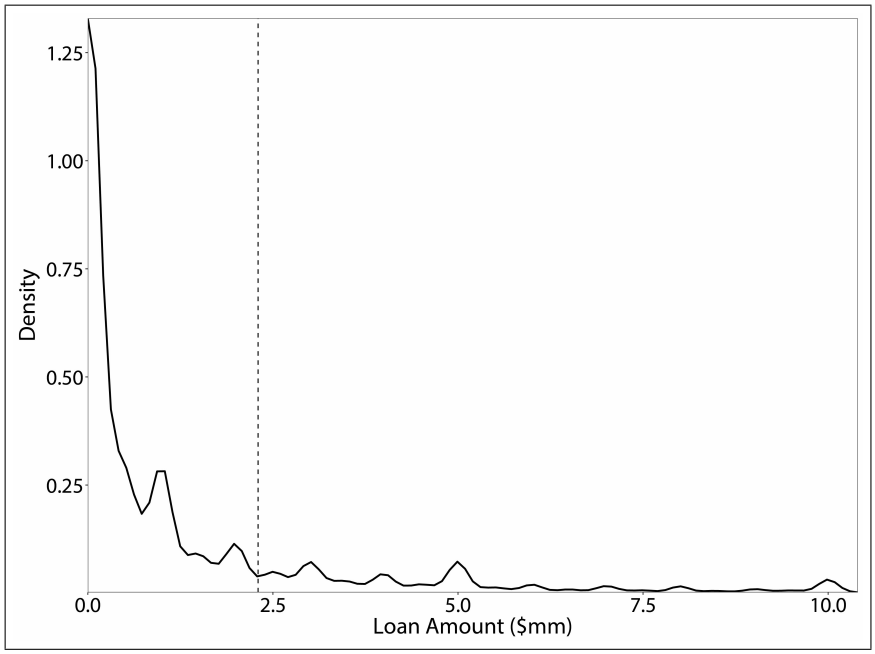
Loan sizes: as was discussed in Section 1, the median primary-credit loan is much smaller than the average. This is because the largest loans in our sample were granted through the primary-credit facility, which also made a significant proportion of relatively small loans. Even when we abstract from the majority of the small test loans, as we did in Table 2, we see that the average size of primary-credit loans is four times larger than the median.

Figure 2 plots the distribution of loan sizes for all loans smaller than or equal to $10 million. It seems clear from the figure that most loans are in the size range that is below $5 million. In fact, during our sample period, 40 percent of all primary-credit loans were $10,000 or less, and more than half were $100,000 or less. On the other end of the distribution, 10 percent of loans were greater than $5 million.

The distribution of loan sizes in Figure 2 appears to be approximately log-normal (with additional concentration of mass at certain amounts). To shed further light on these patterns, we plot in the Appendix (see Figure A1) the distribution of the log size of loan amounts, which should have the shape of a normal distribution if the distribution in Figure 2 was log-normal. While the resulting distribution is relatively symmetric, the figure makes more evident the high concentration of mass around certain specific amounts, such us $1,000, $10,000, $100,000, and $1 million. Some of these spikes in the distribution are likely to be explained by the preponderance of small-amount test loans, but there may also be a tendency from banks to borrow round-number amounts.

To gain some perspective on the financial stakes at play when a bank decides to borrow from the primary-credit program, we calculate

Figure 2 Distribution of Loan Sizes. Primary Credit

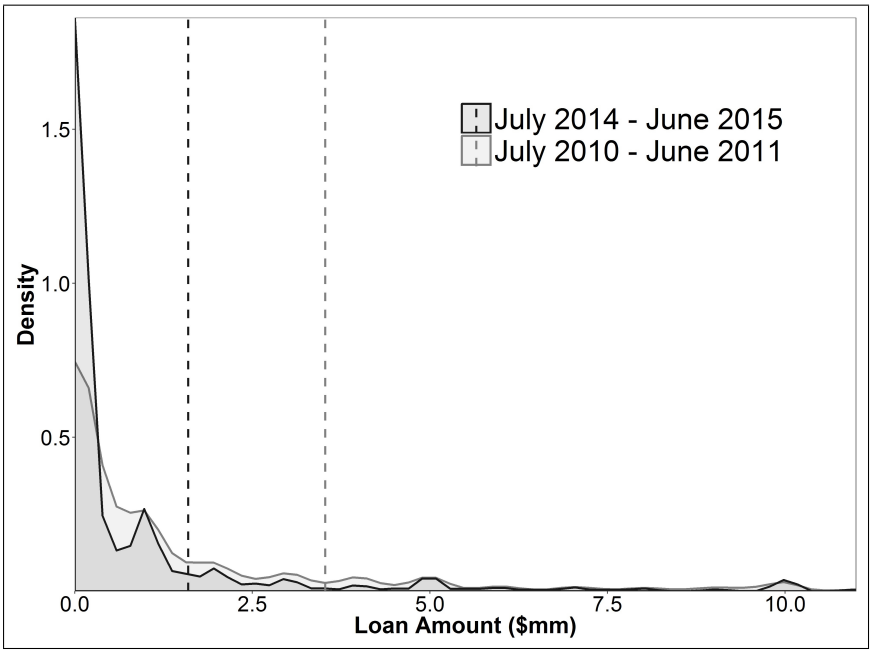


Note: We plot the distribution of loan sizes for loans of $10 million or less. We omit loans greater than $10 million from the plot to reduce the length of the right tail. We then scaled the distribution by the proportion of loans less than $10 million. There are 436 loans that are greater than $10 million in the sample. The vertical dashed line represents the mean of all loans ($2.3 million).

the interest expense on a $10 million overnight loan at the interest rate relevant for all the primary-credit loans in our sample, 75 basis points. The amount charged is approximately $200 in interest payments per day. Given that almost 90 percent of primary-credit loans are overnight loans (see Table 6), we can conclude that a significant portion of the activity at the primary-credit facility constitutes relatively small loans with no significant interest cost for the borrower.

However, there are also some very large primary-credit loans being made in the period under consideration. For example, the largest primary-credit loan in the sample was for $1.02 billion and was granted by the St. Louis Fed on Wednesday, November 24, 2010, to First Tennessee Bank, N.A. The term of the loan was two days, which is relatively unusual, as most primary-credit loans are overnight. First Tennessee

Figure 3 Distribution of Loan Sizes–Evolution Over Time



Note: We consider the first and the last twelve-month period in our data. For the period July 2010 to June 2011, we exclude from our density calculations one clear outlier: a loan for (approximately) $1 billion (which is ten times larger than the second-largest loan in this subsample). As in Figure 2, we plot only the portion of the distribution that corresponds to loans of $10 million or less.

Bank had $24.5 billion in assets and $3.2 billion total equity as of December 2010, so the discount-window loan was a significant financial transaction for the bank. The second-largest primary-credit loan in the sample was for $900 million and was granted by the New York Fed to the New York branch of the Belgian bank KBC on Friday, December 9, 2011. The term of the loan was three days and was repaid at maturity on the following Monday.

Evolution of the distribution over time: As indicated in Table 3, total lending at the primary-credit facility was significantly lower toward the end of our sample period relative to what it was at the beginning. It is also interesting to see that the number of loans increased moderately and, consequently, the mean size of primary-credit loans decreased substantially.

Figure 3 plots the distribution of loan sizes for the first twelve months in our sample period (from July 2010 to June 2011) and compares it with the distribution of loan sizes for the last twelve months of that period (from July 2014 to June 2015). This confirms the pattern suggested by Table 3: in the later part of our sample period, the proportion of loans that are small is much higher. Figure A2 in the Appendix provides a more complete picture of the evolution of the distribution of loan size over time.

During 2010 and 2011, the US banking system was still recovering from the financial crisis. Furthermore, economic turmoil in Europe at the time impacted the financial condition of some banks in the US (recall that foreign-bank subsidiaries have access to the Fed's discount window). This may partly explain the elevated levels of primary-credit activity during that period and the relatively large size of the average loan. In contrast, by the second half of 2014, the banking system in the US was in much better financial shape, a fact that is likely to have contributed to reducing banks' demand for emergency credit.

Term to maturity: Primary-credit loans are mainly overnight loans. Loans made on a Friday, however, are held for a minimum of three days and are charged interest accordingly (that is, for at least three days). Furthermore, in certain specific situations, loans are granted for longer periods of time. While primary credit is provided relatively automatically ("no questions asked"), when banks request loans of a longer maturity than overnight, some extra administrative oversight may occur.

Table 6 provides a description of the maturity profile of primary-credit loans in our sample. There are a lot of loans in the two- to four-day maturity term. Many of those are associated with holidays and weekends. Of the 962 loans with a three-day maturity term, 955 were granted on a Friday. Similarly, many (but not all) of the four-day loans involved a weekend, preceded or followed by a holiday.

Table 6 Primary-Credit Loans. Maturity Term

	Term of loan (days)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**12**	**14**
Number of loans	10208	110	962	122	5	3	15	2	1	1
Number of "test" loans	4339	26	269	21	0	0	0	0	0	0
Total amt. lent ($mm)	20528.381	1361.845	3566.472	686.669	6.000	27.000	87.850	8.000	10.000	4.000
Avg. size of "nontest" loans ($mm)	2.011	12.380	3.707	5.628	1.200	9.000	5.857	4.000	10.000	4.000

Note: The term of the loans is expressed in days, and the amounts are all in millions of dollars. We call "test" loans all loans for amounts less than or equal to $10,000.

We see in Table 6 that there are a few loans lower than or equal to $10,000 and with maturity terms higher than one day. This may be indicative of the fact that not all loans in that size range constitute test loans, in the sense that it seems unlikely that a bank would choose a Friday to conduct a test given that the interest cost is higher than when the test is done with an overnight loan.

The longest maturity term for a primary-credit loan in our sample was a fourteen-day loan granted by the Dallas Fed to First National Bank of Rotan, Texas, in January 2015. The loan was for $4 million, which is close to the average amount of all overnight loans greater than $10,000 and not much different from many of the other longer-term loans in our sample. Interestingly, the loan was actually paid back early, after nine days.

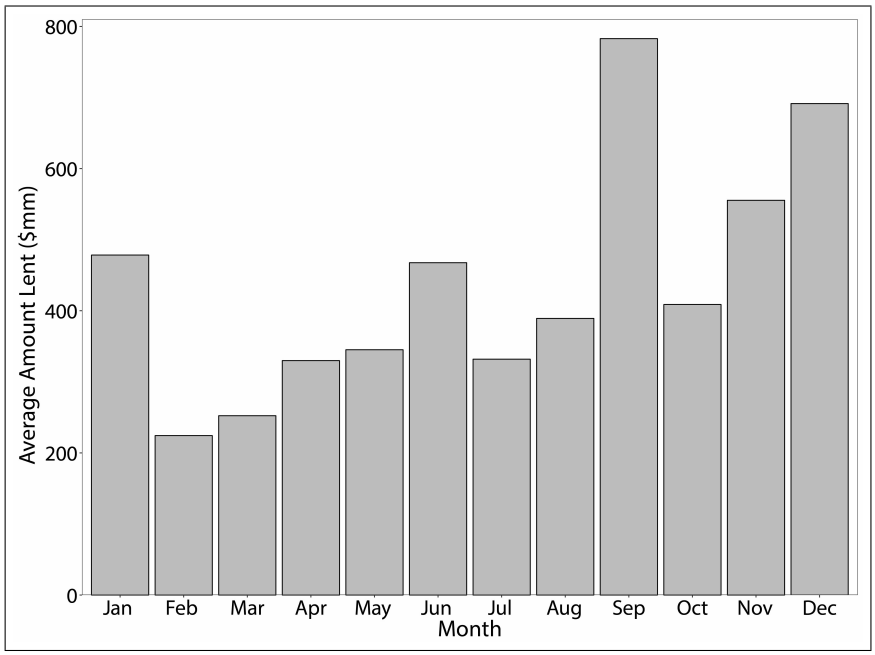
Of the more than 10,000 primary-credit loans in our sample, only ten of them were paid before maturity. We provide a list of these loans in the Appendix (Table A2). The loans that are paid before maturity tend to be relatively large: the average size is $4.8 million, and all but one of them are larger than $100,000. Their maturity is not concentrated in any particular term, and it does not appear to be the case that only the longest maturity loans get repaid earlier.

Seasonality: The average amount lent in each month of the year during our sample period seems to increase as the end of the year approaches. Figure 4 shows that after fairly high levels of total borrowing in January, borrowing slows down in February and only gradually increases as the year progresses.

The high level of lending observed in September is partly driven by an event in September 2012 when a bank in Texas took out several loans of a significant amount during that month. This episode is discussed in more detail below.[3] Even when we remove the large loans to the Texas bank, the month of September exhibits more lending than August and October. In effect, June, September, and December (end-of-quarter months) seem to be months where lending tends to increase. We confirm this below in Table 7.

[3] It is important to realize here that the way these numbers are calculated tends to overweight the importance of short-term loans that get rolled over. As an example, if a bank borrows $100,000 for two days and another bank takes a loan of $100,000 for one day and then another loan of the same amount the following day, the second lending strategy would add up to $200,000 of lending, while the first will only count as $100,000 toward the total amount lent in that month.

Figure 4 Average Total Amount Lent of Primary Credit in Each Month



Note: Each month appears five times in our sample period. For each month, we sum the total amount lent during that month and average that amount among the five corresponding totals. That is, for example, for the month of January we sum all the lending done in January in each of the years in our sample and then take the average over the five Januaries in our sample.

Quarterly and monthly frequency: The first row of Table 7 shows how much primary-credit lending takes place during the beginning, middle, and end days of the month. The pattern suggests that lending is higher at the beginning and end of the month and slows down during the days in the middle. A similar pattern arises when we compute the average amount lent in the first, second, and third month of the quarter (second row of Table 7): lending picks up noticeably toward the end of the quarter.

Daily frequency: Table 8 shows the average total number of loans granted in each day of the week during our sample period and the average number of test loans. We see that banks tend to avoid "testing" on Fridays and other special days such as the end of the month, quarter,

Table 7 Total Amount Lent at Different Times of the Month and the Quarter ($mm)

	Beginning	Middle	End
Month	136.316	111.110	194.801
Quarter	387.220	378.519	548.571

Note: For the monthly calculations, we divide the month into three periods of (approximately) ten days each and call the first ten days the beginning of the month, the next ten days the middle of the month, and the last ten days the end of the month. For the quarterly calculations, the beginning of the quarter is the first month of the quarter, the middle is the second month, and the end is the third month. We calculate the total amount lent in each subperiod, and we report the average across all the corresponding subperiods in the sample.

and year. This makes sense to the extent that conducting a test on a Friday would be more costly.

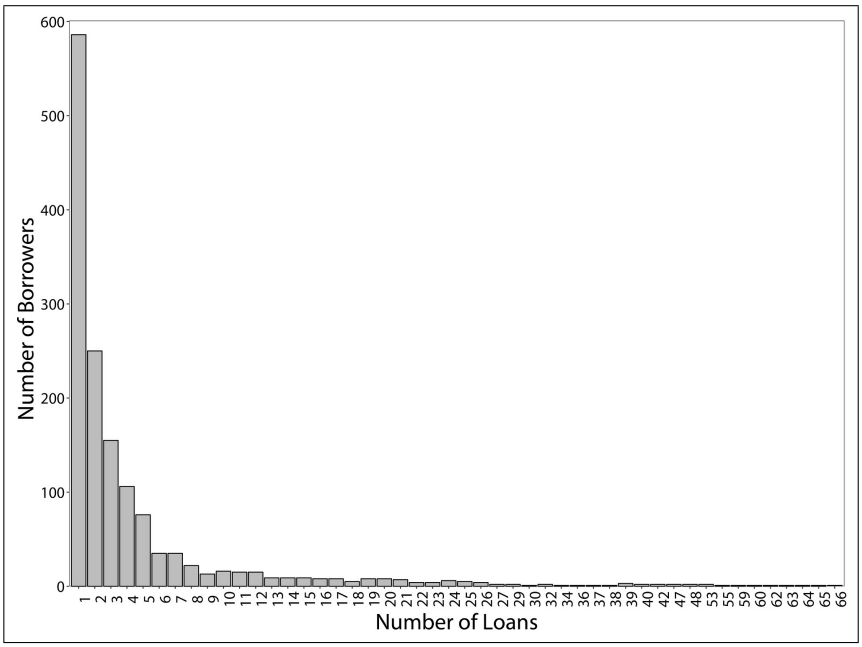
We also report in Table 8 the total amount lent on average in each of those days (third row) and the average size of loans granted each day that were greater than $10,000 (fourth row). The last three columns report similar statistics for the day that corresponds to the end of a month, a quarter, or a year in our sample period.

In general, there is less lending happening on regular Fridays, but the size of the loans are larger on average. Also, the table shows that there is a lot more lending happening on days that are the end of a month, quarter, or year, and the average size of those loans is larger than for other regular weekdays.

Intensity of use: A lot of banks accessed the primary-credit facility during our sample period. Overall, a total of 2,758 different banks (identified by the corresponding ABA number) borrowed at least once from the facility. Of these, 1,756 of them borrowed at least twice, and some banks borrowed over sixty times during our five-year sample. Many of these loans are likely to be tests, of course. For this reason, we also look at intensity of use after restricting our sample to loans greater than $10,000. The total number of different banks that borrowed from the facility (at least once) reduces to 1,450 after we restrict our sample to nontest loans.

Figure 5 presents the histogram of the number of times that individual banks borrowed loans for amounts greater than $10,000 from the primary-credit facility during our sample period.

Figure 5 Histogram of the Number of Primary-Credit Loans by Individual Banks



We compute the number of individual nontest transactions (loans greater than \$10,000) that each individual bank entered at the primary-credit facility and plot the histogram of these numbers; that is, each bar in the histogram denotes the number of banks that borrow n amounts of times from the primary-credit facility, with n = 1,2,3... up to sixty-six, which is the maximum number of nontest primary-credit loans taken by an individual bank in our sample during the five-year period under consideration. We identify a bank by its ABA number because it is a more consistent identifier than the bank's name.

We see in the figure that almost 600 banks took one "nontest" loan during the period, but there were also twenty-eight banks that took thirty or more "large" (nontest) loans from the discount-window primary-credit program in the span of five years. In other words, some banks appear to be relatively frequent users of the facility.[4]

[4] Even if we restrict attention to loans greater than \$100,000 (given that it is possible that some Reserve Banks, such as Dallas, ask for test loans of an amount equal to \$100,000), the number of banks that borrowed at least thirty loans greater than \$100,000 in the space of five years is equal to twenty-six.

Table 8 Total Amount Lent in a Day (Average Over Sample)

Averages	Mon.	Tues.	Wed.	Thurs.	Fri.	End of month	End of quarter	End of year
Number of loans made	11	11	10	10	4	8	9	12
Number of "test" loans	4	4	4	4	1	1	1	0
Daily amount lent ($mm)	23.773	18.502	26.968	20.444	16.499	42.892	57.493	98.217
Size of nontest loans ($mm)	3.806	2.851	4.603	3.554	5.392	6.581	7.141	8.184

Note: We identify the number of days of each day category (Monday, Tuesday, etc.) when the discount window was open during our sample period. We then count the number of loans and the amount lent in each of those days and report the average for each day category (taking into account that, for example, most of the holidays—for which the discount window is closed—fall on a Monday). We also compute the average size of the nontest loan made in each day category, where by "nontest" we mean loans greater than $10,000.

In terms of intensity of use of primary credit, one noticeable case is the situation of Texas Capital Bank, N.A., which in September 2012 borrowed significant amounts overnight for several consecutive days from the Dallas Fed, with balances as high as $250 million (on September 5). Between October 2011 and January 2013, Texas Capital Bank borrowed nineteen times, and in ten of those instances the loans were for an amount larger $100 million (See Table A3 in the Appendix for details).

While the calculations used to produce Figure 5 can be informative, they are not able to reveal some particularly relevant details of the lending patterns observed in the data. For example, some institutions may be rolling over a loan for several days with each loan counted as a new loan while, in a sense, the credit event could be considered to be just one event. Other institutions, instead, may be "repeat users" in that they periodically take an overnight loan and repay it at maturity (and only take another loan after some time). Additionally, some institutions may be repeat users, yet of loans of a very small amount.

There are, of course, many ways to present information that speaks to these issues. We pursue several ones here. For example, we compute the number of primary-credit loans that were the result of a rollover of a previous loan. There were 1,731 loans that were followed by a loan by the same institution on the day that the first loan matured. This amounts to approximately 15 percent of all primary-credit loans in our sample.

Of the loans that were rolled over, only 302 were rolled over for the same dollar amount. Figure 6 plots the amount of the original loan and the amount of the subsequent loan (the rolled-over amount) to get a sense of whether or not loans get rolled over to smaller amounts (as a way to pay down the loan gradually). To make the figure readable, we only plot those loans that were $20 million or less.[5]

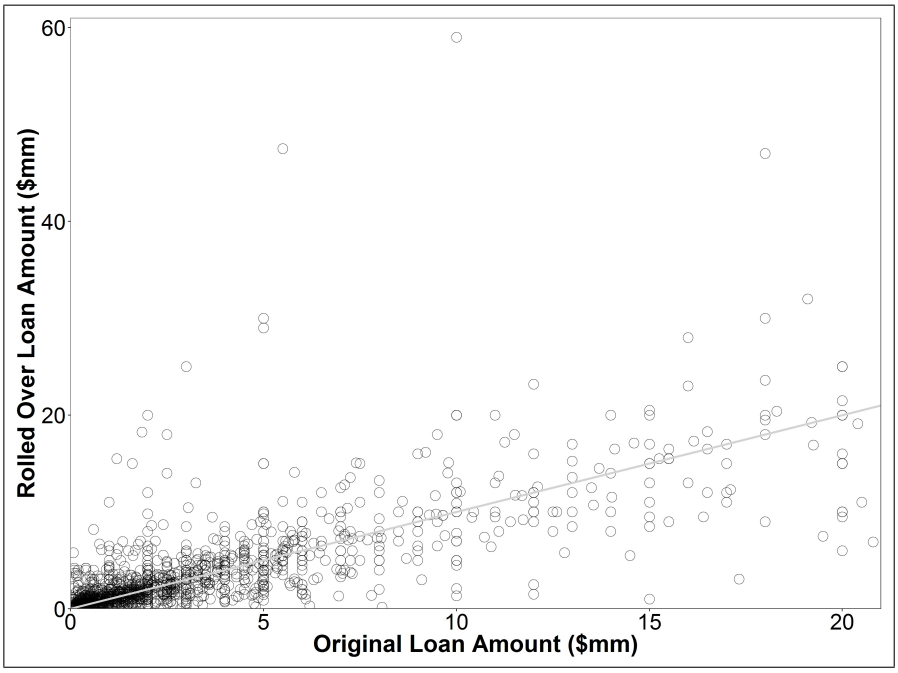
The figure shows that there is no clear pattern: some loans get rolled over into larger amounts, some into smaller amounts, and some into the same amount (these last are represented by the dots that fall exactly over the forty-five-degree line plotted in the figure).[6]

Of the 11,429 primary-credit loans, 9,000 of them were loans that were not followed by another loan from the same institution on the maturity day (that is, they were not rolled over). Of these loans, 4,650 were loans of an amount less than or equal to $10,000. Given that, in

[5] Note also that loans that were initially larger than $20 million but were rolled over into amounts smaller than $20 million do not appear in the figure.

[6] We looked at larger loans than $20 million, and there is no indication that larger loans produce a more definitive pattern.

Figure 6 Rollovers in Primary Credit



Note: The solid diagonal line is the forty–five-degree line. We only consider consecutive loans that were originally for an amount less than or equal to $20 million.

total, there were 4,655 primary-credit loans of an amount lower than or equal to $10,000, that tells us most of those smaller loans, predominantly "test" loans, were not followed by a consecutive loan.

