Misallocation and Productivity in the Great Depression

Nicolas L. Ziebarth*

January 16, 2014

Abstract

 Aggregate productivity fell by 18% between 1929 and 1933. Existing explanations for this decline have focused on unobserved shifts in factor inputs such as labor hoarding. I develop a new hypothesis that focuses on the role of resource misallocation. Using a novel plant-level dataset covering a wide swath of manufacturing industries, I find that dispersion in revenue productivity increases by approximately 7.5% between 1929 and 1933 indicative of greater misallocation as in Hsieh and Klenow (2009). I then study the cement industry in detail because of its homogeneous product. I decompose the overall change in industry-level productivity into efficient productivity shifts and misallocation. Increases in misallocation between 1929 and 1931 can explain at a minimum 50% of the decline in productivity for the cement industry, around 20% between 1931 and 1933, and 10 to 15% for 1933 to 1935. In order to explain these findings, I develop a model of financial frictions that relates misallocation to dispersion in working capital interest rates. I argue that these empirical and theoretical results provide another role for the non-monetary effects of the banking crisis during the Depression (Bernanke, 1983a): the collapse of aggregate productivity.

1 Introduction

Why did productivity decline so much during the Great Depression? Measured aggregate TFP fell by 18% relative to trend from 1929 to 1933 before beginning to recover in 1935 (See Figure 1). Amaral and MacGee (2002) find a decline closer to 20% when the government and agricultural sectors are excluded. Ohanian (2001) estimates that only 1/3 of the fall can be explained by factors such as capacity utilization, factor input quality, labor composition, labor hoarding, and increasing returns to scale.1 The decline in productivity is even more striking when put in the

---

*University of Iowa and NBER. E-mail: nicolas.lehmannziebarth@gmail.com. I thank my dissertation committee members at Northwestern University: Joel Mokyr, Jonathan Parker, and Joe Ferrie for detailed comments and for funding part of the data collection. I also received useful feedback from Mark Carlson, Mark Chicu, Bob Gordon, Paul Grieco, Josh Hausman, Matthias Kehrig, Erik Loualiche, Ryan Marsh, Lee Ohanian, Giorgio Primiceri, Chris Vickers, Thijs van Rens, Hal Cole and seminar participants at numerous seminars. Other portions of the data collection were funded by the Graduate School and the Center for the Study of Industrial Organization. William Creech at the National Archives provided useful information on the data. Stephen Karlson kindly provided some data on the cement industry. Mary Hansen helped organize data collection at the National Archives. Doug Bojack and Joanna Gregson did yeoman’s work photographing the original schedules. Dan Thomas was absolutely indispensable. I also thank the Robert Eisner Memorial Fellowship as well as the Sokoloff Fellowship for providing support.

1This is not completely uncontroversial. Solow (1957) as well as Fabricant (1940) estimated productivity adjusting for utilization and found that productivity did not fall. Inklaar et al. (2011) find large increasing returns for manufacturing and conclude on this basis that productivity did not fall.
context of the major innovations being made at this time. To the point, Field (2011) has argued
that the Depression was actually a “great leap forward” for the U.S. economy. Bernstein (1987)
documented a number of innovations such as Teflon in petrochemicals industries and household
appliances including the radio and refrigerator. These breakthroughs should have had a positive
impact on productivity all else equal. This makes the decline that much more striking and puzzling.²

This feature of the Depression is not unique to the U.S. experience. Calvo et al. (2006) find
in a broad set of emerging market crises that output collapses and recoveries are associated with
collapses and recoveries in TFP much like the U.S. Great Depression. In an even broader set of
depressions not just those associated with sudden stops, a similar pattern of sharp productivity
debutes emerges (Kehoe and Prescott, 2002) including the French Great Depression (Bridji, 2013).
In addition to these declines in TFP and output, usually there is an associated deterioration in credit
market conditions. What these papers all highlight is that in an accounting sense, understanding
depressions is closely tied to understanding the dynamics of aggregate productivity.