If we call a credit event a sequence of consecutive loans taken by a bank where one loan is followed by another loan (of potentially a different amount) at the time of the maturity of the previous loan, then there were 350 events involving one rollover (see Appendix), 146 events involving two rollovers, and there are events of all numbers of rollovers up to sixteen, the maximum observed in our sample (only one event involved sixteen consecutive loan rollovers).

The average size of the loans that get rolled over at least once is $4.7 million—somewhat larger than the average for all primary-credit loans with amounts greater than $10,000, which is equal to $3.8 million, as reported in Table 2. This indicates that banks that end up rolling over primary-credit loans tend to borrow larger amounts.

We also look at situations where a bank takes a loan that we can consider an isolated situation, in the sense that the same bank does not take a loan in the two months around the loan in question (one month prior and one month after). The number of isolated loans identified this way is 7,078, but presumably many of them were taken as test loans. Indeed, only 2,871 of these isolated loans were for an amount greater than $10,000. This means that within the subsample of loans for amounts greater than $10,000, 42 percent of them were isolated. When we look at this subsample of loans (i.e., loans that are greater than $10,000 and isolated), the average size was $3.7 million, which is similar to the average size of all loans greater than $10,000 (Table 2). In other words, isolated loans that are unlikely to be test loans do not tend to be much different in size than other nontest loans.

3. SECONDARY CREDIT

Secondary credit is used much less often than primary credit. There are 650 loans in our five-year sample, of which only thirty-nine are greater than $10,000. Virtually all secondary-credit loans were overnight loans. The volume in the secondary-credit program has been trending down over the years under consideration, with the total amount lent and the number of loans both decreasing over time (see Table 3).

The number of test loans has probably been decreasing as a result of banks moving out of the "poorly capitalized" category, either by becoming eligible for primary credit—as banks (and the economy) exit the crisis state—or by exiting the industry (through mergers or liquidations). Similarly, the intensity of use of the secondary-credit program, beyond just testing, is also likely to be strongly influenced by prevailing financial conditions: as those improved over the years in the sample period, usage declined.

Loan Sizes: In general, secondary-credit loans tend to be smaller than primary-credit loans. The largest loan at the secondary-credit facility during our sample period is much smaller than the largest primary-credit loan ($17 million and $1 billion, respectively). Except for a loan granted by the Philadelphia Fed to Nova Bank in December 2010 for $17 million, all other secondary-credit loans in our sample were under $10 million in value.

This pattern is stronger toward the later part of our sample period. Indeed, since the end of 2012, there has been very little lending of significant amounts at the secondary-credit program. If we consider only larger loans, using a $1 million threshold, then there was only one loan at or above that amount in 2013, three in 2014, and again only one in the first half of 2015. For the whole sample, only twenty-nine

Table 9 Secondary-Credit Loans. Maturity Term

	Term of loan (days)			
	1	**2**	**3**	**4**
Number of loans	587	4	48	11
Number of "test" loans	558	3	43	7
Total amount lent ($mm)	73.249	17.003	6.010	22.716
Average size of "nontest" loans ($mm)	2.489	17.000	1.180	5.675

Note: The term of the loans is expressed in days, and the quantities are all in millions of dollars.

loans were for amounts greater than or equal to $1 million. Clearly, secondary credit is used in a meaningful way outside crisis periods only in rare occasions.

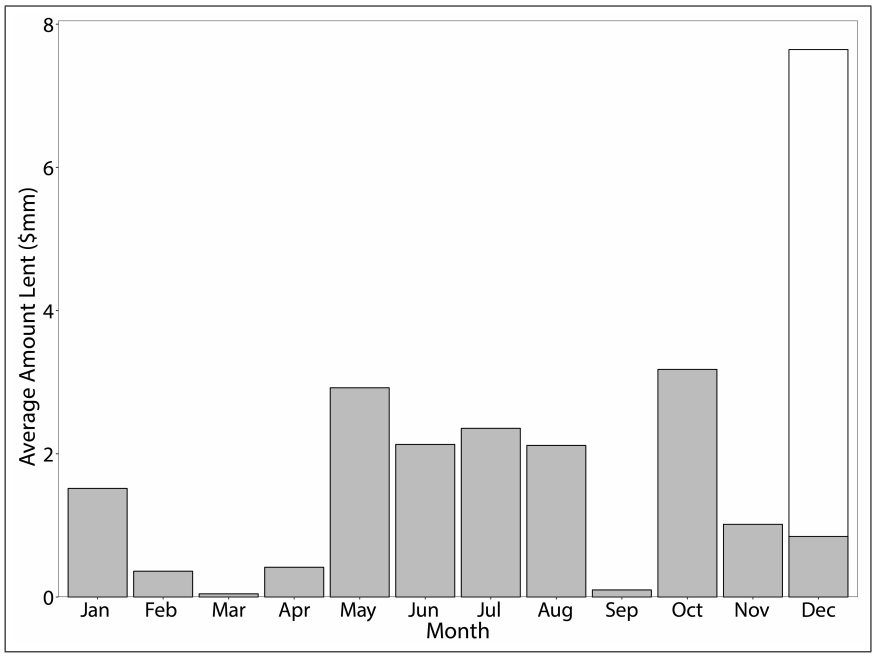
Term to maturity: Most secondary-credit loans have a one- or three-day maturity term (and all the three-day loans were taken on Fridays). Contrary to primary credit, there are no secondary-credit loans of longer maturity than four days (see Table 9) and all four-day loans were made on Fridays followed by a holiday. Potentially, the extra scrutiny by loan officers at Reserve Banks, and the fact that the financial institution seeking the loan is not in obviously sound financial condition, drives this pattern of low maximum maturity length.

When reading Table 9, it is worth noting that there are only four secondary-credit loans of two-day maturity (middle column in the table), of which three of them are for $1,000, and one of them is for $17 million. This fact drives the high average value of loans of the two-day maturity term but it seems reasonable to attribute the anomaly to the small sample size.

The $17 million two-day loan is unusual in that it is the only secondary-credit loan in our sample that was charged an interest rate higher than the standard 125 basis points (that is, fifty basis points higher than the primary-credit rate). The loan was granted by the Philadelphia Fed to Nova Bank, of Berwyn, Pennsylvania, on December 28, 2010, and the interest rate charged was 6.25 percent. Interestingly, this loan was a rollover loan from a loan of the same amount granted for four days on December 24, 2010, at the standard interest rate of 125 basis points. There are no other loans to Nova Bank around that time, which suggests that the bank paid the loan in full on December 30, 2010.[7]

[7] On Friday, October 26, 2012, the Pennsylvania Department of Banking and Securities closed NOVA Bank, and the FDIC was named receiver.

Figure 7 Average Total Amount Lent of Secondary Credit in Each Month



Note: This figure is constructed in the same manner as Figure 4. Each month appears five times in our sample period. For each month, we sum the total amount lent during that month and average that amount among the five corresponding totals. That is, for example, for the month of January, we sum all the lending done in January in each of the years in our sample and then take the average over the five Januaries in our sample. For the month of December, the grey bar represents the average after we exclude the two large loans to Nova Bank in December 2010. The white bar is the average including those two large loans.

The initial loan for $17 million to Nova Bank on December 24, 2010, is also responsible for the high average size of the four-day loan category (see last column of Table 9). Excluding that loan, the average four-day loan size is much lower ($1.9 million). Here, again, the small sample size is a significant limiting factor, but it is worth pointing out that this amount is smaller than the average size for all overnight secondary-credit loans ($2.489 million).

Seasonality: Figure 7 shows that secondary credit is concentrated around the middle of the year, with activity being very moderate in the initial months of the year. The December average is heavily influenced

Table 10 Total Amount Lent at Different Times of the Month and the Quarter

	Beginning	Middle	End
Month	0.464	0.384	1.149
Quarter	1.866	1.603	2.479

Note: We use the same methodology as used in the construction of Table 7.

by the large loans to Nova Bank in December 2010. After excluding these two loans, the average total amount lent in December at the secondary-credit program is similar to the average level for November and relatively small compared with the middle months of the year. The sensitivity of the December average to the exclusion/inclusion of two loans, however, underscores the fact that the sample for secondary credit is relatively small and, as a result, idiosyncratic events may be biasing some of the statistics computed here.

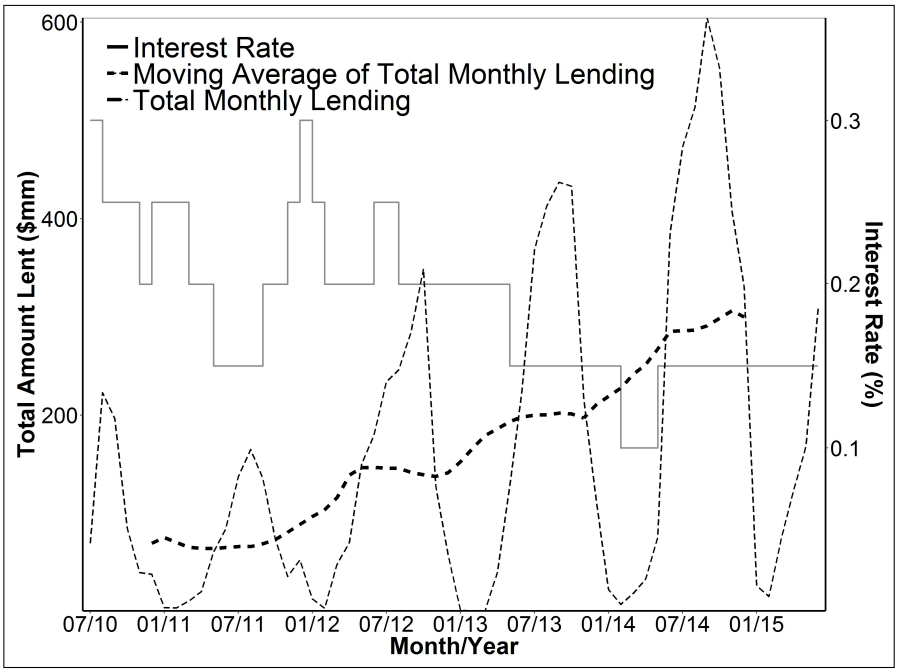
Just for completeness, Table 10 shows the average level of lending at the secondary-credit program at the beginning, middle, and end of the month and of the quarter (to compare it with Table 7). The pattern shows that lending tends to increase at the end of the month and the quarter.

We do not compute a table comparable to Table 8 in the primary-credit section because the subsamples needed to deal with daily patterns (given the presence of test loans) are too small to draw any meaningful inferences. For computing total amounts lent, the test loans play very little role: for example, the 611 test loans (equal to or lower than $10,000) at the secondary-credit program account for $1.2 million of the total $119 million lent over the sample period (see Table 1). Effectively, then, the total amounts lent reflect the thirty-nine nontest loans in the sample. As we split the sample into subsamples, the number of (nontest) loans in each category becomes too small to make any reliable inference.

4. SEASONAL CREDIT

Even though seasonal credit presents a strong annual seasonal pattern, it is clearly discernable from Figure 8 that there has been an upward trend in the total amount lent over the years of our sample. This is also evident in Table 3. Of course, the strong seasonal pattern is not surprising given the objective of this program: to provide ample access

Figure 8 Total Monthly Seasonal Credit ($ million)



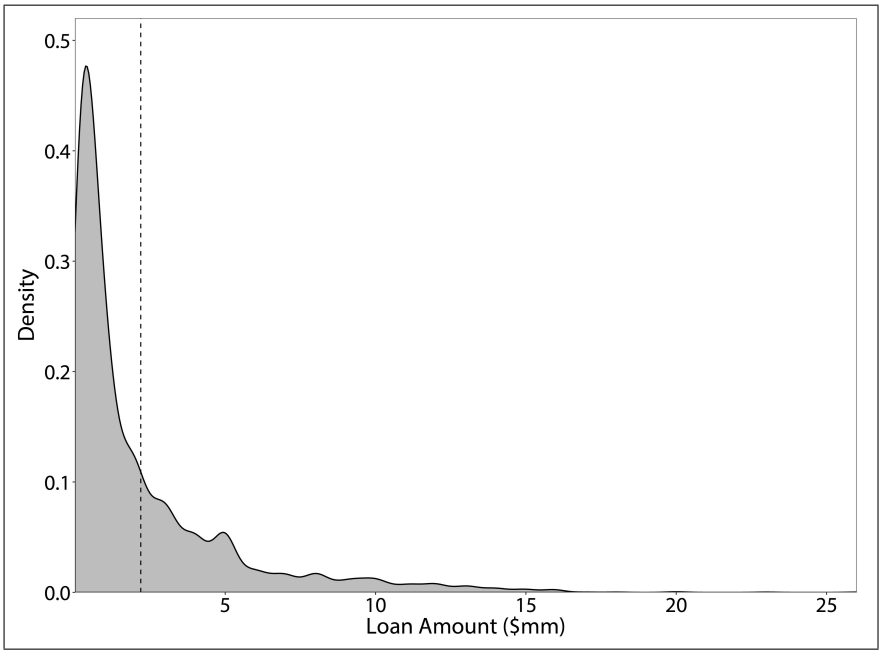
Note: For each month in our sample, we aggregate all loans in that month to generate the monthly time series plotted as the lighter dashed line. We also compute a symmetric, twelve-month moving average, which is the darker dashed line. We overlay the most frequent (discount window) interest rate in the period as the solid line.

to credit to smaller institutions with a predictable (and demonstrable) seasonal pattern in their funding needs.

We included in the figure the interest rate charged for seasonal credit. The rate is calculated as an average of market rates, and it has fluctuated over the years while exhibiting a moderate downward trend. It may be the case that the gradual decrease in the cost of borrowing over the sample period is partly responsible for the (also gradual) increase in lending shown in the figure (although, of course, the cost of other sources of funding probably move together with market rates as well, making those other sources of funding also more attractive).

Loan Sizes: The distribution of loan sizes is plotted in Figure 9. As we know from Table 2, the median size of seasonal-credit nontest loans is similar to the median for primary-credit nontest loans, but the

Figure 9 Distribution of Loan Sizes. Seasonal Credit



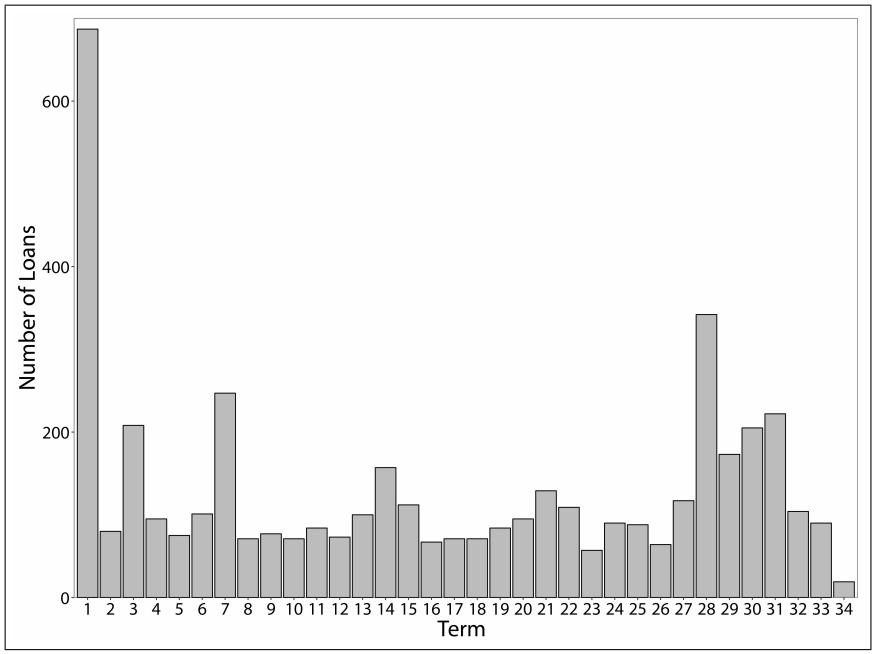
Note: We plot the distribution of loan sizes for all seasonal-credit loans in the sample. The dotted vertical line is the mean.

maximum loan amount is much smaller. As a result, the distribution shows more density in intermediate values, such as loans between $1 million and $10 million. Since seasonal credit is not provided at a penalty interest rate, the nature of its usage is likely to be very different than for primary credit, with midsize loans being more common (as reflected in the distribution of loan sizes).

Testing also seems less common in the case of seasonal credit. In fact, there are no seasonal-credit loans in our sample for an amount equal to or less than $1,000 and eleven loans for an amount between $1,000 and $10,000. This is of course in sharp contrast with the patterns observed in primary and secondary credit where a large proportion of the loans fell within that range of very small amounts.

Term to maturity: Aside from the difference in the distribution of loan sizes between primary, secondary, and seasonal credit, the maturity of seasonal-credit loans is also much less concentrated around

Figure 10 Seasonal Credit Loans. Maturity Term



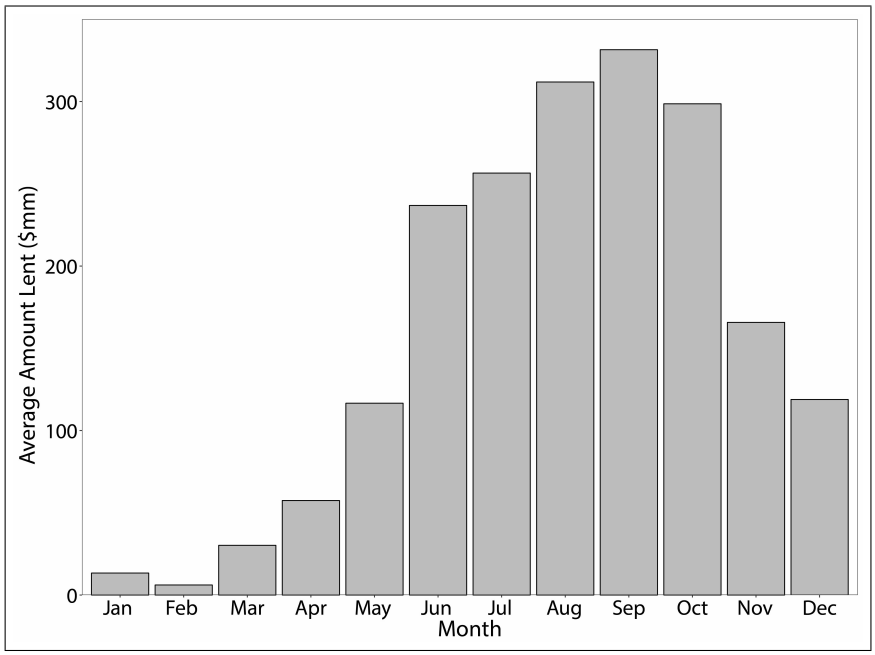
Note: We compute the number of loans of each maturity term from one day (overnight) to thirty-four days, the maximum maturity observed in the sample, and we plot the histogram of the resulting data.

short periods (overnight and three days) than the maturity term of primary- and secondary-credit loans.

Figure 10 presents a histogram of the maturity term of seasonal-credit loans in our sample. Interestingly, loans that are approximately a month long tend to be relatively common among this class of loans. But this relative concentration does not mean that other maturities are not used. In fact, there are many loans for each of the maturity terms in the range between one and thirty-four days. It is still the case, though, that overnight loans are the most common in this subsample, representing 15 percent of the total.

Seasonality: As a complement to Figure 8, in Figure 11 we plot the average amount of seasonal credit provided during each month of the year. The seasonal pattern at the annual frequency is again evident in this figure, with the amount of seasonal credit increasing during the second half of the year and being minimal during the winter months

Figure 11 Average Total Amount Lent of Seasonal Credit in Each Month



Note: See note for Figure 4.

of January and February. While one might suspect that some of this credit follows the agricultural cycle, it is interesting to note that some of the Reserve Banks in agricultural areas, such as St. Louis and Kansas City, are not large providers of seasonal credit (see Table 5).

5. COLLATERAL

Depository institutions can choose to enter credit agreements with their respective Reserve Banks, allowing them to borrow from the discount window if they need to do so at some point. Not all institutions enter credit agreements with their Reserve Bank. When they do, as part of such agreements, depository institutions pledge collateral with Reserve Banks, a process that requires the submission of all the relevant information that would allow loan officers at Reserve Banks to assess the value of the corresponding collateral. Many institutions do not take any loans from the discount window in a given period of time even

when they have an outstanding credit agreement and potentially large amounts of pledged collateral.

Our data only include collateral information as of the time when a bank actually takes a loan from the discount window. In that sense, the information about pledged collateral in our dataset is not comprehensive. To be more specific, for each loan in the sample, the dataset includes information on the value and composition of the collateral available to the borrower with the corresponding Reserve Bank at the time of the loan. The reported value of collateral is adjusted using the appropriate margins (haircuts) so that, in principle, the borrower could receive a loan for the total amount of the reported collateral.

To get a sense of how collateral is being used in these credit relationships, we look at the average composition of the collateral and the level of utilization (how much lending actually occurs relative to the total amount of collateral available). For this latter calculation, we use the total amount of loans outstanding, including the new loan, which is also reported in our data. The collateral pledged at the discount window can be used in any of the discount-window programs so, for example, a bank can take a loan at the primary-credit program and another loan, simultaneously, at the seasonal-credit program. The reported amount of collateral in our dataset is usable in both types of loans.

None of the banks that took a secondary-credit loan during our sample period had any other outstanding loans from the discount window. Eligibility criteria for borrowing at the seasonal credit porgram generally rule out secondary-credit institutions. For those banks taking loans at the primary-credit program, instead, having loans outstanding is more common but still not very prevalent.

The situation is much different for those banks taking seasonal-credit loans, as more than 40 percent of those banks have loans outstanding at the time of taking the loan reported in our dataset. While most of the outstanding loans are previous seasonal-credit loans, we also verified that some banks were borrowing simultaneously from the primary and seasonal programs during our sample period.

Collateral composition. The composition of collateral varies systematically across credit programs, as reflected in Table 11. For primary-credit loans, almost 50 percent of the collateral is accounted for by commercial and consumer loans, with high-quality securities such as US Treasuries and agency debt and mortgage-backed securities taking roughly another 20 percent of the total. Loans granted under the secondary-credit program, instead, have almost 60 percent of the collateral in the form of commercial real estate (CRE) loans. Finally,

Table 11 Type of Collateral (as a Proportion of Total)

Type of collateral	Primary credit	Secondary credit	Seasonal credit
Commercial loans	0.277	0.135	0.422
Consumer loans	0.199	0.046	0.002
CRE loans	0.069	0.585	0.281
US Treasury/Agency securities	0.067	0.070	0.091
Agency MBS	0.123	0.057	0.092
ABS	0.083	0.039	0.000
Other	0.182	0.068	0.112

Note: For each individual loan taken by a bank, the data include the total amount of collateral available to the borrowing bank and its composition. We aggregate across all loans in our dataset and compute the proportion of each collateral component. There are other categories of collateral not reported in the table and for which the proportions were relatively small. These categories are all aggregated under the label "other." Note that if a bank takes several loans, its collateral is counted multiple times in the aggregation, one time for each time that the bank took a loan.

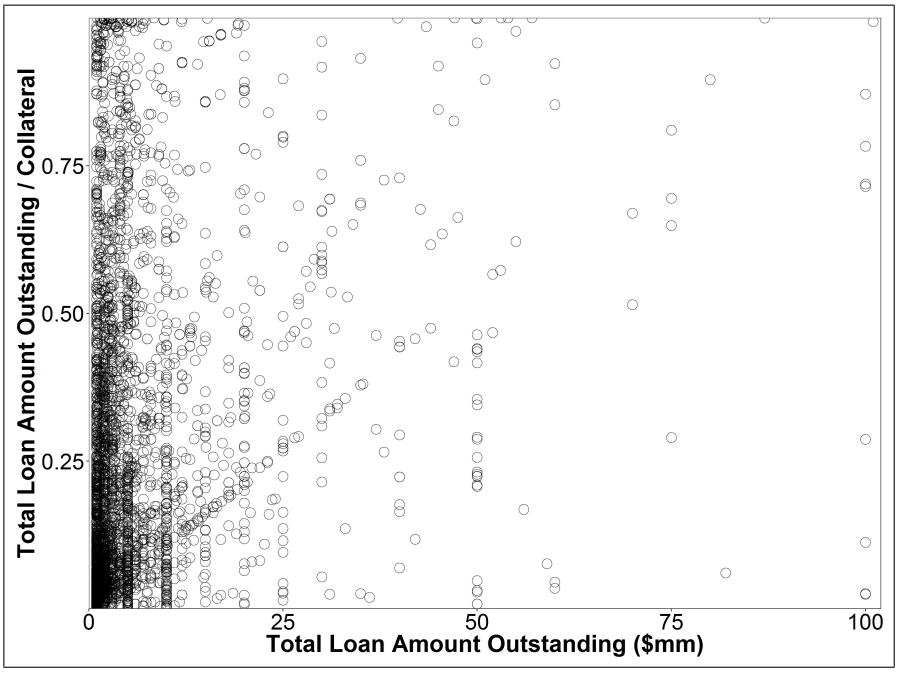
the collateral of banks taking loans in the seasonal-credit program is mainly composed of commercial loans and CRE loans.

Collateral utilization. We measure collateral utilization as the ratio of total outstanding loans for a given bank and its total collateral pledged with the corresponding Reserve Bank. To understand collateral utilization, it is important to keep in mind that the large proportion of test loans in the primary- and secondary-credit programs, by their nature, tend to use a very low proportion of the collateral available to the borrowing bank.

For primary credit, we concentrate attention on loans that are greater than $1 million, since the predominance of test loans can be expected to be very low at those levels of borrowing. Figure 12 shows the level of utilization in the vertical axis and the amount of outstanding loans in the horizontal axis (we plot outstanding amounts up to $100 million to make the figure more readable, but the pattern is similar for all loans over $1 million).[8] We can see from the figure that (contrary to what might be expected) there are a lot of large primary-credit loans that use a significant proportion of the collateral available

[8] The amount of outstanding loans includes the amount of the loans for which the data are being reported. For this reason, the amount of outstanding loans is always greater than $1 million in the figure. Since, for this restricted subsample, only thirty-three out of 2,829 loans were made when previous loans to the borrowing bank were outstanding, the figure basically reflects the utilization ratio calculated using the amount of the current loan.

Figure 12 Collateral Utilization Ratio. Primary Credit



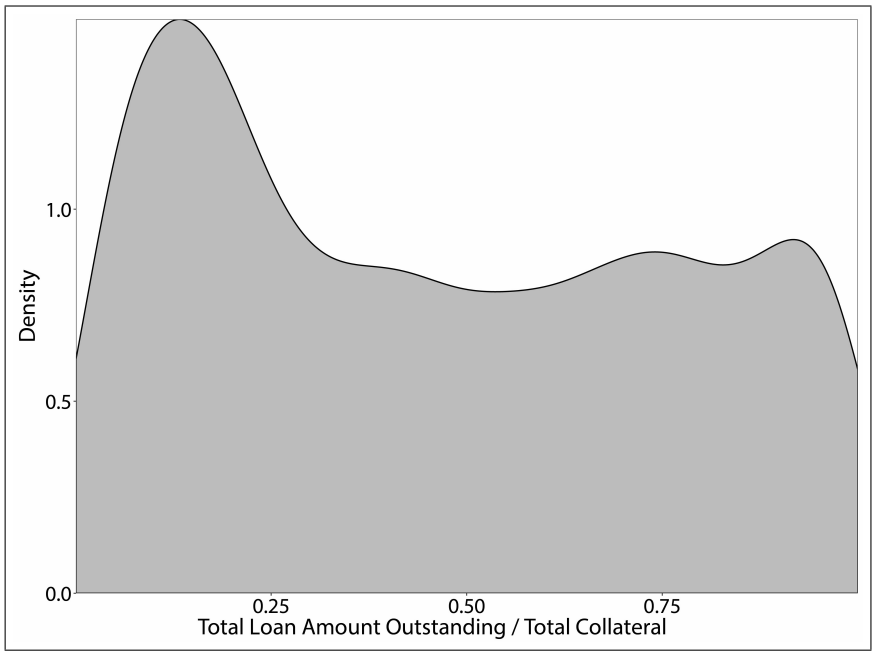
Note: We consider loans at the primary-credit program that, added to the outstanding loans, sum to an amount between $1 million and $100 million. For each loan, we compute the ratio of the total loan amount outstanding and the total collateral available to the borrower. For each loan, then, we plot the level of utilization on the vertical axis and the outstanding loan amount on the horizontal axis.

to the borrowing bank. Smaller loans also can have relatively high utilization ratios. Furthermore, the figure suggests that there is no strong correlation between the size of the loan and the level of collateral utilization.

For secondary credit, we restrict attention to loans greater than $10,000, which in our sample amount to a total of thirty-nine loans. Of those thirty-nine loans, nineteen of them used more than half of the collateral available to the borrower. Collateral utilization for the thirty-nine loans is not concentrated in any particular value and instead is (roughly) evenly spread in the unit interval.

Just as with primary-credit loans, for secondary-credit loans there is no indication of a tight correlation between size of the loans and

Figure 13 Distribution of Collateral Utilization Ratio. Seasonal Credit



Note: For each loan, we compute the ratio of the total loan amount outstanding and the total collateral available to the borrower.

their collateral utilization. The two largest loans in this subsample of thirty-nine loans—those made to Nova Bank in 2010 for $17 million each—have a collateral utilization ratios of 55 percent. On the upper end of the utilization margin (for loans greater than $10,000), there are three loans with utilization ratio greater than 90 percent.

For seasonal credit, collateral utilization is spread (roughly) evenly over values between zero and one. Figure 13 shows the density of the collateral utilization ratio for seasonal-credit loans. Since test loans seem to be much less common in the seasonal-credit program, we include all loans in the figure. For seasonal credit, it is critical to measure utilization using outstanding loan amounts, given that in most cases these loan amounts can be much higher than the amount of the new individual loan that originates the reporting.

6. CONCLUSION

In the five years of now-available discount-window transactions data, from mid-2010 to mid-2015, the Federal Reserve made over 16,000 loans for a total amount of more than $36 billion. Most of the lending was done through the primary-credit program. The seasonal-credit program was also a significant source of funding for banks during our sample period. In the span of those five years, the amount of primary credit shows a slow secular decline and the amount of seasonal credit has been gradually trending up but with large seasonal swings.

A significant proportion of the transactions reported in the data are likely to be "test" loans. These test loans tend to be for very small amounts and do not have a significant impact on the total amounts lent when aggregated across all loans. Still, for understanding the typical size of loans and other relevant aspects of the data, it is helpful to minimize the influence of test loans in the results. We choose a threshold of $10,000 and exclude the loans at or below that threshold when we want to focus on the characteristics of nontest loans. In the case of secondary credit, this procedure leaves us with a very small number of nontest loans, which reduces our ability to draw robust inferences.

In contrast, for the case of the primary-credit program, the sample of nontest loans is large. The size of primary-credit loans is relatively widespread with some very large loans present in the sample. Most loans are overnight, but there are loans with up to a fourteen-day maturity term. A lot of banks accessed the primary-credit program during our sample period, and many of those banks used the program on several occasions. Primary-credit loans are more common at the end of the month, the quarter, and the year, but in general there is credit extended at all times. Most of the primary-credit loans constitute one-time events, in the sense that they do not get rolled over. Those loans that do get rolled over tend to be somewhat larger on average and do not always get rolled over into equal or smaller amounts (sometimes they are rolled over into larger amounts).

Very large secondary-credit loans are rare. Abstracting from test loans, there are actually not many loans being granted through this program. Comparing secondary-credit loans with primary-credit loans, we see that the former tend to be smaller and of shorter term. This is consistent with the fact that those are loans offered at a higher interest rate, and the corresponding borrowers are subject to more supervisory scrutiny.

Consistent with its name, the volume of seasonal credit presents a strong annual seasonal pattern with lending generally picking up significantly in the second half of the year. Test loans are a lot less common in the seasonal-credit program, and loan size is more uniformly

spread in the range of $1 million to $15 million. The maturity term is also widespread between one and thirty-plus days.

Collateral utilization is not concentrated around a certain level and does not tend to reflect a correlation with the size of the loans. Because collateral is measured adjusted for the appropriate haircuts and the utilization ratio is below unity for most of the loans, the level of credit risk involved in these transactions seems likely to be fairly low.

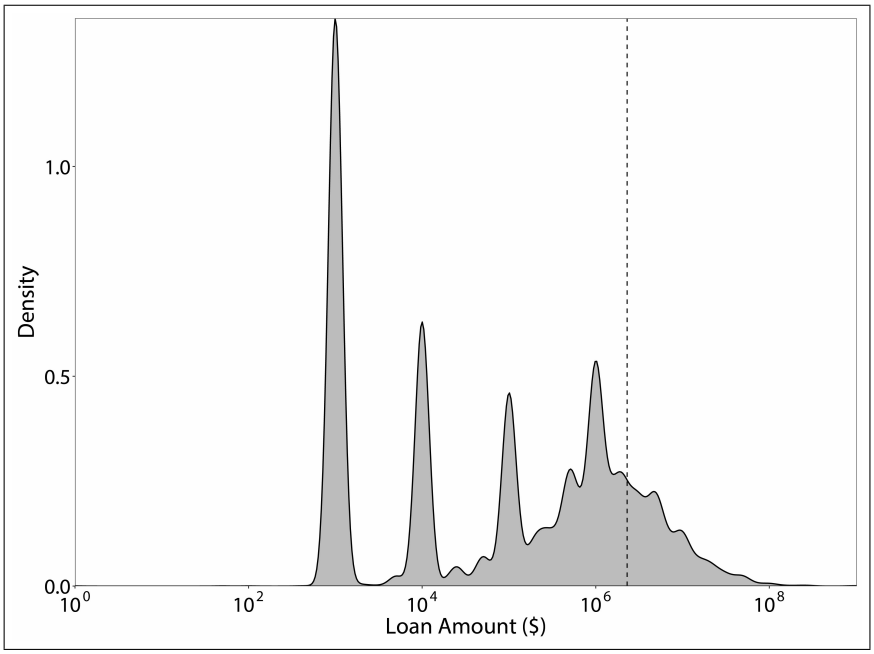
It is, of course, difficult to determine the appropriate amount of discount-window credit needed by US banks. Those needs are likely to also depend on the state of aggregate financial conditions. The period we study includes some episodes of heightened financial turmoil (such as at the peak of the European crisis in 2011) but not a full-blown financial crisis (in the US). Also, during our sample period, banks were holding significant amounts of excess reserves (and liquidity more generally), which likely reduced banks' needs for emergency funding. Still, the data we have analyzed here show that many banks do access the discount window regularly, suggesting that the routine provision of backup funding by the central bank is a valuable option for many participants in the US banking sector.

APPENDIX

Table A1 Annual Average of Daily Loan Amount Outstanding ($mm)

Year	Primary credit	Secondary credit	Seasonal credit	Total
2003	34.038	1.264	62.698	98.000
2004	42.115	0	110.923	153.038
2005	52.673	4.212	142.462	199.346
2006	58.962	0.154	165.538	224.654
2007	479.750	3.192	104.846	587.788
2008	31,817.943	34.642	39.698	31,892.283
2009	40,221.212	192.731	45.115	40,459.058
2010	4,429.154	309.923	39.212	4,778.288
2011	25.942	0.173	35.942	62.058
2012	22.981	0.019	49.442	72.442
2013	13.173	0.019	65.462	78.654
2014	12.830	0	104.755	117.585
2015	17.250	0.019	107.404	124.673
2016	17.808	0	83.615	101.423
2017	13.212	0.019	81.846	95.077

Figure A1 Distribution of Primary-Credit Loan Sizes. Log Amounts



Note: We compute the distribution of primary-credit loan sizes after taking log of each corresponding amount.

Table A2 Primary-Credit Loans Paid Before Maturity

Borrower	Amount ($mm)	Date of loan	Term	Date of repayment	Location	Lender
Community First Bank	0.300	8/5/2010	6	8/6/2010	Fairview Heights, IL	St. Louis
Community First Bank	0.950	8/26/2010	5	8/30/2010	Fairview Heights, IL	St. Louis
Community First Bank	0.050	9/2/2010	5	9/3/2010	Fairview Heights, IL	St. Louis
Community First Bank	0.800	9/7/2010	3	9/9/2010	Fairview Heights, IL	St. Louis
Douglas National Bank	25.000	12/29/2014	7	1/2/2015	Lawrence, KS	Kansas City
First National Bank	4.000	1/13/2015	14	1/22/2015	Rotan, TX	Dallas
First National Bank	1.350	1/5/2015	7	1/12/2015	Chadron, NE	Kansas City
First State Bank	5.000	6/17/2013	8	6/19/2013	Norton, KS	Kansas City
Lakeside Bank	10.000	1/31/2011	4	2/3/2011	Chicago, IL	Chicago
University & Com. FCU	0.700	6/28/2011	3	6/30/2011	Stillwater, OK	Kansas City

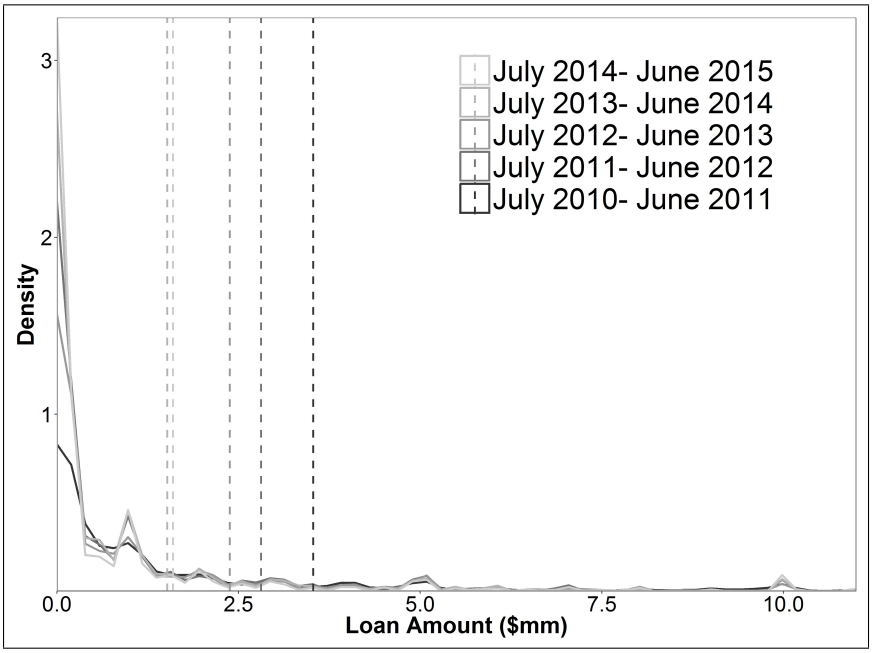
Table A3 Primary-Credit Loans from Texas Capital Bank

Date	Term	Loan amount ($mm)
10/31/2011	1	15.000
11/2/2011	1	60.000
11/15/2011	1	35.000
11/30/2011	1	82.000
12/1/2011	1	7.000
12/30/2011	4	132.000
1/31/2012	1	7.000
4/30/2012	1	115.000
5/1/2012	1	115.000
5/2/2012	1	296.000
5/3/2012	1	7.000
8/31/2012	4	115.000
9/4/2012	1	200.000
9/5/2012	1	250.000
9/6/2012	1	215.000
9/7/2012	3	150.000
9/10/2012	1	160.000
9/12/2012	1	50.000
1/30/2013	1	20.000

Table A4 Loan Rollovers

Number of rollovers (x)	Number of credit events involving x rollovers	Average size of loans
0	9000	1.649
1	350	3.277
2	146	3.631
3	69	6.377
4	37	2.594
5	27	10.366
6	16	6.427
7	18	5.209
8	7	3.101
9	6	8.751
10	7	4.405
11	1	9.575
12	5	3.58
13	4	7.075
14	2	0.85
15	2	3.097
16	1	13.929

Figure A2 Distribution of Primary Loan Sizes. Evolution over Time



Note: We compute the distribution of primary-credit loan sizes for each twelve-month period starting in July 2010. The distribution is calculated after excluding from the data all loans greater than $100 million (fifteen loans in total for all five years). This helps to make the estimated distributions smoother. We then plot only the portion of the distribution that corresponds to loans of $10 million or less.

REFERENCES

Anbil, Sriya. Forthcoming. "Managing Stigma During a Financial Crisis." *Journal of Financial Economics.*

Armantier, Olivier, Eric Ghysels, Asani Sarkar, and Jeffrey Shrader. 2015. "Discount Window Stigma During the 2007–2008 Financial Crisis." *Journal of Financial Economics* 118 (November): 317–35.

Benmelech, Efraim. 2012. "An Empirical Analysis of the Fed's Term Auction Facility." *Cato Papers on Public Policy* 2: 57–91.

Berger, Allen N., Lamont K. Black, Christa HS Bouwman, and Jennifer Dlugosz. 2017. "Bank Loan Supply Responses to Federal Reserve Emergency Liquidity Facilities." *Journal of Financial Intermediation* 32 (October): 1–15.

Courtois, Renee, and Huberto M. Ennis. 2010. "Is There Stigma Associated with Discount Window Borrowing?" Federal Reserve Bank of Richmond *Economic Brief* 10-05 (May).

Drechsler, Itamar, Thomas Drechsel, David Marques-Ibanez, and Philipp Schnabl. 2016. "Who Borrows from the Lender of Last Resort?" *Journal of Finance* 71 (October): 1933–74.

Ennis, Huberto M. 2016. "Models of Discount Window Lending: A Review." Federal Reserve Bank of Richmond *Economic Quarterly* 102 (First Quarter): 1–50.

Ennis, Huberto M., and Alexander L. Wolman. 2015. "Large Excess Reserves in the United States: A View from the Cross-Section of Banks." *International Journal of Central Banking* 11 (January): 251–89.

Furfine, Craig. 2003. "Standing Facilities and Interbank Borrowing: Evidence from the Federal Reserve's New Discount Window." *International Finance* 6 (November): 329–47.

Kleymenova, Anya. 2016. "Consequences of Mandated Bank Liquidity Disclosures." Chicago Booth Research Paper 16-04 (November).

Vossmeyer, Angela. 2017. "Analysis of Stigma and Bank Behavior." Manuscript.

Economic Quarterly—Volume 103, Numbers 1-4—First-Fourth Quarter 2017—Pages 81–137

Using the Richmond Fed Manufacturing Survey to Gauge National and Regional Economic Conditions

Nika Lazaryan and Santiago M. Pinto

Several Regional Banks in the Federal Reserve System conduct regional surveys of business conditions in an effort to obtain real-time information about changes in local economic conditions. To the extent that the performance of the national economy is related to the performance of its regions, the regional surveys may provide useful information about national economic conditions as well. The results of the monthly regional surveys receive attention among analysts and other organizations that assess and forecast economic conditions because they are typically available one or two weeks prior to the release of the national and regional data. Considering how the survey data are used, it is extremely important, first, to understand what the surveys actually measure and, second, to determine how well they measure changes in economic conditions. This paper intends to offer some insights on these issues by carefully analyzing the underlying survey data and investigating their ability to precisely gauge economic conditions observed at both national and regional levels.

The present work focuses on the information content of the Regional Surveys of Business Activity conducted by the Federal Reserve Bank of Richmond (FRBR). We specifically examine the survey that

Federal Reserve Bank of Richmond: nika.lazaryan@rich.frb.org; Federal Reserve Bank of Richmond: santiago.pinto@rich.frb.org. The views expressed in this paper are those of the authors and should not necessarily be interpreted as those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

DOI: https://doi.org/10.21144/eq1030103

tracks the manufacturing sector, the Fifth District Survey of Manufacturing Activity. In general, the surveys collect qualitative data from businesses in the Fifth District on several items that are supposed to convey information about recent changes in economic activity. For instance, survey participants are asked if they have observed an increase, a decrease, or no change in their levels of employment, shipments, orders, and wages, as well as other indicators. The responses are summarized into a statistic, called a diffusion index, that essentially captures the breadth of the changes taking place along the relevant economic dimensions in the time period under consideration. A few individual diffusion indices are combined into a composite diffusion index. In this paper, we evaluate the performance of these individual and composite diffusion indices by examining how closely they track overall national and regional economic conditions.

As a measure of national economic activity, this paper uses the manufacturing diffusion index produced by the Institute of Supply Management (ISM). The individual diffusion indices calculated by the ISM are based on questions that are very similar to those included in the FRBR survey. Previous work, such as Harris et al. (2004), shows that the composite ISM index is a good gauge of national economic activity based on its ability to track the national gross domestic product (GDP) and personal income. Their work also shows that there is indeed a strong correlation between the ISM and the indices produced from regional surveys by the Richmond and Philadelphia Federal Reserve Banks.[1]

Our indicator of regional economic activity in the Fifth District is a weighted average of state payroll manufacturing employment growth (MEG) rates. Only a few papers have attempted to evaluate the accuracy of the diffusion indices produced by regional Reserve Banks in describing economic changes at the regional or local level. The limited work in this area includes Harris et al. (2004) and Pinto et al. (2015b). Even though the analysis in Harris et al. (2004) focuses on the national economy, it also briefly assesses the extent to which the FRBR composite diffusion index helps explain changes in personal consumption expenditures (PCE) in the Fifth District. We use MEG instead of PCE because, among other reasons, the former series are available monthly, whereas the latter are available only at quarterly intervals. Our work is also related to Pinto et al. (2015a). This paper assesses the ability of certain specific individual diffusion indices (employment and wage diffusion indices) to explain employment and wage growth rates. It argues that, in general, growth rates include information both about

[1] It would be useful, in future analysis, to benchmark the FRBR indices against other measures of economic activity, such as the Industrial Production Index.

the "intensive margin," or the size or magnitude of a change, and the "extensive margin," or the spread or breadth of a change. As a result, diffusion indices would describe fairly well how a variable changes over time, if those changes are predominantly explained by a higher proportion of participants reporting a change (up, down, or remain the same). Among other things, this paper shows that while the FRBR employment diffusion index tends to track quite well regional employment growth, it does not do a good job at tracking wage growth.

The analysis performed in this paper departs from the previous work in at least two fundamental ways. First, we carefully examine the behavior of *all* diffusion indices currently reported by the FRBR, with the intention of gaining a much broader understanding of the information conveyed by the FRBR survey. The conclusions of this study may be used to verify whether the FRBR indices capture what they intend to capture and to determine which indices are more informative depending on the specific objective. Second, this exercise offers useful insights that could guide the design and construction of alternative indicators of economic activity and provide feedback on what kind of data to gather (for instance, which survey questions are more relevant, or how representative the survey sample is).

Our analysis proceeds as follows. We first evaluate how well the FRBR composite diffusion index tracks national and regional economic activity, as measured by the ISM diffusion index and the Fifth District MEG, respectively. We next explore the differential contribution of the individual components of the composite diffusion index. Finally, we explore the benefits of including information that is currently available from the FRBR surveys but is not considered in the calculation of the composite index. Our findings show, among other things, that while the reported composite index contains useful information about economic activity at both the national and regional levels, models that incorporate additional survey information may improve the predictive power of the FRBR indices.