This paper makes two contributions to understanding the evolution of productivity during the
Great Depression. First, I offer empirical evidence for an increase in misallocation and then docu-
ment for a particular industry the fraction of the productivity decline due to increased misallocation.
Second, there is theoretical and empirical evidence to believe that this increase in misallocation is
driven by the banking crisis and, in particular, the collapse in interbank lending between 1929 and
1933.

To measure misallocation, I build on the methodology used in both the growth (Hsieh and
Klenow, 2009) (HK) and business cycle (Chari et al., 2007) literatures that measures deviations
from a neoclassical benchmark.³ The crucial distinction necessary for measuring misallocation is
the difference between physical productivity TFPQ and revenue productivity TFPR, which by
definition is equal to TFPQ * p, where p is the price a plant charges. The difference is in the units
productivity is measured in. The first is in terms of raw output, the second in terms of dollars.

²The seeming disconnect between Field (2011) and Ohanian (2001) is more apparent than real since the time
periods they study are slightly asynchronous. Field studies the whole decade of the 1930s while Ohanian focuses on
the first half of the 1930s and in fact, notes the remarkably strong productivity growth after 1933.
³At the same time, these two strands of the literature have used this framework to different ends. For example, in
the business cycle literature, Chari et al. (2002) argue that a “wedge” between the marginal rate of substitution and
the real wage during the Great Depression can explain the collapse in employment. In the growth literature, Hsieh
and Klenow (2009) have suggested that 30-50% of the differences in TFP between the U.S. and China and India can
be explained by greater misallocation in the developing countries.
An efficient allocation of resources along the intensive margin, taking as given the distribution of \( TFPQ \) will equalize \( TFPR \) across heterogeneous plants. Therefore, differences in \( TFPR \) across plants can be interpreted as a misallocation of resources.\(^4\)

To estimate the empirical relevance of misallocation, I build a novel plant-level dataset from the Census of Manufactures taken in 1929, 1931, 1933 and 1935 of 11 industries. I construct a measure of labor revenue productivity due to limitations on information for plant-level capital stocks. I find that the dispersion of this variable increases by approximately 7.5%. This is driven by a spreading out in both tails rather than only in the left tail. I then turn to a detailed analysis of one of the industry: cement. This industry has a number of convenient features. First, its product is homogeneous making price and productivity comparisons valid. Second, for this industry, I observe a good measure of a plant’s capital stock and physical output, so I can estimate plant productivity directly. This is important because HK, who do not have output measures, have to lean heavily on a particular demand structure to infer physical output from revenue. Depending on the assumed elasticity of substitution between varieties for cement, an increase in misallocation can explain anywhere between 50% and 100% of the overall decline in productivity between 1929 and 1931. Between 1929 and 1933, increases in misallocation still play a crucial role explaining around 30%. Misallocation continues to be important for changes in productivity in the cement industry explaining between 10 and 20% for 1935.

I stress the role of financial frictions in rationalizing these changes in misallocation and their effects on productivity. Not surprisingly, the literature on the role of financial frictions on business cycles has been strongly influenced by the experience of the Depression. The classic work of Friedman and Schwartz (1971) argued the banking crisis played a major causal role in the collapse of the real economy. However, they assigned a purely monetary role to the wave of bank failures between 1929 and 1933. The work of Bernanke (1983b) was the first to argue that the effect of bank failures was not solely felt through a decline in the money supply but also in more direct ways through a loss of informational capital embodied in banks.

\(^4\)I do not specify whether these differences are related to so-called output, labor, or materials wedges. In addition to the dispersion in \( TFPR \), there are other dimensions of misallocation. For one, I consider the correlation between \( TFPR \) and \( TFPQ \). Restuccia and Rogerson (2008) emphasize this correlation as particularly deleterious for productivity since when the correlation is positive, the most physically productive firms also face the highest implicit taxes. In a calibrated model similar in spirit to HK, they show that \( i.i.d. \) dispersion in \( TFPR \) across plants does not lead to large losses in productive efficiency, but a strong positive correlation does.
Two empirical observations motivate my theoretical focus on financial frictions. First, I document that the dispersion of state-level lending rates, a proxy for working capital rates, showed a sharp increase between 1929 and 1933 followed by a decline between 1933 to 1935. This timing exactly matched the changes in productivity and misallocation. At a time of more or less unit banking, the interbank market was crucial for moderating regional differences in lending rates. So the second observation reemphasizes the well-known fact that the interbank market for funds collapsed between 1929 and 1933. Taken together, these facts show the causal chain from banking to real outcomes I attempt to capture in my formal model.