The rest of this paper is organized as follows. Section 1 describes the survey and the data used in the analysis. Section 2 evaluates the ability of FRBR diffusion indices to describe the behavior of the national economy. Section 3 examines the relationship between the FRBR diffusion indices and economic conditions in the Fifth District. Section 4 summarizes and discusses the results.

1. PRELIMINARY ANALYSIS OF THE DATA

The work in this paper focuses on the Fifth Federal Reserve District, which includes Virginia, most of West Virginia, Maryland, North

Carolina, South Carolina, and the District of Columbia. Moreover, the analysis is centered on the manufacturing sector. Three pieces of data are used: the diffusion indices produced by the FRBR, constructed from the information collected by the Fifth District Survey of Manufacturing Activity; the manufacturing diffusion indices reported by the ISM, used as a proxy of national economic activity in the manufacturing sector; and Fifth District payroll MEG, used as a measure of regional manufacturing activity. The data are monthly, and the sample covers the period from May 2002 to June 2017.

The FRBR conducts monthly surveys within the Fifth Federal Reserve District states to assess business conditions in two sectors: manufacturing and service. The manufacturing survey is designed to approximate the distribution of manufacturing firms by state, industry type, and firm size.[2] It inquires about various aspects related to the economic conditions faced by firms, including questions on employment, shipments, new orders, backlogs, inventories, prices, etc.[3] The survey is qualitative in nature, in the sense that firms are asked whether they experienced an increase, decrease, or no change in each variable of interest from the preceding month. The responses to each question are then combined into what is referred to as a diffusion index.[4]

The diffusion index calculated by the FRBR is similar to the manufacturing diffusion index reported by the ISM. The latter, however, is based on information collected through a survey of more than 300 purchasing managers of manufacturing companies across the US. This survey is nationally representative and captures the various manufacturing categories by their relative contribution to the GDP.

In general, diffusion indices are summary statistics of the form

$$D_t = 100 \times \left(w^u \frac{N_t^u}{N_t} + w^s \frac{N_t^s}{N_t} + w^d \frac{N_t^d}{N_t} \right), \tag{1}$$

where N_t^u, N_t^s, and N_t^d denote, respectively, the number of survey participants who report that the relevant economic variable has increased, stayed unchanged, or declined from period $(t-1)$ to period t. The

[2] On average, and during the sample period under consideration, the number of respondents in the manufacturing survey has oscillated around 100.

[3] The survey also includes questions on vendors, average workweek hours, wages, business expenditures, inventories of raw materials and finished goods, as well as capacity utilization.

[4] The FRBR currently reports diffusion indices at the Fifth District level. Sample sizes are at the moment too small to report informative diffusion indices at the state level. In this paper, the focus is precisely on the regional diffusion index because, as explained earlier, one of the goals is to determine whether information about the economic performance of the region conveys useful information about the national economy. See Pinto et al. (2015b) for a thorough discussion about the information content of state-level diffusion indices.

FRBR, for example, reports diffusion indices with $w^u = 1, w^s = 0$, and $w^d = -1$, so the range of the index is $[-100, 100]$. Note that, in this case, the diffusion index is simply the difference between the fraction of respondents who reported an increase and the number of respondents who reported a decrease in a particular measure of economic activity. Therefore, a positive (negative) reading indicates that the proportion of participants who report an increase is higher (lower) than the proportion of those who report a decline in the variable. A larger value of the index (in absolute terms) indicates that the change in the economic variable is widely spread out and broadly observed among respondents.[5] The diffusion indices reported by the ISM use $w^u = 1, w^s = 0.5$, and $w^d = 0$. The range of this index is $[0, 100]$, so in this instance, the series are centered at 50: a reading of the index above (below) 50 indicates that the percentage of responses reporting an increase is higher (lower) than the percentage of responses indicating a decline. Both the FRBR and ISM also report composite indices that consist of a weighted average of several individual diffusion indices, each one tracking different indicators of economic activity. Specifically, the composite index reported by the FRBR is given by

$$RIC_t = 0.27 \times RIC_t^E + 0.33 \times RIC_t^S + 0.40 \times RIC_t^O, \tag{2}$$

and the ISM composite index by

$$ISM_t = 0.20 \times \left(ISM_t^E + ISM_t^S + ISM_t^O + ISM_t^P + ISM_t^I\right), \tag{3}$$

where RIC_t^i and ISM_t^i are the FRBR and ISM individual diffusion indices for category i, and the superscript i stands for E: employment, S: shipments, O: orders, P: production, and I: inventories.[6] When comparing the FRBR diffusion index to the ISM, we normalize the FRBR diffusion index in order to compare both indices on the same scale. Finally, to measure regional economic activity in the manufacturing sector, we use the series of payroll MEG obtained from the Bureau of Labor Statistics (BLS). We calculate the Fifth District MEG as a weighted average of the states' MEG.[7] For the purposes of our analysis, the closest counterpart to the manufacturing FRBR diffusion indices is regional MEG, not only because MEG closely tracks changes

[5] See Pinto et al. (2015a) or Pinto et al. (2015b) for a thorough explanation of diffusion indices.

[6] The weights currently used in the FRBR composite index RIC_t were obtained from Harris et al. (2004). Note, however, that the weights in that paper were chosen with the explicit goal of comparing the ISM and FRBR diffusion index series. In other words, the FRBR composite index, with those specific weights, was intended to track changes at the national level.

[7] We apply our own seasonal adjustment process to the ISM, MEG, and FRBR diffusion index series to preserve uniformity.

Table 1 Summary Statistics

	Variable	Mean	Std. Dev.	N
ISM_t	Composite ISM	52.765	4.810	182
MEG_t	Manufacturing employment growth	-0.172	0.386	181
RIC_t	FRBR manu. composite diffusion index	50.366	5.790	182
RIC_t^S	Shipments	50.820	6.462	182
RIC_t^O	Orders	50.363	6.904	182
RIC_t^E	Employment	49.817	4.980	182
RIC_t^B	Backlog	45.410	5.667	182
RIC_t^C	Capacity	49.304	5.847	182
RIC_t^V	Vendors	52.032	3.354	182
RIC_t^H	Hours	49.531	4.978	182
RIC_t^W	Wages	54.826	2.828	182
RIC_t^{IF}	Inventory finished goods	59.627	4.036	182
RIC_t^{IR}	Inventory raw materials	58.006	3.422	182

in the manufacturing sector, but also because the data are available at monthly intervals.[8]

Descriptive analysis of the series

We start our examination with a simple descriptive analysis of the series. Consider first the RIC_t and ISM_t series. The summary statistics reported in Table 1 indicate that, for the period under consideration, the average value of ISM_t is higher than the average of RIC_t, but the volatility of RIC_t is more pronounced.[9] Moreover, from Figures 1a and 1b, it appears that the series follow each other very closely.

Figure 2a describes the behavior of MEG_t (measured on the left axis) and RIC_t (measured on the right axis), and Figure 2b shows a scatter plot of the two series.[10] It appears from the two figures that the FRBR composite diffusion index tracks fairly well the Fifth District MEG. In Pinto et al. (2015b), we find that the FRBR employment diffusion index also follows very closely the behavior of regional employment growth. In Section 3, we will evaluate in more detail the differential contribution of each one of the FRBR diffusion index series in explaining regional economic changes.

[8] When comparing the regional MEG and the FRBR diffusion indices, we use the currently reported version of the FRBR diffusion index with $w^u = 1, w^s = 0$, and $w^d = -1$.

[9] The higher volatility of RIC_t may be partly attributed to a smaller sample size.

[10] In Figures 2a and 2b, we use the FRBR diffusion index series centered at zero.

Figure 1 ISM_t **and** RIC_t

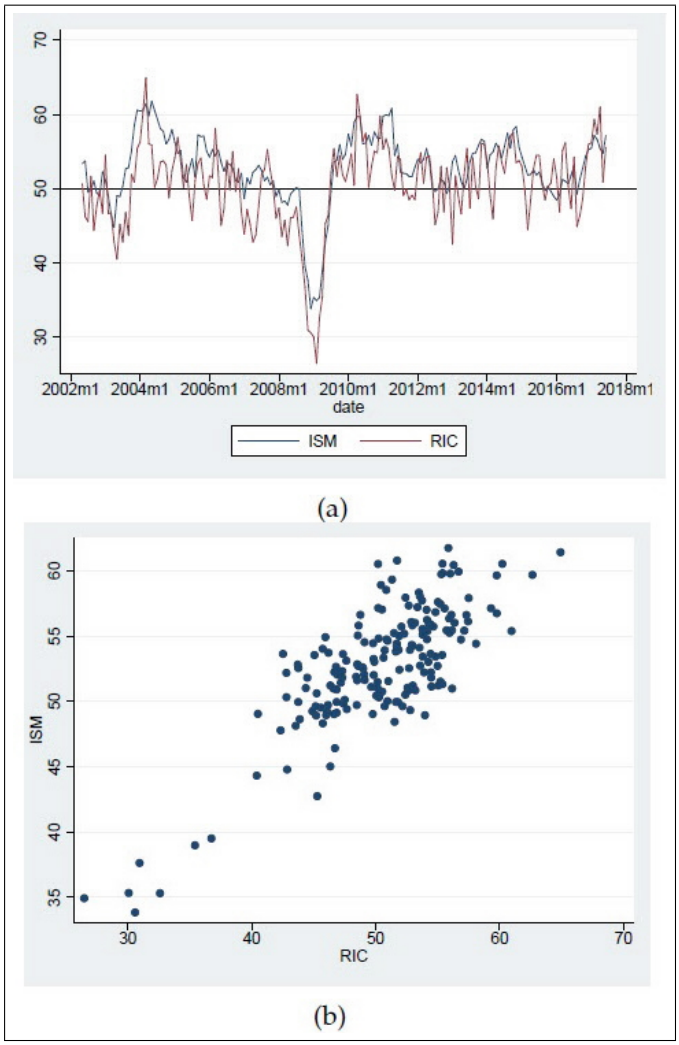
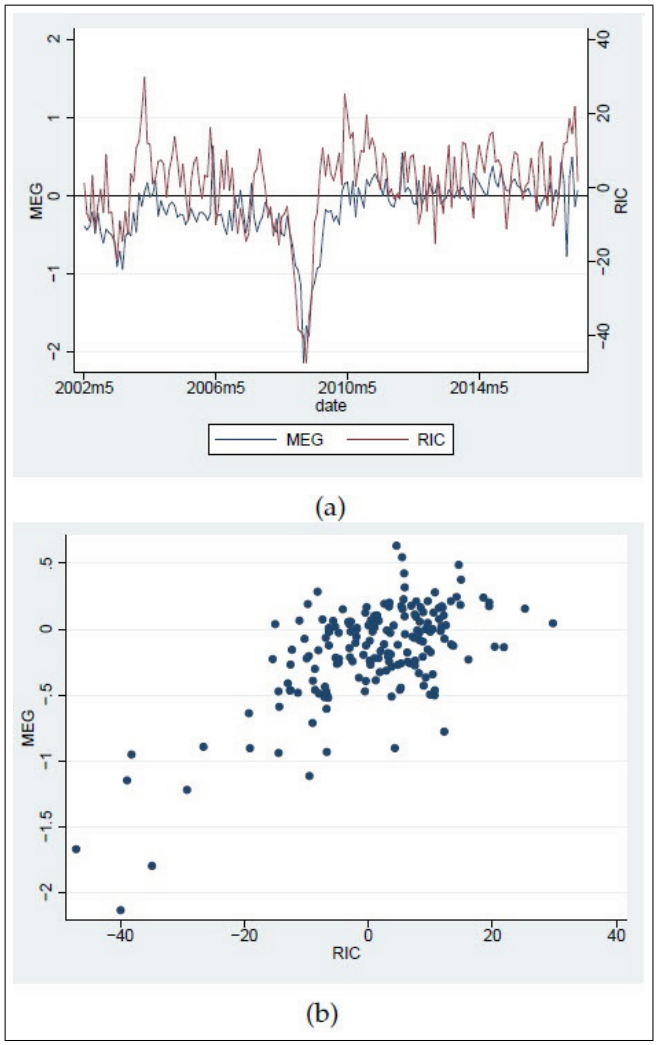


Figure 3a shows the evolution of the individual components of the composite index RIC_t: RIC_t^E, RIC_t^S, and RIC_t^O. While the RIC_t^O series shows the largest volatility, the RIC_t^E series shows the least. Additional diffusion indices obtained from other questions in the FRBR survey are shown in Figure 3b. Among all the series, RIC_t^W has the lowest standard deviation.

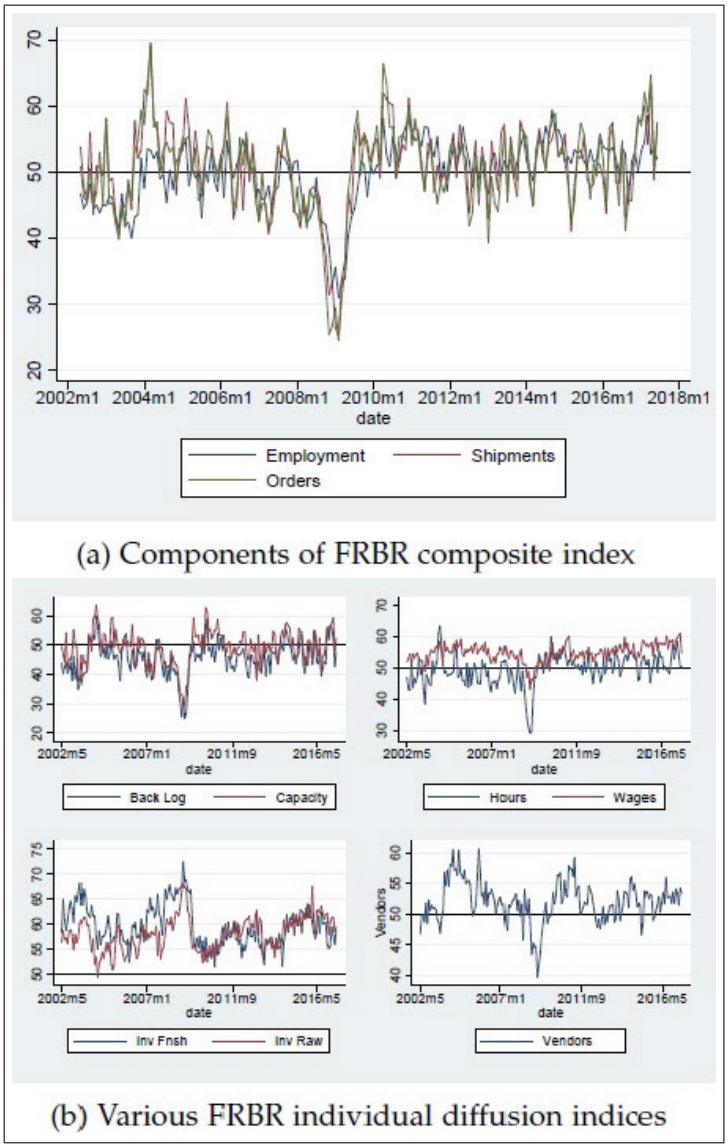
Figure 2 MEG_t **and** RIC_t



A key feature of the series considered in the analysis is that they exhibit high levels of persistence.[11] This behavior is evident from the autocorrelation (ACF) and partial autocorrelation functions (PACF) of the series. Figure 4 shows the ACF and PACF of the ISM_t, RIC_t,

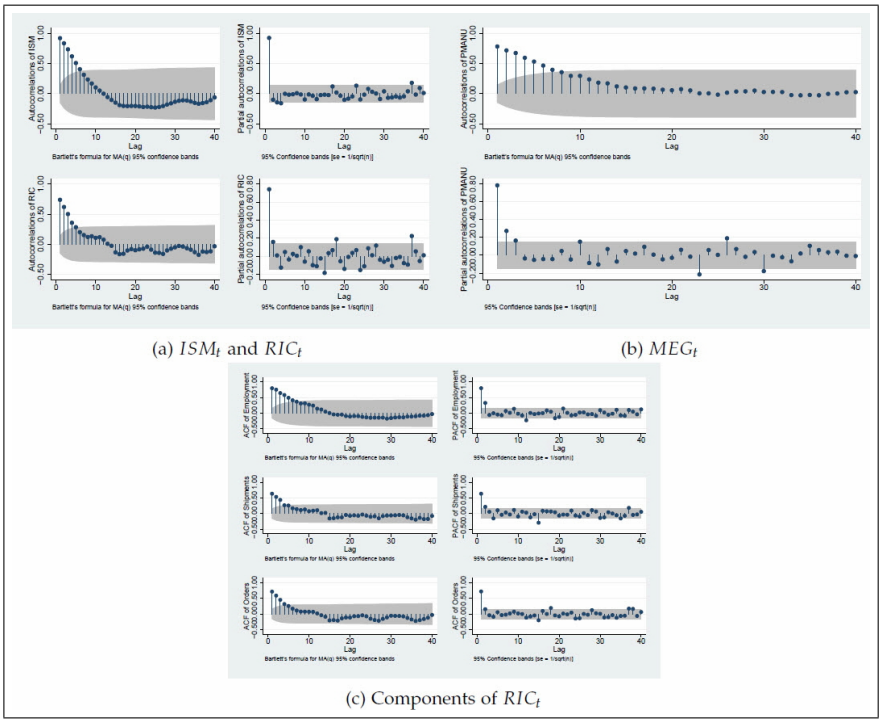
[11] We calculate and report in Table 11 in the Appendix the results of several unit root tests for all variables. In all cases, the tests reject the presence of a unit root.

Figure 3 FRBR Diffusion Indices



MEG_t, and the individual diffusion indices used to calculate RIC_t. In every case, the ACF and PACF indicate a strong autocorrelation at the first three of four lags. The latter is relevant because it suggests that when considering models that explain and predict the evolution of ISM_t and MEG_t, it would become critical to incorporate their dy-

Figure 4 ACF and PACF



(a) ISM_t and RIC_t (b) MEG_t

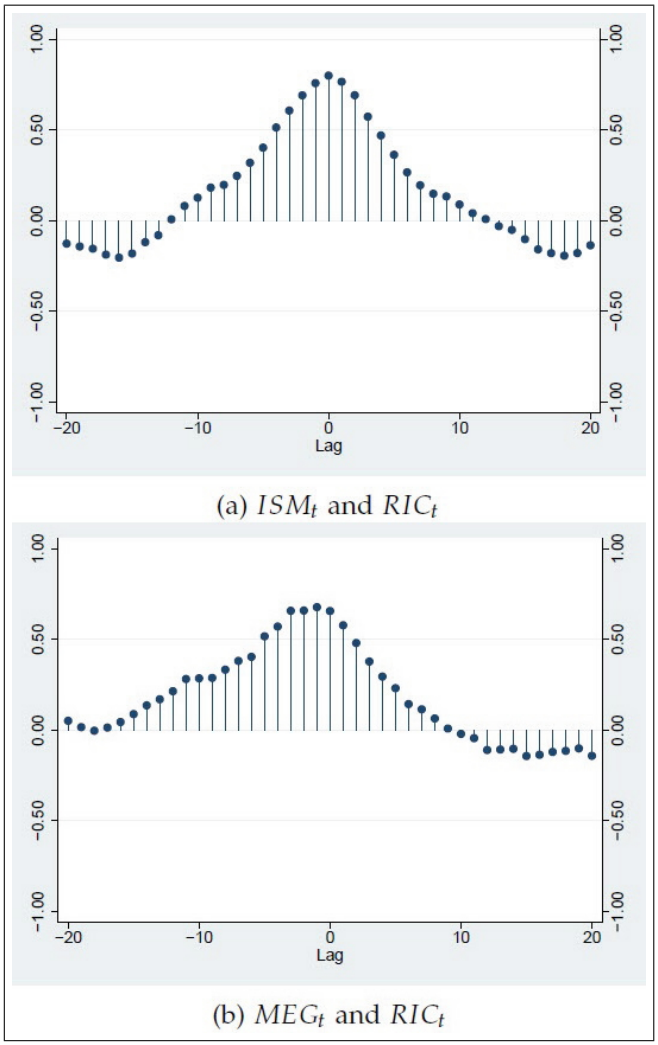
(c) Components of RIC_t

namic behavior in addition to the dynamic behavior of RIC_t and RIC_t^i. We evaluate several univariate dynamic models that explain the behavior of the series in Sections 2.1 and 3.1.

Cross-correlations

As a first approximation to the analysis of the dynamic relationship between the FRBR diffusion indices and national and regional economic indicators, we examine the cross-correlations between the variables $X_t = ISM_t, MEG_t$, defined as $Corr(X_t, RIC_{t+h})$, for different values of $h = -20, ..., -1, 0, 1, ...20$. The cross-correlograms between ISM_t and RIC_t and MEG_t and RIC_t are shown in Figure 5. Figure 5a indicates a very strong contemporaneous correlation between ISM_t and RIC_t, with a correlation equal to 0.80 (at $h = 0$). The highest correlation between the MEG_t and RIC_t series, shown in Figure 5b, occurs at $h = -1$ and is equal to 0.68. In other words, this last cross-

Figure 5 Cross-correlograms



correlogram seems to indicate that RIC_t tends to lead MEG_t, so RIC_t may contain information about the future behavior of the MEG_t series.

We also examine the cross-correlations between ISM_t and MEG_t and the individual FRBR diffusion indices, and among the individual FRBR diffusion indices themselves. The results, which are reported in Table 12 in the Appendix, can be summarized as follows. First, the cross-correlograms between ISM_t and the FRBR individual diffusion indices RIC_t^i generally reflect a strong contemporaneous relationship

(at $h = 0$), with the exception of RIC_t^E and RIC_t^W (in both cases, the highest correlation occurs at $h = 1$; it is 0.71 for RIC_t^E and 0.58 for RIC_t^W). Second, the highest correlation is between ISM_t and RIC_t^V, with a value of 0.78 at $h = 0$, followed by RIC_t^O, with a value of 0.77 also at $h = 0$. Third, for some diffusion indices, specifically RIC_t^{IF} and RIC_t^{IR}, the correlation is negative. Fourth, the results are somewhat different when comparing the cross-correlations between MEG_t and RIC_t^i in Table 13. For instance, the highest $Corr(MEG_t, Y_{t+h})$ is observed when $Y_{t+h} = RIC_{t+h}^E$ and $h = 0$, with a value of 0.76. Other diffusion indices, however, tend to lead MEG_t series (when $Y_{t+h} = RIC_{t+h}^O$, the highest value is observed at $h = -3$, and when $Y_{t+h} = RIC_{t+h}^S$ the highest value is $h = -1$). Finally, the RIC_t^i series also tend to move together. The cross-correlations between RIC_t^E, RIC_t^O, and RIC_t^S, and the other individual diffusion indices are reported in the Appendix in Tables 14, 15, and 16. Note that the correlations between RIC_t^E and the other diffusion indices are relatively low. The correlations are higher for the series RIC_t^O (for instance, the contemporaneous correlations between RIC_t^O and RIC_t^S and between RIC_t^O and RIC_t^C are, respectively, 0.92 and 0.90). The main takeaway from this preliminary analysis is that the information contained in the FRBR survey seems to be highly correlated with changes in both national and regional economic conditions. However, based on the correlations observed between the series, various composite indices based on different series and weights may be constructed to more accurately explain national and regional economic changes.

2. PREDICTING THE NATIONAL ECONOMY

We now proceed to a formal analysis of how well the FRBR composite index RIC_t tracks the national economy. As mentioned earlier, we use the ISM_t series as a gauge of national economic activity. We begin by first evaluating the predictive ability of a number of univariate dynamic models of ISM_t for benchmarking purposes. We next estimate several linear and vector autoregressive models (VARs) that incorporate the diffusion indices obtained from the FRBR manufacturing survey, and we examine how this additional information improves the models' predictive ability. We specifically compare the predictive power of models that use the composite diffusion index RIC_t to other less-constrained models in which the components of RIC_t are considered individually, as well as models that incorporate other diffusion indices, not part of RIC_t, calculated from currently available FRBR survey data.

Univariate models of ISM

The descriptive analysis of Section 1.1 suggests that the ISM_t series is highly persistent. In order to formally examine its behavior, we first estimate several univariate dynamic models of ISM_t and determine how well these models predict ISM_t.[12] These univariate dynamic models are used as the benchmark against which we evaluate the performance of models that incorporate the information from the FRBR survey. We consider models that assume a general ARMA(p,q) representation of the form

$$ISM_t = a + \sum_{j=1}^{p} \phi_j ISM_{t-j} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}, \tag{4}$$

where ε_t is assumed to be an i.i.d. white noise process, and $\phi = [\phi_1, \ldots, \phi_p]$ and $\theta = [\theta_1, \ldots, \theta_q]$ are the autoregressive and moving average coefficients, respectively. Table 17 in the Appendix presents estimates of several univariate models fitted to the ISM_t series, together with the goodness-of-fit statistics AIC and BIC. Based on the estimation results, AR(4) has the best AIC statistics, whereas AR(1) produces the best BIC statistics.[13] These models produce residuals with RMSE of 1.841 for AR(1) and 1.789 for AR(4). Figure 6 shows one-step-ahead predictions of ISM_t for the AR(1) and AR(4) models. Overall, the results suggest that past values of ISM_t explain fairly well the behavior of the series. The contribution of the information contained in the FRBR survey should, therefore, be assessed by comparing different models that include the FRBR diffusion indices to the performance of these very simple univariate dynamic models.

Linear models

We now present estimates of several linear models that explain the behavior of ISM_t using the information collected from the FRBR survey series. First, we consider very simple models that assume a contemporaneous relationship between the variables of the form

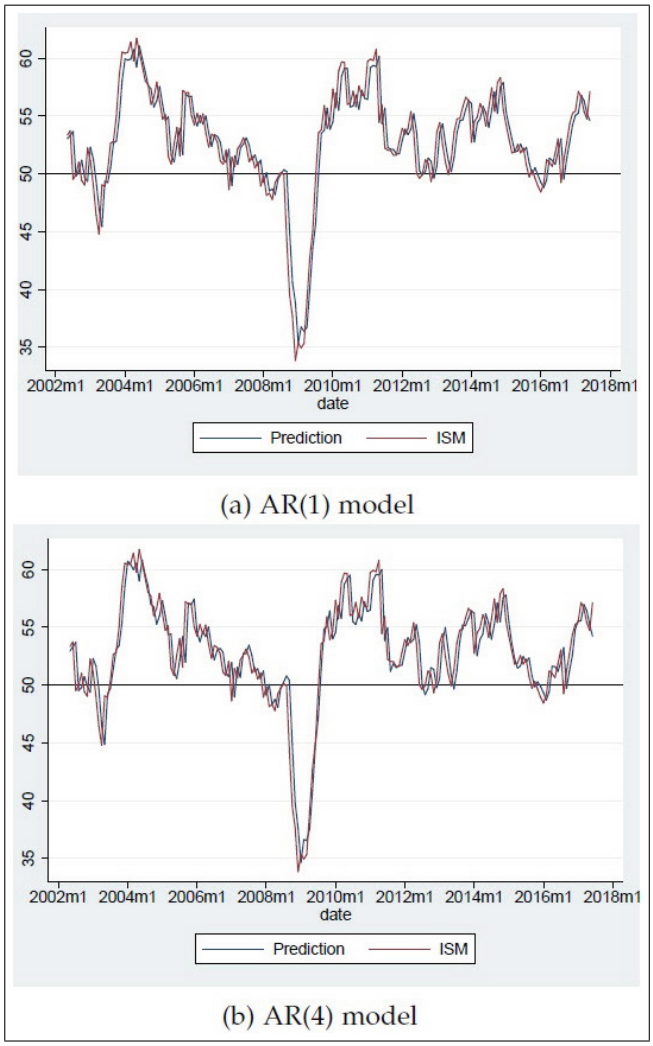
$$ISM_t = \alpha + X_t\beta + \varepsilon_t, \tag{5}$$

where X_t is a vector of regional predictors and ε_t is an error term assumed to be an i.i.d. white noise process. Table 2 presents the estimates of four alternative model specifications depending on the variables

[12] Throughout the paper, we compare models based on their predictive accuracy measured by the root mean squared error (RMSE).

[13] Several alternative ARMA(p,q) were estimated; those reported in Table 17 have the smallest AIC and BIC statistics.

Figure 6 One-Step-Ahead Predictions of ISM_t



included in X_t. Model (1) includes only the composite index of the FRBR series constructed as a weighted average of RIC_t^E, RIC_t^S, and RIC_t^O, given by (2); in other words, $X_t = RIC_t$ in this case. In model (2), each one of the components of the composite index are included in an unconstrained form, meaning that $X_t = [RIC_t^E, RIC_t^S, RIC_t^O]$. Model (3) adds additional components not used in the composite index but available in the FRBR survey, and model (4) is basically a

Table 2 Linear Models of ISM: Contemporaneous Regressors

	(1)	(2)	(3)	(4)
RIC_t	0.665***			
	(0.054)			
RIC_t^E		0.269**	0.143*	0.149*
		(0.089)	(0.068)	(0.064)
RIC_t^S		0.184*	0.106	
		(0.089)	(0.078)	
RIC_t^O		0.246*	0.096	0.168***
		(0.101)	(0.096)	(0.048)
RIC_t^B			0.037	
			(0.070)	
RIC_t^C			-0.061	
			(0.089)	
RIC_t^V			0.479***	0.521***
			(0.087)	(0.082)
RIC_t^H			-0.007	
			(0.071)	
RIC_t^W			0.083	
			(0.095)	
RIC_t^{IV}			-0.184*	-0.193**
			(0.072)	(0.067)
RIC_t^{IR}			-0.239**	-0.220**
			(0.078)	(0.074)
Constant	19.282***	17.618***	32.422***	34.024***
	(2.762)	(3.898)	(8.178)	(7.659)
N	182	182	182	182
Adj-R^2	0.638	0.639	0.768	0.770
RMSE	2.893	2.889	2.318	2.307

Note: Newey-West standard errors in parentheses; *$p < 0.05$, ** $p < 0.01$, ***$p < 0.001$.

refinement of model (3) obtained through a stepwise procedure of regressor selection.[14]

[14] Throughout the paper, we follow a standard stepwise procedure to select the variables of the model. We typically proceed from general to particular: we start with a general model that includes the largest possible set of predictors (in the dynamic versions of the models, we also include lagged values of the variables),then we remove predictors with the highest p-values and refit the model. The procedure also takes into account, when comparing models, their respective AIC and BIC values. Standard errors are produced by a Newey-West regression procedure that corrects potential serial correlation in the error terms. While under serial correlation, OLS still produces unbiased parameter estimates, the standard errors in this case are not efficient. We reestimate the model in (5) using the Newey-West regression procedure that produces serial correlation robust standard errors. The adjusted-R^2 measure is from the OLS regression.

The results show that the behavior of the FRBR series explains considerable variation of ISM_t. Specifically, model (1) shows that the FRBR composite index RIC_t explains, by itself, about 64 percent of variation of ISM_t. Model (1) also tells us that when $RIC_t = 50$, which represents the point at which the percentage of respondents reporting an increase is the same as the percentage reporting a decrease in the FRBR survey, the ISM composite index is 52.53. According to the linear estimates, RIC_t and ISM_t are equal when RIC_t is approximately 57. Moreover, when RIC_t is higher (lower) than 57, then $RIC_t > (<)ISM_t$.

We additionally perform the following exercise. Suppose the goal is to construct a composite index $\overline{RIC}_t$ that includes $\{RIC_t^E, RIC_t^S, RIC_t^O\}$ and tracks as closely as possible the ISM_t series. Specifically, suppose that $\overline{RIC}_t$ takes the functional form $\overline{RIC}_t = \alpha + \beta^E RIC_t^E + \beta^E RIC_t^S + \beta^E RIC_t^O$, and $\{\alpha, \beta^E, \beta^S, \beta^O\}$ are chosen so as to minimize $\sum_{t=1}^{T} \left(ISM_t - \overline{RIC}_t\right)^2$, subject to the constraints $\beta^E + \beta^S + \beta^O = 1$, $\beta_i \geq 0$. The values obtained in this case are: $\alpha = 2.545$, $\beta^O = 0.53$, $\beta^S = 0.33$, $\beta^E = 0.14$. Two remarks are worth making. First, since the ISM_t series seems to be displaced upward, as explained before, $\overline{RIC}_t$ includes a positive constant term (note that the current composite FRBR diffusion index RIC_t does not have a constant term). Second, RIC_t^O should receive the highest weight and RIC_t^E the lowest weight in the composite index, if the objective is to construct a composite index that tracks as closely as possible the ISM_t series.

When each individual component is included as separate regressors in an unconstrained way (model [2]), the fit and predictive power of the model improve, but note that such improvement is relatively small. Also, by comparing the estimates of model (2), RIC_t^E seems to be the most important variable at explaining the behavior of ISM_t, but in the construction of the composite index RIC_t, this variable receives the lowest weight of the three individual diffusion indices. Sizable improvements in fit and predictive ability are observed, however, when we incorporate additional survey information, as evidenced by models (3) and (4). In general, the results of Table 2 confirm that when the objective is to describe or predict the evolution of the national economy, including other information readily available through the FRBR survey would tend to improve the outcome.

Including only contemporaneous values of the FRBR diffusion indices is somewhat restrictive. It is clear from Section 1.1 that the series show high levels of persistence, which suggests that further improvements could be obtained by using models that include a dynamic structure. Thus, we estimate next a set of models that account for this

more general dynamic behavior of the form

$$ISM_t = \alpha + \sum_{j=0}^{3} \beta_j X_{t-j} + \sum_{j=1}^{3} \gamma_j ISM_{t-j} + \varepsilon_t, \tag{6}$$

where ε_t is again assumed to be an i.i.d. distributed white noise process. Table 3 presents the estimates of four models of the type represented by (6): model (1) includes contemporaneous and lagged value of the FRBR composite index RIC_t; model (2) adds lagged values of ISM_t; model (3) includes contemporaneous and lagged values of the components of the RIC_t; and model (4) includes lagged values of ISM_t.[15]

By simply considering lagged values of RIC_t, such as in model (1), the RMSE decreases substantially (the static model [1] of Table 2 has a RMSE equal to 2.893, and this one has a RMSE of 2.539). However, once the model incorporates lagged values of ISM_t, such as in model (2), the explanatory power of RIC_t declines. Also, model (2) has a much better predictive accuracy. Model (3) is an improvement relative to model (1) but not relative to model (2). Considering both a less constrained and richer dynamic behavior undoubtedly increases the fit of the model and improves its predictive power. Model (4), which includes lagged values of the individual diffusion indices RIC_t^i, $i = E, O, S$, and lagged values of ISM_t, has the lowest RMSE among all models, with a value of 1.676. Notice that the FRBR diffusion index that captures changes in orders, RIC_t^O, is always relevant at explaining the behavior of ISM_t, even after accounting for past values of ISM_t.

Finally, Table 4 shows the estimates of a model obtained by a stepwise procedure of regressor selection among the FRBR survey series RIC_t^i, and their respective lags, and lagged values of ISM_t. In this case, the model explains almost 90 percent of the variation in ISM_t. The variables that seem to be most relevant at explaining changes in ISM_t include, in addition to ISM_{t-1}, the regional indicators RIC_{t-2}^E, RIC_{t-2}^O, and RIC_t^V. The linear predictions of this model are shown in Figure 7. The RMSE is 1.554, and this value is the lowest among all the models considered up to this point.

Vector autoregressive models

In this section, we use VAR models to further explore and understand the relationship between the ISM_t series and the indices elaborated by the FRBR. Let $Z_t = [RIC_t^1, RIC_t^2, \cdots, RIC_t^m, ISM_t]$ be a multivariate

[15] Linear predictions of models (1) through (4) are shown in Figure 16 in Appendix A.3.2.

Table 3 Linear Models of ISM: Contemporaneous Lagged Regressors

ISM_t	(1)	(2)	(3)	(4)
ISM_{t-1}		2.313***		2.301***
		(0.196)		(0.202)
ISM_{t-2}		1.002		0.941
		(0.114)		(0.105)
ISM_{t-3}		0.939		1.024
		(0.076)		(0.076)
RIC_t	1.509***	1.190***		
	(0.075)	(0.049)		
RIC_{t-1}	1.200***	0.993		
	(0.050)	(0.038)		
RIC_{t-2}	1.106*	0.984		
	(0.052)	(0.045)		
RIC_{t-3}	1.123*	0.998		
	(0.054)	(0.037)		
RIC_t^E			1.096	0.990
			(0.082)	(0.050)
RIC_{t-1}^E			0.984	0.990
			(0.073)	(0.050)
RIC_{t-2}^E			0.896	0.857**
			(0.063)	(0.050)
RIC_{t-3}^E			1.025	1.085
			(0.075)	(0.046)
RIC_t^O			1.279**	1.188**
			(0.110)	(0.063)
RIC_{t-1}^O			1.116	1.037
			(0.084)	(0.056)
RIC_{t-2}^O			1.048	1.078
			(0.074)	(0.055)
RIC_{t-3}^O			1.041	0.952
			(0.082)	(0.045)
RIC_t^S			1.055	0.963
			(0.079)	(0.043)
RIC_{t-1}^S			1.055	0.977
			(0.072)	(0.044)
RIC_{t-2}^S			1.099	0.969
			(0.074)	(0.046)
RIC_{t-3}^S			1.119	1.051
			(0.086)	(0.050)
Constant	11.965***	4.241**	14.972***	5.844**
	(2.331)	(1.495)	(2.992)	(1.749)
N	179	179	179	179
Adjusted R^2	0.725	0.872	0.736	0.880
RMSE	2.539	1.734	2.489	1.676

Note: Newey-West standard errors in parentheses; $^*p < 0.05$, ** $p < 0.01$, $^{***}p < 0.001$.

time series, where RIC_t^i represents each one of the diffusion indices at time t. The long-run structural relationship between the FRBR series

Table 4 Linear Model of ISM: Stepwise Selection

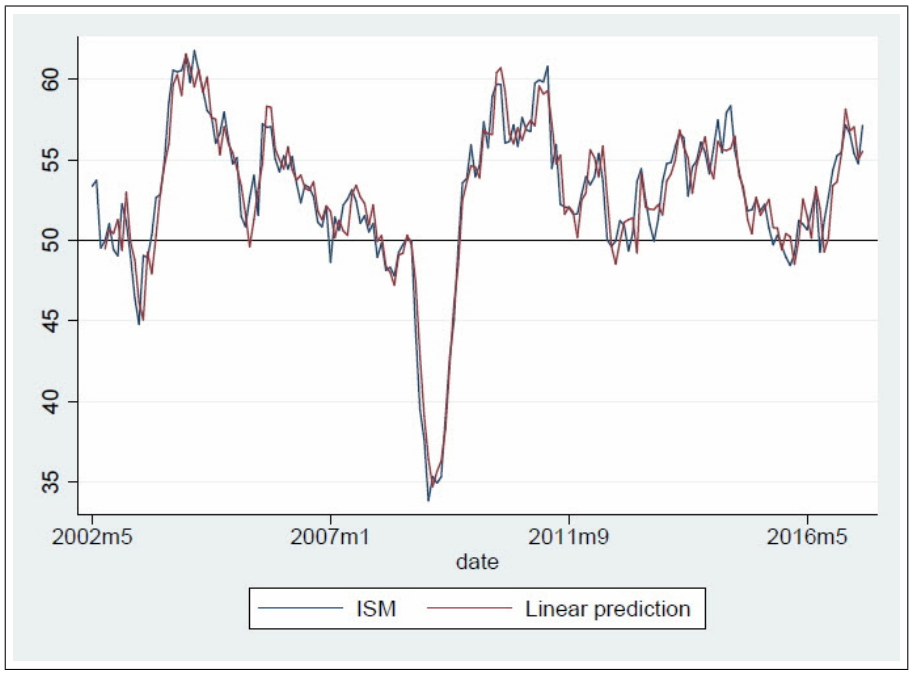
	ISM_t
RIC^E_{t-2}	-0.161**
	(0.049)
RIC^E_{t-3}	0.063
	(0.043)
RIC^O_t	0.145**
	(0.044)
RIC^O_{t-1}	0.069
	(0.052)
RIC^O_{t-2}	0.120*
	(0.052)
RIC^S_{t-2}	-0.069
	(0.050)
RIC^B_{t-1}	0.039
	(0.051)
RIC^C_t	-0.095
	(0.049)
RIC^C_{t-1}	-0.110*
	(0.050)
RIC^W_t	0.057
	(0.067)
RIC^W_{t-1}	-0.101
	(0.067)
RIC^W_{t-2}	-0.081
	(0.068)
RIC^V_t	0.206**
	(0.062)
RIC^{IR}_{t-2}	0.135*
	(0.066)
RIC^{IR}_{t-3}	-0.067
	(0.062)
RIC^{IF}_t	-0.078
	(0.055)
RIC^{IF}_{t-1}	-0.056
	(0.061)
RIC^{IF}_{t-2}	-0.099
	(0.061)
RIC^{IF}_{t-3}	0.062
	(0.057)
ISM_{t-1}	0.720***
	(0.059)
Constant	17.001**
	(6.217)
N	179
Adj-R^2	0.897
RMSE	1.554

Note: Newey-West standard errors in parentheses; $*p < 0.05$, $** \ p < 0.01$, $***p < 0.001$.

and ISM_t is modeled by the pth-order VAR process

$$B\,Z_t = a + \sum_{j=1}^{p} A_j Z_{t-j} + \varepsilon_t, \tag{7}$$

Figure 7 Stepwise Selection Model of ISM: Observed Values and Predictions



where B and A_j are $(m+1)\times(m+1)$ matrices, and $\varepsilon_t = [\varepsilon_t^1, \varepsilon_t^2, \ldots, \varepsilon_t^{m+1}]'$ is a multivariate white noise process with mean zero and variance $I_{(m+1)\times(m+1)}$. Multiplying both sides of (7) by the inverse of B, we obtain

$$Z_t = \alpha + \sum_{j=1}^{p} \Phi_j Z_{t-j} + e_t, \tag{8}$$

where $\alpha = B^{-1}a, \Phi_j = B^{-1}A_j$ and $e_t = [e_t^1, e_t^2, \ldots, e_t^{m+1}]'$ is a multivariate mean zero white noise process with variance-covariance matrix $\Sigma_e = B^{-1}(B^{-1})'$. The equation in (8) represents a VAR model of order p, which can be estimated by maximum likelihood. With the estimates of the Σ_e matrix, we perform a Cholesky decomposition and obtain a lower triangular matrix P such that $\Sigma_e = PP'$. Premultiplying (8) by P^{-1} yields

$$P^{-1}Z_t = P^{-1}\alpha + P^{-1}\sum_{j=1}^{p} \Phi_j Z_{t-j} + u_t, \tag{9}$$

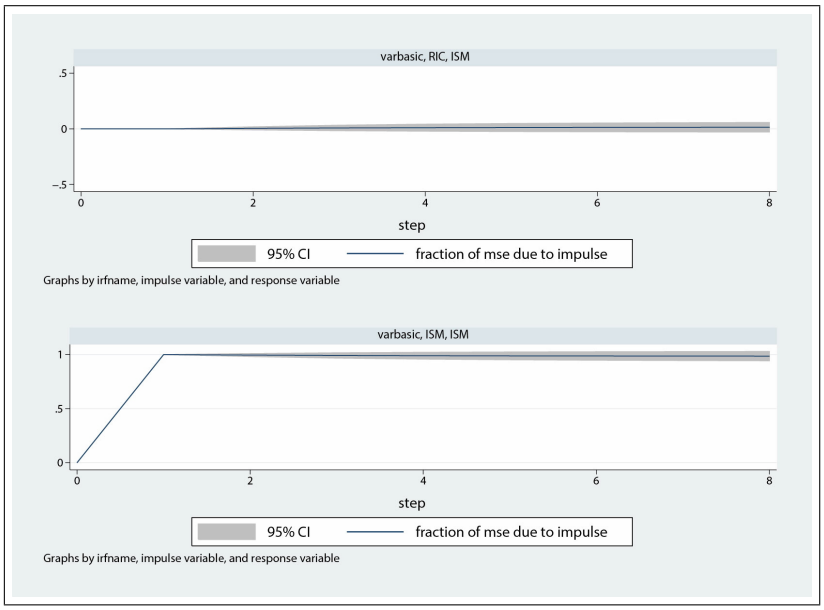
where $u_t = P^{-1}e_t$ is a multivariate white noise process with variance-covariance matrix $I_{(m+1)\times(m+1)}$. The expression in (9) gives us $(m+1)$ equations in the FRBR series and ISM_t, in addition to their past values. Because P is lower triangular, so is P^{-1}, thus the $(m+1)$-st equation of (9) contains all current and past values of the multivariate time series. Also note that the error term u_t^{m+1} is the linear combination of error terms $(e_t^1, e_t^2, \ldots, e_t^{m+1})$ weighted by the coefficients of the matrix P^{-1}. The expression obtained in the $(m+1)$-st equation is what is known as the structural equation of ISM_t. This equation represents ISM_t as a linear function of its past values (up to p-th lag), as well as contemporaneous and lagged values of the FRBR series. Using this equation, we then construct "predictions" of the value of the ISM diffusion index under the following premise: at time t, when regional survey results have become known, we can use this information to obtain a reasonable prediction of the value of the ISM diffusion index for the current time period.

Bivariate VAR model: ISM and FRBR composite diffusion indices

We first estimate a VAR(1) model for bivariate series consisting of the composite diffusion indices ISM_t and RIC_t. The selection of lags is based on AIC and BIC statistics. The parameter estimates of the VAR(1) model are shown in Table 18 in the Appendix, together with the variance-covariance matrix of the error terms and its Cholesky decomposition.[16] Using the inverse of the lower triangular matrix obtained from the Cholesky decomposition, we construct the structural form for ISM_t, which is plotted in Figure 17 in the Appendix. The RMSE of this specification is 1.72. We additionally perform a forecast error variance decomposition (FEVD) to interpret the results of the VAR model. The FEVD quantifies the relative contribution of the variables in the system, in this case ISM and RIC, to the variance of the forecast error of each variable. We focus here on the forecast error variance of ISM_t. The top panel in Figure 8 shows the percentage of the forecast error variance of ISM explained by RIC, and the bottom panel shows the percentage explained by itself. The figure indicates that variations in ISM are mostly explained by shocks to the series itself; the variation explained by RIC is virtually zero.

[16] The table and the figure showing the observed and predicted values are shown in Appendix A.3.3.

Figure 8 VAR Model: ISM_t and RIC_t. Forecast Error Variance Decomposition

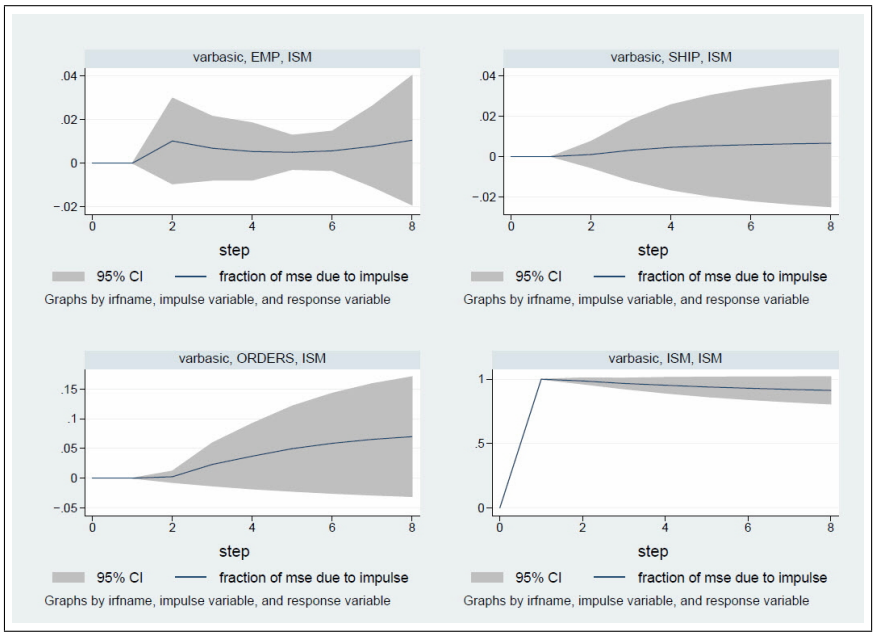


Multivariate VAR: ISM and FRBR individual diffusion indices

We now estimate a VAR model that includes ISM_t and the individual diffusion indices used in the FRBR composite index, RIC_t. The AIC and BIC statistics suggest that a VAR(2) model fits the data best. The parameter estimates of the VAR(2) model are shown in Table 19 in Appendix A.3.4, along with the variance-covariance matrix and its Cholesky decomposition. Using the inverse of the lower triangular matrix obtained from the Cholesky decomposition, we construct the structural form of ISM_t.[17] The prediction errors have a standard deviation of 1.64, which is slightly smaller than the RMSE of the bivariate VAR model considered in the previous section. The FEVD in Figure 9 describes the effect of a shock on the variables RIC^E, RIC^S, RIC^O, and ISM on the forecast error variance of ISM. Once again, the figure indicates that shocks on the variable ISM essentially explain most of

[17] Figure 18 in Appendix A.3.4 compares the predicted values of the model to the observed values.

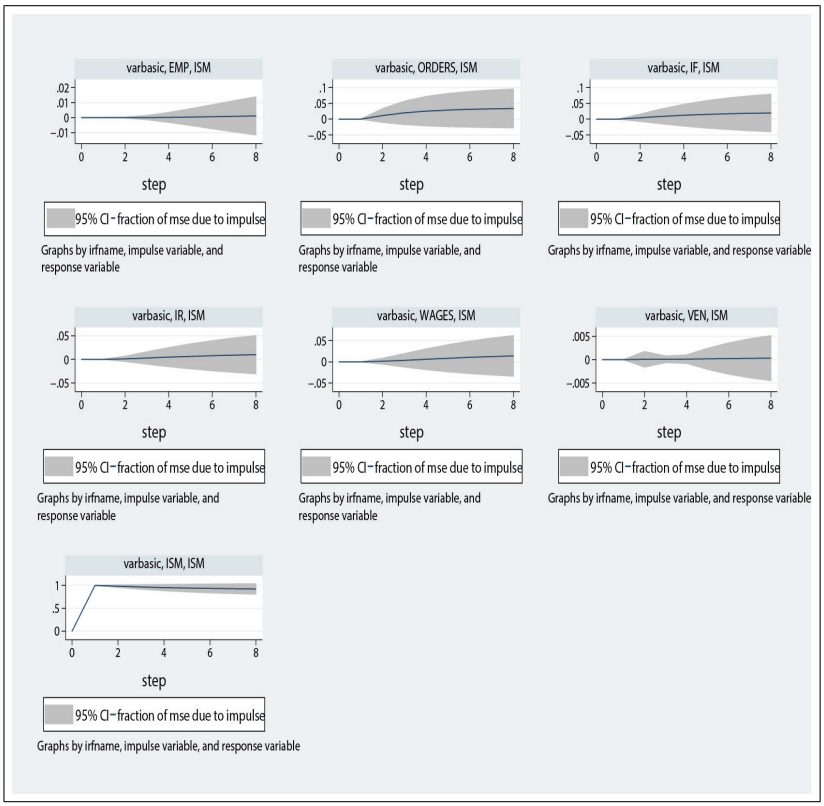
Figure 9 ISM and FRBR Individual Diffusion Indices. Forecast Error Variance Decomposition



the variation of ISM. However, the variable RIC^O now becomes relevant, explaining approximately 8 percent of the variation in ISM after eight periods.

We finally estimate a VAR model that incorporates other diffusion indices available from the FRBR survey. The variables included in the analysis were selected through a stepwise regression procedure similar to the one followed in Section 2.2. Based on AIC and BIC statistics, the VAR(1) model fits the data best. The results are reported in Table 20 in Appendix A.3.4. This specification offers a high level of predictive accuracy with the lowest RMSE, which is equal to 0.85. The FEVD in Figure 10 confirms the importance of the ISM series in explaining its own variation. The FEVD also shows that, among all FRBR diffusion indices, RIC^O is the most important one for explaining variations in ISM.

Figure 10 ISM and FRBR Individual Diffusion Indices. Forecast Error Variance Decomposition



Summary of findings

To summarize our findings, in Table 5 we compare the RMSE of the models discussed thus far. We include, for comparison, the RMSE of the model that includes only the composite index RIC_t currently reported by the FRBR. In light of the predictive accuracy of the models, it is clear that a multivariate VAR model dominates all other alternatives, producing a RMSE equal to 0.85. However, a linear dynamic model that considers information readily available through the FRBR survey but not currently included in the calculation of the composite index RIC_t, offers more accurate predictions, with a RMSE of 1.55. So even the consideration of this last relatively simpler model would entail an important increase in predictive accuracy compared with a

Table 5 RMSE for Selected Models of ISM

Model		RMSE
Composite index	RIC_t	2.89
Univariate	$AR(1)$	1.84
	$AR(4)$	1.79
Linear	Contemporaneous	2.31
	Dynamic	1.55
VAR	Bivariate	1.72
	Multivariate	0.85

model that relies exclusively on the composite index RIC_t, which has the highest RMSE, equal to 2.89.

3. PREDICTING THE REGIONAL ECONOMY

While in Section 2 we examined the extent to which the information collected by the FRBR manufacturing survey helps explain changes in the national economy, we now focus on how well the survey tracks the regional economy. As explained earlier, we use payroll MEG as a measure of regional manufacturing activity for two reasons. First, MEG data are available monthly and for all states. Second, MEG is a good indicator of the economic performance of the manufacturing sector, which is the focus of the present analysis, so MEG serves as a reasonable benchmark against which to assess the predictive ability of the information contained in the FRBR manufacturing survey. Our approach is similar to that in the previous section: we begin by estimating several univariate dynamic models of MEG; next, we compare the performance of these models to the performance of linear and VAR models of MEG that incorporate the diffusion index series from the FRBR manufacturing survey.

Univariate models of MEG

We first examine the predictive power of simple univariate dynamic models that only include the MEG series. The inspection of the MEG autocorrelation and partial autocorrelation functions in Figure 4 reveal that the series show high levels of persistence. To capture such dynamic behavior more formally, we estimate, as we did earlier with the ISM

Table 6 Univariate Models of MEG

MEG_t	(1) AR(2)	(2) AR(3)	(3) ARMA (1,1)	(4) ARMA (1,2)	(5) ARMA (2,1)	(6) ARMA (2,2)
ϕ_1	0.569***	0.526***	0.915***	0.910***	0.862***	0.300
	(0.049)	(0.054)	(0.033)	(0.042)	(0.200)	(1.522)
ϕ_2	0.271***	0.177**			0.0459	0.565
	(0.051)	(0.065)			(0.157)	(1.381)
ϕ_3		0.162*				
		(0.066)				
θ_1			-0.374***	-0.386***	-0.331	0.251
			(0.069)	(0.071)	(0.198)	(1.525)
θ_2				0.0328		-0.253
				(0.065)		(0.534)
Constant	-0.173	-0.172	-0.172	-0.172	-0.172	-0.172
	(0.110)	(0.132)	(0.129)	(0.127)	(0.128)	(0.134)
N	181	181	181	181	181	181
AIC	-9.59	-12.27	-12.51	-10.65	-10.59	-8.67
BIC	3.21	3.72	0.29	5.34	5.40	10.52
RMSE	0.230	0.227	0.228	0.228	0.228	0.228

Note: Standard errors in parentheses; $^*p < 0.05$, ** $p < 0.01$, $^{***}p < 0.001$.

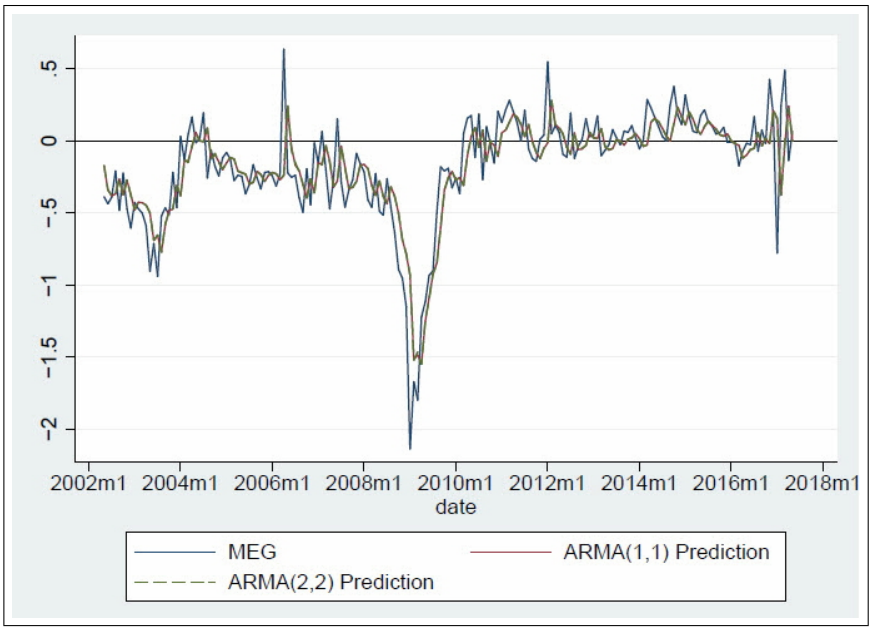
series, several ARMA(p,q) models of the form

$$MEG_t = a + \sum_{j=1}^{p} \phi_j MEG_{t-j} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}, \tag{10}$$

where ε_t is assumed to be an i.i.d. white noise process, and ϕ and θ are the vectors of autoregressive and moving average coefficients, respectively. Table 6 presents the estimates along with goodness-of-fit statistics AIC and BIC.[18] Based on the AIC criterion, the best model specification is an ARMA(1,1), but the BIC criterion chooses the ARMA(2,2). The predictions obtained from these models, shown in Figure 11, are very close to each other. In terms of their predictive accuracy, all models are practically identical, with a RMSE approximately equal to 0.23.

[18] Several alternative ARMA(p,q) models were estimated; Table 6 reports those with the smallest AIC and BIC statistics

Figure 11 ARMA(1,1) and ARMA(2,2) Models. Observed and Predicted Values



Linear models of MEG

We now incorporate the diffusion indices calculated from the FRBR surveys to assess how well they explain economic changes in the Fifth District. We proceed by estimating several linear models of MEG using contemporaneous and lagged values of the FRBR survey series. These models are generally described by the expression

$$MEG_t = \alpha + \sum_{j=0}^{3} \beta_j X_{t-j} + \sum_{j=1}^{3} \gamma_j MEG_{t-j} + \varepsilon_t, \tag{11}$$

where X_t is a vector of diffusion indices produced by the FRBR, and ε_t is an error term assumed to be an i.i.d. white noise process.

Table 7 presents the parameter estimates of five alternative model specifications that only include contemporaneous values of the FRBR series as explanatory variables.[19] In other words, those models assume

[19] In this section, we use the normalization $w^u = w^d = 1$ and $w^s = 0$. This means that when the percentage of participants reporting an increase is equal to the percentage

Table 7 Linear Models of MEG: Contemporaneous Regressors

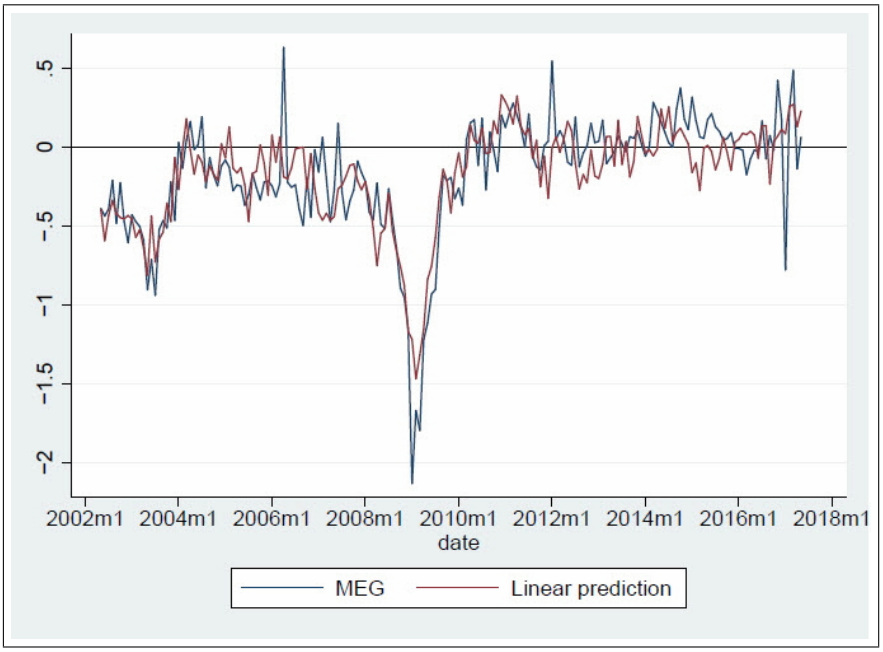
MEG_t	(1)	(2)	(3)	(4)	(5)
RIC_t	0.022***				
	(0.002)				
RIC_t^E		0.030***	0.027***	0.018***	0.018***
		(0.002)	(0.003)	(0.003)	(0.003)
RIC_t^O			-0.003	-0.002	
			(0.004)	(0.005)	
RIC_t^S			0.007	0.007	0.006
			(0.004)	(0.004)	(0.003)
RIC_t^B				0.001	
				(0.004)	
RIC_t^C				-0.010*	-0.010**
				(0.004)	(0.004)
RIC_t^V				0.006	0.006
				(0.004)	(0.004)
RIC_t^H				0.006	0.005
				(0.004)	(0.004)
RIC_t^W				0.011*	0.011*
				(0.005)	(0.004)
RIC_t^{IF}				-0.012***	-0.012***
				(0.003)	(0.003)
RIC_t^{IR}				0.005	0.005
				(0.004)	(0.004)
Constant	-0.186***	-0.160***	-0.170***	-0.167	-0.179*
	(0.022)	(0.019)	(0.019)	(0.090)	(0.075)
N	181	181	181	181	181
Adj-R^2	0.429	0.578	0.587	0.640	0.644
RMSE	0.292	0.251	0.248	0.232	0.231

Note: Newey-West standard errors in parentheses; *$p < 0.05$, ** $p < 0.01$, ***$p < 0.001$.

$\beta_j = \gamma_j = 0$, for $j = 1, 2, 3$. Model (1) only includes the FRBR composite diffusion index, i.e., $X_t = RIC_t$; in model (2), $X_t = RIC_t^E$, which is the FRBR diffusion index that tracks changes in employment; model (3) includes the components of the FRBR composite index, i.e., $X_t = [RIC_t^E, RIC_t^S, RIC_t^O]$; model (4) incorporates additional diffusion

of participants reporting a decrease, the diffusion index is equal to zero. We use this normalization to avoid working with small numbers with many digits. Ideally, we would want to rescale the ISM diffusion index. However, this requires using the ISM raw data, which are not publicly available.

Figure 12 Observed and Predicted Values of MEG: Model (5) (Stepwise Selection)



indices from the FRBR surveys; and model (5) is a refinement of model (4) obtained after a stepwise procedure of regressor selection.[20]

A few remarks are worth making. First, by inspecting model (1), it follows that when the composite diffusion index RIC_t is equal to zero, $MEG_t = -0.186$. In other words, zero employment growth in the district would be consistent with a value of $RIC_t = 8.5$. It is important to note that, in theory, a zero diffusion index does not imply a zero growth rate. A diffusion index captures the breadth of a change measured by the number of respondents experiencing no change, an increase, or a decrease in a specific variable. In other words, a diffusion index tracks changes in the extensive margin. A growth rate, however, in addition

[20] Standard errors are produced by a Newey-West regression procedure that corrects potential serial correlation in the error terms. While under serial correlation, OLS still produces unbiased parameter estimates; the standard errors in this case are not efficient. We reestimate the model in (5) using the Newey-West regression procedure that produces serial correlation robust standard errors. The adjusted-R^2 measure is from the OLS regression.

to changes in the extensive margin, also captures the intensity of the change, or the intensive margin. Pinto et al. (2015a) and Pinto et al. (2015b) show a decomposition of a growth rate into the extensive margin, or a term that includes a diffusion index, and the intensive margin. So the explanatory power of the diffusion indices depends both on the information content of the FRBR surveys, summarized by the diffusion indices, and, more generally, on the extent to which changes in the extensive margin drive changes in the growth rate.

Second, the employment diffusion index RIC_t^E by itself (model [2]) explains about 60 percent of the variation in MEG. In this case, when $RIC_t^E = 5.33$, $MEG_t = 0$. Since information about RIC_t^E is available prior to the release of the MEG_t monthly data (usually, the value of RIC_t^E is known a few weeks earlier), it becomes important to understand such a relationship in order to anticipate the values of MEG_t. Third, adding more information from the FRBR surveys improves the fit of the model, as shown by models (4) and (5). In those cases, the employment diffusion index RIC_t^E is still the most important variable explaining the behavior of MEG_t. Other variables, such as RIC_t^C, RIC_t^W, and RIC_t^{IF}, also contribute to explaining MEG_t. Note that all the models considered in Table 7 have a relatively low adjusted-R^2; model (5) has the highest one, which is equal to 0.644. Figure 12 plots the observed and predicted values of this model. It shows that the two lines only infrequently overlap, confirming the model's low goodness of fit. This suggests, in light of the previous discussion on growth rates and extensive and intensive margins, that in the case of Fifth District employment, diffusion indices (or the extensive margin) only partially explain its growth rate. In other words, a low adjusted-R^2 should not be necessarily used to draw conclusions about the quality of the information content of the FRBR surveys. In fact, Pinto et al. (2015b) show that RIC_t^E tracks fairly well the extensive margin component of the actual employment growth rate in the Fifth District.

Next, we perform a similar exercise as in Section 2.2 but for the MEG series. In this case, we obtain the following results: $\alpha = -0.1$, $\beta^E = 0.83$, $\beta^O = 0.00$, $\beta^S = 0.17$. The main conclusion from this exercise is that a composite index that assigns weights to the individual diffusion indices $\{RIC_t^S, RIC_t^S, RIC_t^O\}$ with the objective of tracking MEG_t as closely as possible should give the highest weight to RIC_t^E, a lower but positive weight to RIC_t^O, and zero weight to RIC_t^O. This composite index is definitely different from the one that is supposed to track the national economy and is also different from the one currently reported by the FRBR.

Up to this point, the models assume a contemporaneous relationship between the variables. The models examined next, summarized in

Table 8 Linear Models of MEG: Contemporaneous and Lagged Regressors

MEG_t	(1)	(2)	(3)	(4)	(5)	(6)
MEG_{t-1}		0.315***		0.320***		0.278***
		(0.075)		(0.078)		(0.079)
MEG_{t-2}		0.097		0.060		0.055
		(0.079)		(0.083)		(0.082)
MEG_{t-3}		0.182*		0.081		0.129
		(0.071)		(0.075)		(0.079)
RIC_t	0.010***	0.008***				
	(0.002)	(0.002)				
RIC_{t-1}	0.006*	0.003				
	(0.003)	(0.002)				
RIC_{t-2}	0.004	-0.000				
	(0.003)	(0.002)				
RIC_{t-3}	0.011***	0.003				
	(0.002)	(0.002)				
RIC_t^E			0.015***	0.010***	0.010**	0.005
			(0.003)	(0.003)	(0.003)	(0.003)
RIC_{t-1}^E			0.010**	0.007*	0.008*	0.005
			(0.003)	(0.003)	(0.004)	(0.003)
RIC_{t-2}^E			0.002	-0.002	0.003	-0.001
			(0.003)	(0.003)	(0.004)	(0.003)
RIC_{t-3}^E			0.008**	0.004	0.008*	0.004
			(0.003)	(0.003)	(0.003)	(0.003)
RIC_t^S					0.003	0.002
					(0.003)	(0.003)
RIC_{t-1}^S					0.004	0.003
					(0.003)	(0.003)
RIC_{t-2}^S					0.004	0.002
					(0.003)	(0.003)
RIC_{t-3}^S					0.002	-0.001
					(0.003)	(0.003)
RIC_t^O					0.002	0.003
					(0.004)	(0.003)
RIC_{t-1}^O					-0.003	-0.001
					(0.004)	(0.003)
RIC_{t-2}^O					-0.005	-0.002
					(0.004)	(0.003)
RIC_{t-3}^O					-0.000	0.002
					(0.004)	(0.003)
Constant	-0.186***	-0.075***	-0.155***	-0.082***	-0.174***	-0.093***
	(0.019)	(0.020)	(0.017)	(0.022)	(0.018)	(0.024)
N	178	178	178	178	178	178
Adj. R^2	0.596	0.713	0.660	0.702	0.673	0.709
RMSE	0.247	0.208	0.227	0.212	0.222	0.209

Note: Newey-West standard errors in parentheses; $*p < 0.05$, $** \ p < 0.01$, $***p < 0.001$.

Table 8, include both contemporaneous and lagged regressors, as specified in expression (11). The results show that by considering a dynamic relationship between the variables, it is possible to improve the fit of the models. Lagged values of RIC_t^E are relevant for explaining the behavior of MEG_t when $RIC_t^E$is the only explanatory variable (model

[3]), when RIC_t^E is combined with MEG_t (model [4]), and when RIC_t^E is included in the regression model along with RIC_t^S and RIC_t^O. Note, however, that RIC_t^E becomes statistically insignificant when all individual diffusion indices and their lags, in addition to lagged values of MEG_t, are included in the model specification (model [6]). The latter result is consistent with the persistent behavior of the MEG_t series described earlier (shown in Figure 4), and the fact that the series RIC_t^E, RIC_t^S, and RIC_t^O are highly correlated. In sum, all the models that include lagged values of MEG_t (specifically, models [2], [4], and [6]) have relatively low RMSEs. However, the lowest RMSE (and also the highest adjusted-R^2) is associated with model (2), with a RMSE equal to 0.208.

Finally, Table 9 presents the best dynamic specification that includes all the diffusion indices calculated by the FRBR. We report the estimates for the model that results from a stepwise variable selection process. Of all the models considered up to this point, this last specification has the highest predictive accuracy with a RMSE of 0.189. In addition to the lagged values of MEG_t, a number of diffusion indices not currently included in the reported composite index, specifically RIC_t^{IR} and RIC_t^{IF}, appear to be significantly different from zero. The model has, however, an adjusted-R^2 equal to 0.76, so it imperfectly fits the data. Figure 13 shows observed and predicted values from this specification.

VAR Models

Bivariate VAR model: MEG and FRBR composite diffusion index

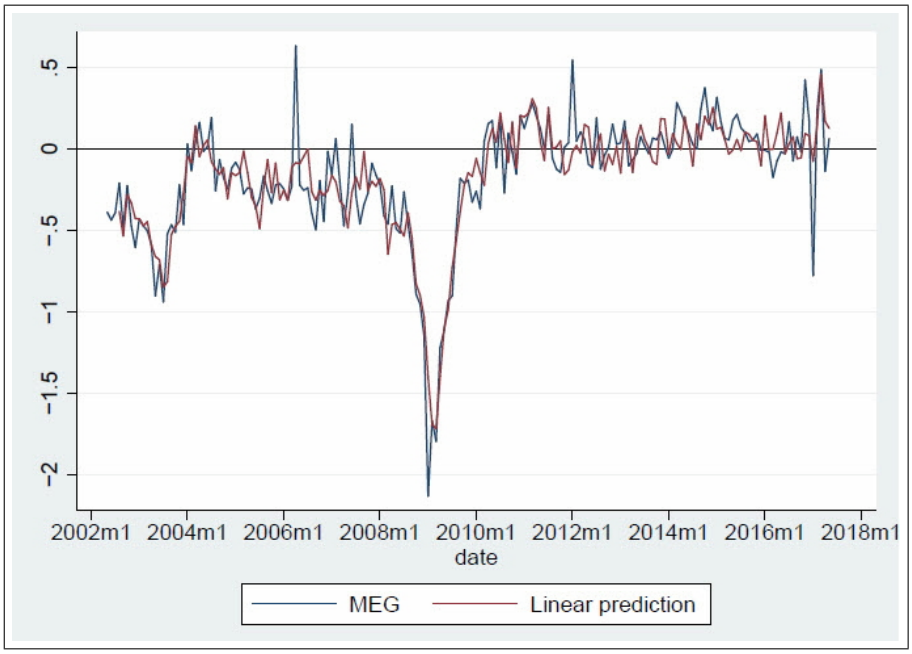
As in the ISM case, we estimate several VAR models, assess their predictive accuracy, and perform a FEVD. We begin by estimating two bivariate VAR models: one includes the FRBR composite index RIC_t and the other the FRBR employment index RIC_t^E. The results of the estimation are shown in Tables 21 and 22 in Appendix A.4.1 in addition to the observed and predicted values obtained from each model (Figures 20 and 21, respectively). Comparing the accuracy of the predictions, the specification that uses the individual diffusion index RIC_t^E has a RMSE equal to 0.211, slightly below the model that includes the composite diffusion index RIC_t, with a RMSE equal to 0.217, and lower AIC and BIC statistics. Note, however, that some of the models considered in the previous section outperform, in terms of predictive accuracy, these two VAR models. Finally, the FEVDs for each model, shown in Figures 14a and 14b, are practically identical. They indicate

Table 9 Linear Model of MEG: Stepwise Selection

	MEG_t
RIC^E_t	0.006
	(0.004)
RIC^E_{t-1}	0.003
	(0.003)
RIC^E_{t-3}	0.003
	(0.003)
RIC^O_{t-2}	0.002
	(0.002)
RIC^O_{t-3}	0.006
	(0.003)
RIC^S_t	0.005
	(0.003)
RIC^S_{t-1}	0.004
	(0.003)
RIC^S_{t-3}	-0.003
	(0.003)
RIC^C_t	-0.003
	(0.003)
RIC^C_{t-1}	-0.006
	(0.003)
RIC^V_{t-1}	0.004
	(0.004)
RIC^V_{t-2}	-0.006
	(0.004)
RIC^W_t	0.005
	(0.004)
RIC^W_{t-1}	0.005
	(0.004)
RIC^W_{t-2}	-0.008
	(0.004)
RIC^H_t	0.004
	(0.003)
RIC^H_{t-1}	-0.002
	(0.003)
RIC^H_{t-3}	0.002
	(0.003)
RIC^{IR}_{t-1}	-0.010*
	(0.004)
RIC^{IR}_{t-2}	0.004
	(0.004)
RIC^{IR}_{t-3}	0.013***
	(0.004)
RIC^{IF}_{t-1}	0.006
	(0.004)
RIC^{IF}_{t-2}	-0.012**
	(0.004)
MEG_{t-1}	0.249***
	(0.072)
MEG_{t-3}	0.155*
	(0.070)
Constant	-0.140
	(0.081)
N	178
Adjusted R^2	0.762
RMSE	0.189

Note: Newey-West standard errors in parentheses; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Figure 13 Dynamic Linear Model of MEG: Stepwise Selection

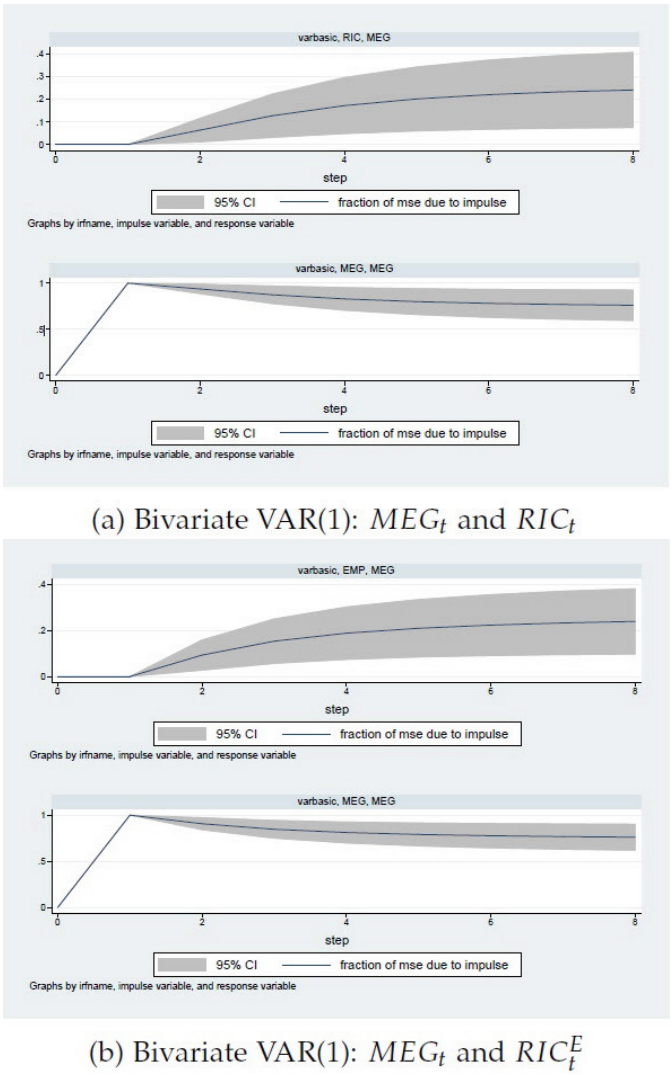


that, even though MEG explains most of the variation in the series, the FRBR diffusion indices are still relevant: RIC_t and RIC_t^E explain, in each case, about 20 percent of the variation of MEG after eight periods.

Multivariate VAR model: MEG and FRBR individual diffusion indices

Finally, we estimate a VAR model that includes additional diffusion indices computed from the FRBR survey. We follow a stepwise regression procedure to select the components considered in the analysis. The estimated values are shown in Table 23, and the predicted values from the structural equation are presented in Figure 22 in Appendix A.4.2. This model has the highest predictive accuracy of all the models considered thus far, with a RMSE of 0.131. Of all the FRBR series included in the model, RIC_t^E is still the one that explains a larger proportion of the variation in MEG_t (around 15 percent of the variance), as shown by the FEVD in Figure 15.

Figure 14 Forecast Error Variance Decomposition



(a) Bivariate VAR(1): MEG_t and RIC_t

(b) Bivariate VAR(1): MEG_t and RIC_t^E

Summary of results

From all the models considered in the previous sections, we present those with the highest predictive accuracy in Table 10, in addition to the model that includes the composite index RIC_t currently reported by the FRBR, for comparison. The VAR model that includes all the FRBR individual diffusion indices has the lowest RMSE. This model,

Figure 15 MEG_t **and FRBR Individual Diffusion Indices. Forecast Error Variance Decomposition**

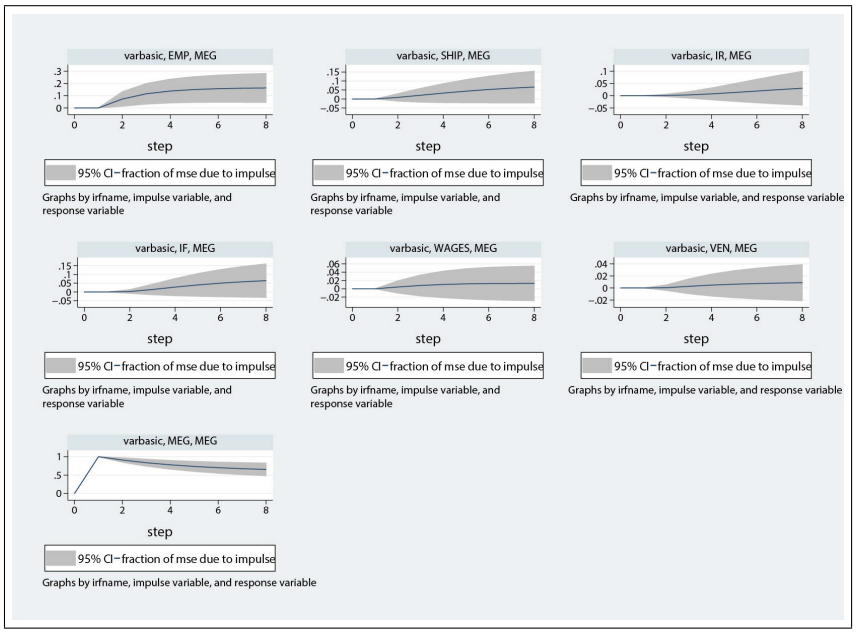


Table 10 Comparison of RMSE for Selected Models of MEG

Model		RMSE
Composite index	RIC_t	0.29
Univariate	$ARMA(1,1)$	0.23
	$ARMA(2,2)$	0.23
Linear	Contemporaneous	0.23
	Dynamic	0.19
VAR	Bivariate RIC_t	0.22
	Bivariate R_t^E	0.21
	Multivariate	0.13

with an RMSE of 0.13, is clearly an improvement compared with the model that relies only on RIC_t. All the other models, however, have approximately the same RMSEs. As in the ISM case, a linear dynamic model that includes readily available information from the FRBR survey performs reasonably well.

4. CONCLUSION

In this paper, we evaluate the information content of the FRBR manufacturing survey to determine the extent to which the diffusion indices based on the survey responses contribute to explaining national and regional economic conditions. We do so by examining the predictive accuracy of a variety of models, some of which include the composite diffusion index reported by the FRBR, and some of which incorporate additional information available from the FRBR survey but not currently employed in the calculation of the composite index.

The findings of the exercise can be summarized as follows. First, the diffusion indices currently reported by the FRBR manufacturing survey perform reasonably well at explaining the national economy, described by the evolution of the ISM diffusion index, and the regional economy, described by the evolution of the MEG. Second, in order to more accurately predict the behavior of the national and regional economy, it becomes essential to consider models that account for a richer dynamic structure given the high persistence of the series under study. And third, there are grounds for improving the predictive power of the FRBR composite index, both at national and regional levels, by adjusting the weights currently used in the calculation and by including other readily available diffusion indices. However, it should be kept in mind that the composite indices that track the national and regional economy would not necessarily be the same. This paper provides a few insights on what those diffusion indices would look like.

Future analysis should study more carefully the design of composite indices based on currently available information, including perhaps the possibility of constructing those indices based on a principal component analysis.

APPENDIX A.1: UNIT ROOT TESTS

Table 11 Unit Root Tests

Variable	Drift		Drift and Trend		ADF Test	
	t-stat	p-value	t-stat	p-value	t-stat	p-value
ISM_t	-2.591	0.095	-2.801	0.058	-3.437	0.047
MEG_t	-4.632	0.000	-5.199	0.000	-3.122	0.021
RIC_t	-5.132	0.000	-5.257	0.000	-3.902	0.012
R_t^E	-4.557	0.000	-5.130	0.000	-3.461	0.044
R_t^O	-5.708	0.000	-5.757	0.000	-4.358	0.003
R_t^S	-6.624	0.000	-6.647	0.000	-4.069	0.007
R_t^B	-6.082	0.000	-6.176	0.000	-4.587	0.001
R_t^C	-6.088	0.000	-6.134	0.000	-4.326	0.003
R_t^H	-5.717	0.000	-6.034	0.000	-5.348	0.000
R_t^W	-6.346	0.000	-6.660	0.000	-3.130	0.099
R_t^{IF}	-4.558	0.000	-4.775	0.000	-3.415	0.049
R_t^{IR}	-4.720	0.000	-5.163	0.000	-3.697	0.023
R_t^V	-5.276	0.000	-5.294	0.000	-3.749	0.019

Note: ADF: Augmented Dickey-Fuller. The number of Δ terms in the ADF is determined by the autoregressive order.

APPENDIX A.2: CROSS-CORRELOGRAMS

Table 12 Cross-correlogram between ISM_t and RIC_t^I

Lag	RIC_t^E	RIC_t^O	RIC_t^S	RIC_t^B	RIC_t^C	RIC_t^H	RIC_t^W	RIC_t^{IF}	RIC_t^{IR}	RIC_t^V
-10	-0.010	0.146	0.163	0.107	0.187	0.130	-0.073	-0.207	-0.136	0.059
-9	0.040	0.200	0.213	0.152	0.225	0.171	-0.025	-0.272	-0.176	0.105
-8	0.090	0.205	.0215	0.156	0.226	0.167	-0.018	-0.337	-0.226	0.121
-7	0.139	0.255	0.254	0.179	0.259	0.190	0.020	-0.361	-0.263	0.174
-6	0.192	0.331	0.318	0.267	0.313	0.229	0.107	-0.417	-0.315	0.245
-5	0.261	0.410	0.398	0.360	0.398	0.326	0.157	-0.467	-0.376	0.324
-4	0.345	0.525	0.499	0.470	0.479	0.415	0.226	-0.509	-0.439	0.414
-3	0.443	0.613	0.576	0.579	0.546	0.492	0.336	-0.546	-0.518	0.530
-2	0.512	0.700	0.649	0.666	0.630	0.578	0.417	-0.615	-0.566	0.641
-1	0.621	0.741	0.709	0.705	0.679	0.641	0.492	-0.638	-0.622	0.720
0	0.676	0.770	0.749	0.730	0.722	0.675	0.559	-0.655	-0.629	0.777
1	0.713	0.710	0.713	0.682	0.698	0.663	0.579	-0.651	-0.626	0.757
2	0.690	0.616	0.646	0.590	0.616	0.606	0.550	-0.616	-0.578	-0.711
3	0.633	0.481	0.536	0.460	0.497	0.490	0.527	-0.567	-0.512	0.670
4	0.583	0.380	0.417	0.363	0.399	0.392	0.455	-0.488	-0.430	0.585
5	0.507	0.274	0.313	0.268	0.303	0.287	0.418	-0.383	-0.335	0.492
6	0.437	0.189	0.206	0.175	0.222	0.193	0.367	-0.299	-0.247	0.419
7	0.362	0.128	0.138	0.118	0.148	0.111	0.342	-0.239	-0.163	0.367
8	0.305	0.087	0.099	0.105	0.113	0.092	0.324	-0.152	-0.106	0.353
9	0.269	0.080	0.092	0.093	0.104	0.059	0.318	-0.104	-0.046	0.314
10	0.237	0.035	0.049	0.062	0.050	0.017	0.301	-0.050	0.007	0.257

Table 13 Cross-correlogram between MEG_t and RIC_t^O

Lag	RIC_t^E	RIC_t^O	RIC_t^S	RIC_t^B	RIC_t^C	RIC_t^H	RIC_t^W	RIC_t^{IF}	RIC_t^{IR}	RIC_t^V
-10	0.351	0.245	0.238	0.235	0.259	0.286	0.116	-0.409	-0.064	0.110
-9	0.371	0.253	0.221	0.217	0.252	0.285	0.156	-0.436	-0.058	0.171
-8	0.410	0.287	0.274	0.266	0.309	0.284	0.172	-0.478	-0.115	0.138
-7	0.481	0.344	0.288	0.341	0.327	0.351	0.239	-0.517	-0.151	0.216
-6	0.481	0.355	0.334	0.338	0.335	0.367	0.272	-0.534	-0.166	0.273
-5	0.549	0.479	0.438	0.467	0.431	0.461	0.321	-0.552	-0.224	0.340
-4	0.623	0.524	0.479	0.506	0.473	0.485	0.436	-0.563	-0.247	0.381
-3	0.679	0.608	0.570	0.582	0.550	0.611	0.502	-0.595	-0.245	0.456
-2	0.704	0.597	0.572	0.573	0.562	0.635	0.506	-0.637	-0.301	0.476
-1	0.758	0.597	0.590	0.579	0.551	0.634	0.593	-0.587	-0.326	0.533
0	0.762	0.569	0.566	0.551	0.539	0.640	0.618	-0.601	-0.278	0.529
1	0.741	0.479	0.482	0.483	0.452	0.572	0.611	-0.527	-0.246	0.520
2	0.681	0.374	0.391	0.390	0.358	0.474	0.578	-0.507	-0.188	0.434
3	0.637	0.259	0.289	0.275	0.272	0.388	0.520	-0.440	-0.096	0.357
4	0.576	0.180	0.205	0.205	0.211	0.318	0.466	-0.352	0.024	0.269
5	0.476	0.126	0.165	0.134	0.146	0.236	0.460	-0.300	0.085	0.205
6	0.436	0.048	0.054	0.077	0.059	0.168	0.388	-0.190	0.160	0.139
7	0.352	0.027	0.055	0.067	0.037	0.129	0.395	-0.143	0.209	0.122
8	0.306	-0.009	-0.007	0.017	-0.019	0.083	0.359	-0.078	0.249	0.108
9	0.228	-0.057	-0.046	-0.009	-0.067	0.028	0.333	-0.029	0.300	0.056
10	0.215	-0.081	-0.086	0.055	-0.077	0.024	0.307	0.020	0.312	-0.036

Table 14 Cross-correlogram of RIC_t^E with FRBR Diffusion Indices

Lag	RIC_t^O	RIC_t^S	RIC_t^B	RIC_t^C	RIC_t^H	RIC_t^W	RIC_t^{IF}	RIC_t^{IR}	RIC_t^V
-10	0.215	0.209	0.243	0.252	0.259	0.080	-0.409	-0.098	0.101
-9	0.208	0.185	0.210	0.249	0.266	0.053	-0.446	-0.123	0.121
-8	0.239	0.231	0.231	0.271	0.261	0.123	-0.501	-0.127	0.155
-7	0.282	0.260	0.268	0.267	0.264	0.198	-0.483	-0.174	0.199
-6	0.367	0.329	0.341	0.331	0.314	0.278	-0.526	-0.172	0.278
-5	0.423	0.398	0.425	0.377	0.379	0.337	-0.532	-0.210	0.331
-4	0.517	0.485	0.502	0.474	0.502	0.406	-0.565	-0.243	0.389
-3	0.580	0.520	0.553	0.504	0.517	0.483	-0.606	-0.315	0.460
-2	0.635	0.576	0.640	0.580	0.627	0.546	-0.640	-0.340	0.553
-1	0.668	0.613	0.655	0.653	0.709	0.554	-0.670	-0.350	0.567
0	0.681	0.635	0.660	0.683	0.774	0.624	-0.625	-0.332	0.564
1	0.556	0.532	0.538	0.559	0.655	0.608	-0.609	-0.257	0.498
2	0.389	0.399	0.378	0.418	0.481	0.552	-0.538	-0.197	0.458
3	0.303	0.331	0.289	0.319	0.391	0.496	-0.485	-0.117	0.347
4	0.175	0.189	0.187	0.191	0.298	0.458	-0.366	-0.009	0.291
5	0.109	0.121	0.122	0.105	0.195	0.383	-0.293	0.078	0.237
6	0.056	0.058	0.053	0.030	0.114	0.384	-0.211	0.138	0.162
7	0.036	0.026	0.094	0.030	0.106	0.364	-0.154	0.211	0.160
8	0.029	0.031	0.074	0.018	0.117	0.365	-0.115	0.223	0.151
9	0.011	0.033	0.077	0.004	0.134	0.351	-0.097	0.242	0.117
10	0.001	0.017	0.044	0.020	0.110	0.352	-0.090	0.266	0.060

Table 15 Cross-correlogram of RIC_t^O with FRBR Diffusion Indices

Lag	RIC_t^E	RIC_t^S	RIC_t^B	RIC_t^C	RIC_t^H	RIC_t^W	RIC_t^{IF}	RIC_t^{IR}	RIC_t^V
-10	0.001	0.074	0.093	0.101	0.085	-0.068	-0.087	0.020	0.019
-9	0.011	0.088	0.084	0.132	0.083	-0.071	-0.161	-0.041	0.022
-8	0.029	0.093	0.031	0.116	0.054	-0.056	-0.227	-0.054	-0.025
-7	0.036	0.103	0.034	0.095	0.068	-0.018	-0.203	0.082	0.040
-6	0.056	0.124	0.080	0.105	0.058	0.015	-0.270	-0.121	0.108
-5	0.109	0.202	0.175	0.172	0.121	0.067	-0.249	-0.127	0.135
-4	0.175	0.237	0.290	0.244	0.181	0.169	-0.302	-0.200	0.204
-3	0.303	0.374	0.402	0.351	0.298	0.217	-0.399	-0.315	0.274
-2	0.389	0.476	0.553	0.477	0.463	0.359	-0.475	-0.354	0.437
-1	0.556	0.629	0.646	0.646	0.632	0.416	-0.505	-0.467	0.548
0	0.681	0.921	0.897	0.903	0.814	0.546	-0.577	-0.548	0.676
1	0.668	0.647	0.670	0.689	0.669	0.541	-0.576	-0.517	0.639
2	0.635	0.575	0.505	0.555	0.555	0.464	-0.536	-0.459	0.600
3	0.580	0.457	0.376	0.450	0.420	0.451	-0.500	-0.404	0.499
4	0.517	0.291	0.241	0.294	0.350	0.379	-0.421	-0.318	0.455
5	0.423	0.236	0.173	0.227	0.241	0.365	-0.339	-0.243	0.373
6	0.367	0.153	0.131	0.141	0.139	0.284	-0.240	-0.173	0.306
7	0.282	0.101	0.106	0.097	0.089	0.285	-0.195	-0.048	0.308
8	0.239	0.060	0.101	0.046	0.071	0.278	-0.143	-0.037	0.315
9	0.208	0.114	0.135	0.098	0.067	0.297	-0.119	-0.003	0.279
10	0.215	0.057	0.108	0.082	0.062	0.315	-0.073	0.040	0.256

Table 16 Cross-correlogram of RIC_t^S with FRBR Diffusion Indices

Lag	RIC_t^E	RIC_t^O	RIC_t^B	RIC_t^C	RIC_t^H	RIC_t^W	RIC_t^{IF}	RIC_t^{IR}	RIC_t^V
-10	0.017	0.057	0.085	0.116	0.127	-0.056	-0.117	0.015	0.039
-9	0.033	0.114	0.072	0.152	0.075	-0.071	-0.183	-0.049	0.040
-8	0.031	0.060	0.022	0.115	0.041	-0.066	-0.204	-0.031	-0.035
-7	0.026	0.101	0.045	0.112	0.077	-0.025	-0.209	-0.073	0.032
-6	0.058	0.153	0.093	0.131	0.093	0.017	-0.271	-0.156	0.145
-5	0.121	0.236	0.214	0.189	0.142	0.059	-0.295	-0.167	0.164
-4	0.189	0.291	0.321	0.277	0.225	0.169	-0.316	-0.235	0.256
-3	0.331	0.457	0.434	0.388	0.331	0.225	-0.406	-0.338	0.332
-2	0.399	0.575	0.556	0.505	0.473	0.385	-0.480	-0.355	0.463
-1	0.532	0.647	0.596	0.615	0.625	0.422	-0.526	-0.460	0.536
0	0.635	0.921	0.809	0.897	0.780	0.521	-0.609	-0.542	0.633
1	0.613	0.629	0.608	0.634	0.624	0.488	-0.589	-0.526	0.607
2	0.576	0.476	0.442	0.466	0.483	0.424	-0.516	-0.451	0.555
3	0.520	0.374	0.347	0.410	0.378	0.413	-0.480	-0.410	0.489
4	0.485	0.237	0.179	0.267	0.307	0.347	-0.424	-0.317	0.427
5	0.398	0.202	0.157	0.212	0.235	0.349	-0.344	-0.266	0.351
6	0.329	0.124	0.113	0.150	0.123	0.274	-0.263	-0.179	0.288
7	0.260	0.103	0.100	0.114	0.077	0.255	-0.230	-0.061	0.293
8	0.231	0.093	0.122	0.084	0.074	0.307	-0.161	-0.050	0.323
9	0.185	0.088	0.142	0.089	0.066	0.299	-0.117	-0.012	0.263
10	0.209	0.074	0.095	0.105	0.047	0.291	-0.112	-0.002	0.238

APPENDIX A.3: NATIONAL ECONOMY: VAR MODELS

A.3.1 UNIVARIATE MODELS OF ISM

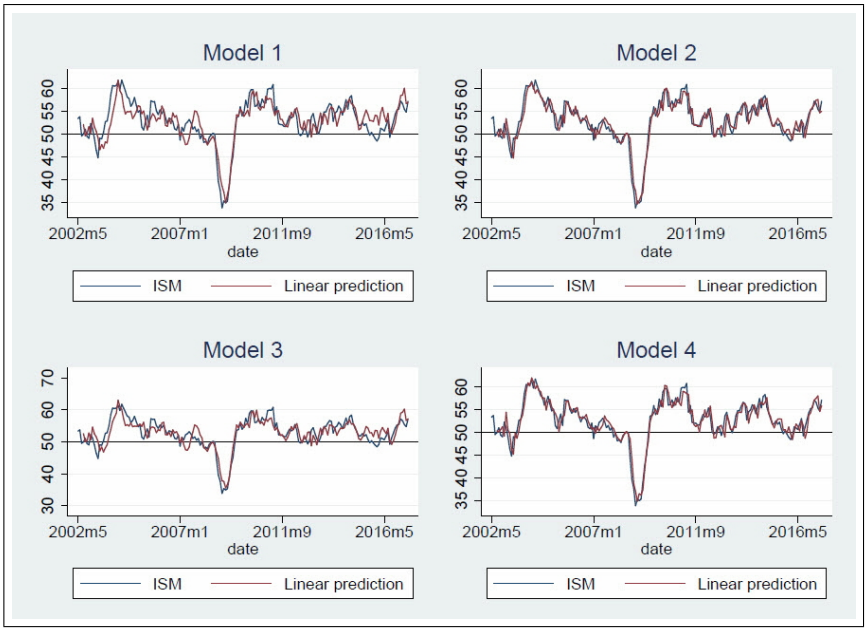
Table 17 Univariate Models of ISM_t

	(1) AR(1)	(2) AR(4)	(3) ARMA(1,1)	(4) ARMA(1,2)	(5) ARMA(4,1)
ISM_{t-1}	0.921***	0.972***	0.909***	0.892***	0.984
	(0.023)	(0.065)	(0.028)	(0.037)	(0.556)
ISM_{t-2}		0.059			0.047
		(0.101)			(0.543)
ISM_{t-3}		0.016			0.015
		(0.120)			(0.128)
ISM_{t-4}		-0.160*			-0.158
		(0.078)			(0.115)
ϵ_{t-1}			0.084	0.070	-0.012
			(0.073)	(0.077)	(0.568)
ϵ_{t-2}				0.137	
				(0.084)	
Constant	53.045***	52.953***	53.018***	53.018***	52.953***
	(1.818)	(1.285)	(1.696)	(1.596)	(1.283)
N	182	182	182	182	182
AIC	746.537	742.370	747.163	746.366	744.369
BIC	756.150	761.594	759.979	762.386	766.797
RMSE	1.841	1.789	1.834	1.820	1.789

Note: Standard errors in parentheses; * $p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

A.3.2 LINEAR MODELS

Figure 16 Linear models of ISM with Contemporaneous and Lagged Regressors. Observed Values and Predictions

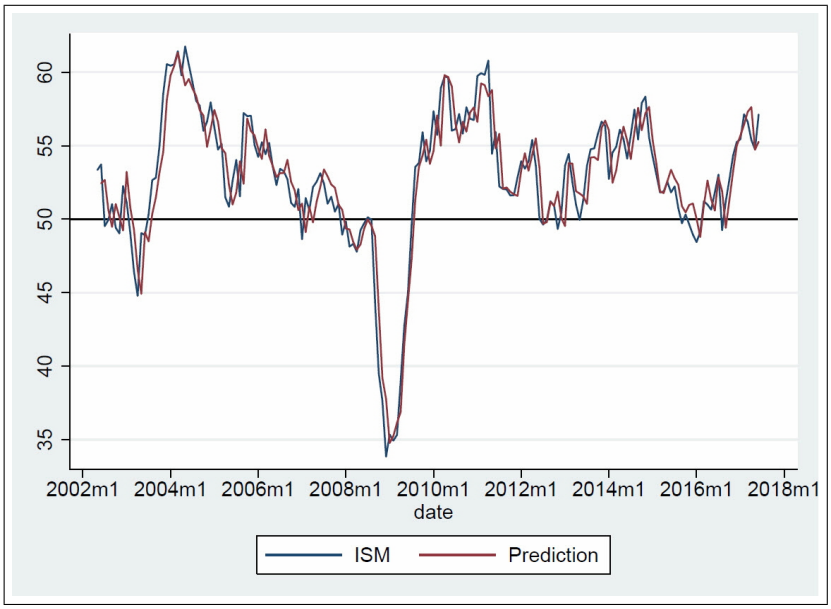


A.3.3 BIVARIATE VAR: ISM AND FRBR COMPOSITE DIFFUSION INDICES

Table 18 VAR(1): ISM_t and RIC_t. Estimates, Variance-Covariance Matrix, and Cholesky Decomposition

Φ	RIC_t	0.35***	0.59***
	ISM_t	0.05	0.88***
α	RIC_t	1.62	
	ISM_t	3.97***	
Σ	RIC_t	12.27	
	ISM_t	2.31	3.38
$p-1$	RIC_t	0.29	0
	ISM_t	-0.11	0.58
N		181	
AIC		9.33	
BIC		9.43	
RMSE		1.72	

Figure 17 VAR(1): ISM_t and RIC_t. Observed and Predicted Values



A.3.4 MULTIVARIATE VAR: ISM AND FRBR INDIVIDUAL DIFFUSION INDICES

Table 19 VAR(2): ISM_t, RIC_t^E, RIC_t^O, and RIC_t^S. Estimates, Variance-Covariance Matrix, and Cholesky Decomposition

	RIC_t^E	0.38***	0.12***	-0.06	0.43***
Φ_1	RIC_t^O	0.32***	0.31**	-0.07	0.98***
	RIC_t^S	0.19	0.22*	-0.07	0.80***
	ISM_t	0.07	0.05	-0.02	0.94***
	RIC_t^E	0.29***	0.03	-0.08	-0.18
Φ_2	RIC_t^O	-0.35***	0.39***	-0.36***	-0.34*
	RIC_t^S	-0.28*	0.28**	-0.21*	-0.07
	ISM_t	-0.16***	0.13**	-0.09	-0.06
	RIC_t^E	3.06			
α	RIC_t^O	3.93			
	RIC_t^S	5.61			
	ISM_t	6.99***			
	RIC_t^E	6.86			
Σ	RIC_t^O	4.48	18.68		
	RIC_t^S	3.28	15.34	18.40	
	ISM_t	0.67	2.65	2.00	3.05
	RIC_t^E	0.38	0	0	0
$p-1$	RIC_t^O	-0.16	0.25	0	0
	RIC_t^S	0.03	-0.35	0.42	0
	ISM_t	0.0001	-0.10	0.02	0.61
N		180			
AIC		19.17			
BIC		19.43			
RMSE		1.64			

Figure 18 VAR(2): ISM, RIC_t^E, RIC_t^O, and RIC_t^S. Observed and Predicted Values

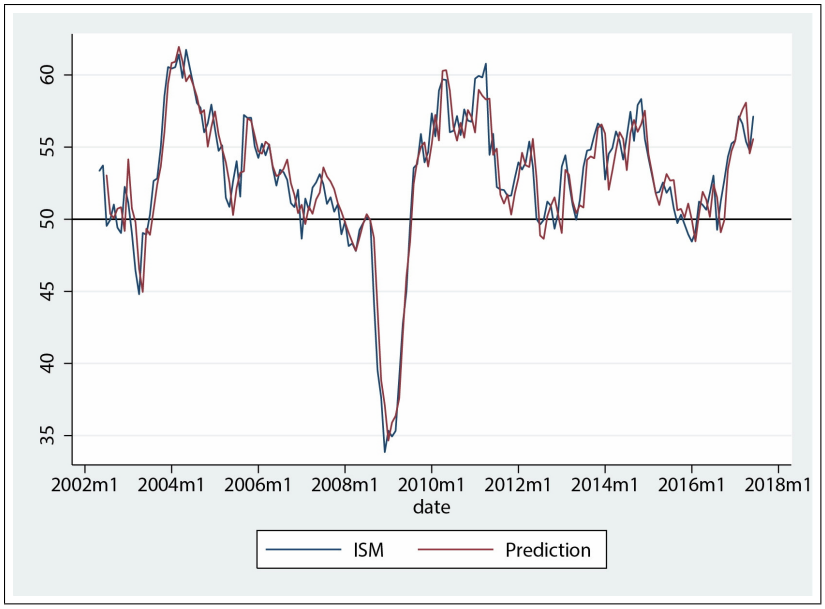
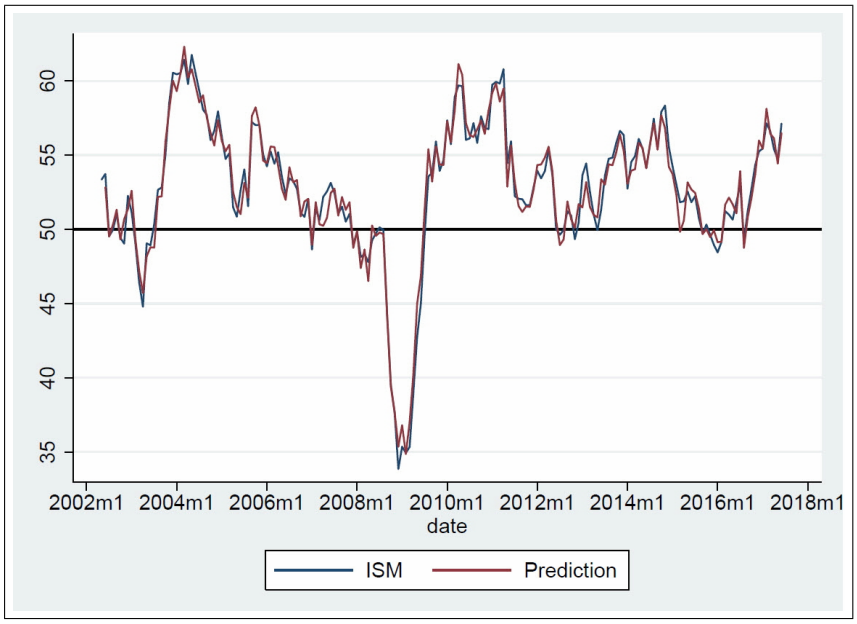


Table 20 VAR(1): ISM and FRBR Diffusion Indices. Estimates, Variance-Covariance Matrix, and Cholesky Decomposition

	RIC_t^{IR}	0.54***	-0.60*	0.03	0.17**	-0.03	0.13***	-0.26***
	RIC_t^{O}	-0.01	0.35***	0.01	-0.10	-0.14	0.08	0.69***
	RIC_t^{IF}	0.09	-0.30	0.57***	-0.05	0.19*	-0.10	-0.19**
Φ	RIC_t^{W}	0.13**	0.04	0.00	0.33***	0.02	0.11**	0.18**
	RIC_t^{V}	-0.11*	0.04	0.03	0.10	0.31***	-0.05	0.29***
	RIC_t^{E}	0.25***	0.08	-0.30***	0.03	0.02	0.43***	0.29***
	ISM_t	-0.04	0.06	-0.06	-0.06	0.01	-0.03	0.84***
	RIC_t^{IR}	26.60***						
	RIC_t^{O}	4.91						
	RIC_t^{IF}	29.16***						
α	RIC_t^{W}	12.92**						
	RIC_t^{V}	20.79***						
	RIC_t^{E}	11.84						
	ISM_t	15.24***						
	RIC_t^{IR}	3.84						
	RIC_t^{O}	-2.20	21.02					
	RIC_t^{IF}	1.35	-2.03	5.31				
Σ	RIC_t^{W}	0.27	2.01	0.36	3.82			
	RIC_t^{V}	-0.45	3.01	-0.16	1.23	3.86		
	RIC_t^{E}	-0.04	4.11	-0.17	0.95	0.55	6.85	
	ISM_t	-0.33	3.38	-0.61	0.50	1.18	0.31	3.26
	RIC_t^{IR}	1.96	0	0	0	0	0	0
	RIC_t^{O}	-1.12	4.44	0	0	0	0	0
	RIC_t^{IF}	0.70	-0.28	2.17	0	0	0	0
$p-1$	RIC_t^{W}	0.14	0.49	0.18	1.88	0	0	0
	RIC_t^{V}	-0.23	0.63	0.08	0.50	1.78	0	0
	RIC_t^{E}	-0.02	0.92	0.05	0.26	-0.09	2.43	0
	ISM_t	-0.17	0.72	-0.13	0.10	0.36	-0.14	1.59
N		181						
AIC		31.49						
BIC		32.48						
RMSE		0.85						

Figure 19 VAR(1): ISM and FRBR Diffusion Indices. Observed and Predicted Values



APPENDIX A.4 REGIONAL ECONOMY: VAR MODELS

A.4.1 BIVARIATE VAR: MEG AND FRBR COMPOSITE INDEX

Table 21 Bivariate VAR(1): MEG_t and RIC_t. Estimates, Variance-Covariance Matrix, and Cholesky Decomposition

	RIC_t	0.635***	4.851**
Φ	MEG_t	0.010***	0.596***
	RIC_t	1.081	
α	MEG_t	-0.074***	
	RIC_t	58.520	
Σ	MEG_t	0.448	0.050
	RIC_t	0.131	0
$p-1$	MEG_t	-0.035	4.619
N		180	
AIC		6.751	
BIC		6.858	
RMSE		0.217	

Figure 20 Bivariate VAR(1): MEG_t and RIC_t. Observed and Predicted Values

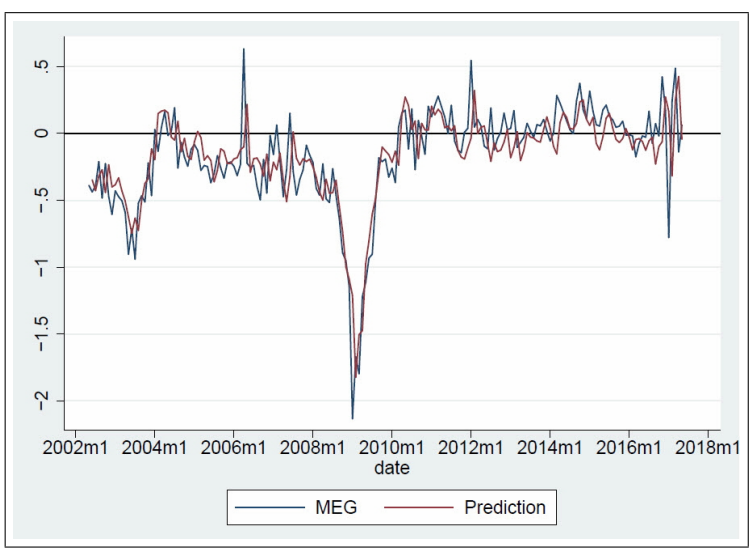
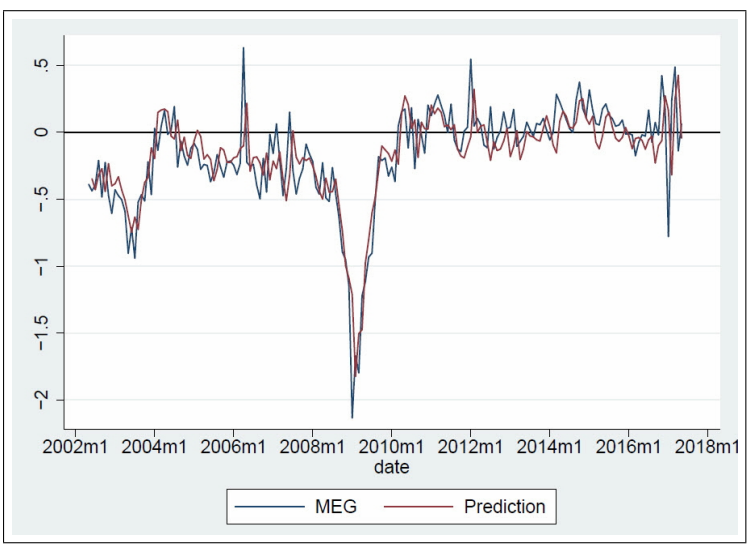


Table 22 Bivariate VAR(1): MEG_t and RIC_t^E. Estimates, Variance-Covariance Matrix, and Cholesky Decomposition

	RIC_t^E	0.542***	8.534***
Φ	MEG_t	0.015***	0.490***
	RIC_t^E	1.349**	
α	MEG_t	-0.079***	
	RIC_t^E	32.348	
Σ	MEG_t	0.349	0.048
	RIC_t^E	0.176	0
$p-1$	MEG_t	-0.051	4.761
N		180	
AIC		6.098	
BIC		6.205	
RMSE		0.211	

Figure 21 Bivariate VAR(1): MEG_t and RIC_t^E. Observed and Predicted Values



A.4.2 MULTIVARIATE VAR MODEL: MEG AND FRBR INDIVIDUAL DIFFUSION INDICES

Figure 22 VAR(1) model: MEG and FRBR Individual Diffusion Indices. Observed and Predicted Values

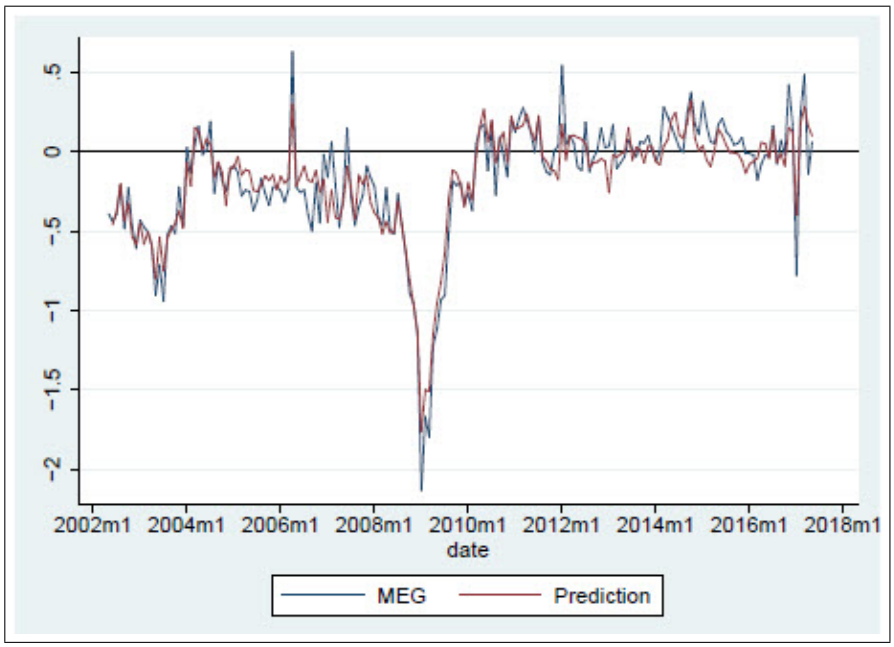


Table 23 VAR(1): MEG and FRBR Individual Diffusion Indices. Estimates, Variance-Covariance Matrix, and Cholesky Decomposition

Φ	RIC_t^{IR}	0.598***	-0.744**	0.096*	0.162**	-0.177**	0.11**	-0.68
	RIC_t^{S}	-0.279*	0.311***	-0.201	0.308*	0.163	0.101	-0.294
	RIC_t^{IF}	0.113**	-0.045	0.621***	0.001	0.066	-0.114**	0.391
	RIC_t^{W}	0.002	0.022	-0.008	0.454***	0.058	0.049	2.855**
	RIC_t^{V}	-0.204***	0.073**	0.003	0.115	0.445***	-0.076	3.085***
	RIC_t^{E}	0.050	0.082	-0.247***	0.178*	0.111	0.398***	4.434***
	MEG_t	0.001	0.001	0.005	0.003	0.001	0.009***	0.484***
α	RIC_t^{IR}	3.810***						
	RIC_t^{S}	5.765*						
	RIC_t^{IF}	5.151***						
	RIC_t^{W}	5.777**						
	RIC_t^{V}	4.811***						
	RIC_t^{E}	2.254						
	MEG_t	-0.028						
Σ	RIC_t^{IR}	18.564						
	RIC_t^{S}	-12.655	117.450					
	RIC_t^{IF}	6.973	-9.740	23.974				
	RIC_t^{W}	-1.30	13.24	0.32	20.07			
	RIC_t^{V}	-4.096	15.485	-1.958	6.481	17.618		
	RIC_t^{E}	-0.906	21.629	-0.698	5.271	3.672	32.152	
	MEG_t	-0.029	0.501	-0.225	0.119	0.074	0.208	0.048
$p-1$	RIC_t^{IR}	0.232	0	0	0	0	0	0
	RIC_t^{S}	0.065	0.096	0	0	0	0	0
	RIC_t^{IF}	-0.076	0.010	0.218	0	0	0	0
	RIC_t^{W}	0.003	-0.027	-0.015	0.233	0	0	0
	RIC_t^{V}	0.037	-0.023	0.002	-0.069	0.267	0	0
	RIC_t^{E}	-0.015	-0.033	-0.004	-0.028	-0.004	0.190	0
	MEG_t	-0.020	-0.014	0.045	-0.017	0.003	-0.020	4.830
N		180						
AIC		36.876						
BIC		37.870						
RMSE		0.131						

REFERENCES

Harris, Matthew, Raymond E. Owens, and Pierre-Daniel G. Sarte. 2004. "Using Manufacturing Surveys to Assess Economic Conditions." Federal Reserve Bank of Richmond *Economic Quarterly* 90 (Fall): 65–93.

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

Pinto, Santiago, Sonya Ravindranath Waddell, and Pierre-Daniel G. Sarte. 2015b. "Monitoring Economic Activity in Real Time Using Diffusion Indices: Evidence from the Fifth District." Federal Reserve Bank of Richmond *Economic Quarterly* 101 (Fourth Quarter): 275–301.