The focus of my model is quite distinct from the classic work by Bernanke (1983b) and almost all subsequent work on financial frictions and the real economy. The papers in this literature have tended to focus on the role played by banks in funding long-term investments (and potentially consumer durable purchases). A central insight of these models was that small changes in the level of bank capital could have large effects on investment through constraints on leverage due to limited commitment and moral hazard problems. This view of how financial frictions matter has led people to solely focus on anomalous moves in investment as evidence for financial frictions. For example, Carlstrom and Fuerst (1997) show that the classic financial frictions model of Bernanke and Gertler (1989) can be written as the standard neoclassical growth model with a time-varying “tax” on investment. In that model, there is no impact of financial frictions on productivity. This line of thinking was brought to fruition in the business cycle accounting framework of Chari et al. (2007). They applied their methodology to the Great Depression concluding that financial frictions were of limited importance.

Recently, new literature including Buera and Moll (2013), Gilchrist et al. (2010), and Khan and Thomas (2013) has argued that financial frictions can manifest themselves in not only the so-called investment wedge but also the labor wedge and, for my purposes, productivity. Relative to the existing literature, I explicitly attempt to model the joint dynamics of the $TFPR$ and $TFPQ$ distributions including first and second moments as well as the correlation between the two. Existing literature has abstracted away one or more of these moments. First, all the previous work has considered competitive models with no price dispersion. This prevents them from distinguishing between $TFPR$ and $TFPQ$. Second, part of the literature such as Buera and Moll (2013) has
focused on what I would call misallocation on the extensive margin.\textsuperscript{5} By this I mean that aggregate productivity is solely a function of some threshold level of plant-level productivity. This allows for a characterization of the dynamics of the mean and variance of $TFPQ$. At the other extreme, Khan and Thomas (2013) and Gilchrist et al. (2010) assume exogenous entry and exit, which excludes the possibility of endogenous changes in moments of $TFPQ$, but they gain the ability to understand misallocation on the intensive margin.\textsuperscript{6}

The richness in my model comes at a cost of focusing on a basically static problem. In particular, all borrowing is one period debt to fund working capital for labor. Furthermore, there is no internal saving on the part of the firms (or banks). I model the friction as falling on banks though this is simply for convenience. Banks face a convex cost of taking on more leverage by borrowing from interbank markets to fund working capital loans. In an efficient allocation of resources, the distribution of bank leverage ratios “matches” the distribution of plant productivities. Banks with good opportunities should be highly leveraged and those without opportunities should fund the good banks. The friction that limits bank leverage leads to differences in working capital interest rates and, hence, limits the amount of reallocation from unproductive to productive plants. This in turn generates dispersion in $TFPR$, a positive correlation between $TFPR$ and $TFPQ$, and an inefficiently low level of productivity.

The shock that I emphasize is one to the marginal cost of leverage (a “credit crunch”). This leads resources to become stuck in areas where demand is relatively low or with plants that are relatively physically unproductive. The shock leads to a flattening of the distribution of leverage ratios across banks. In addition, dispersion in working capital interest rates rises, thereby increasing dispersion in $TFPR$. Moreover, the correlation between $TFPR$ and $TFPQ$ increases as the higher cost of leverage falls most severely on the most productive plants. This generates an endogenous decline in productivity. Note that this mechanism does not emphasize changes in the overall leverage ratio of the financial sector nor the amount of bank capital \textit{per se}. These actually have no effect on dispersion and productivity in the model with no entry and exit and limited effects in the model with an extensive margin. Instead the relevant summary statistic for the health of the banking

\textsuperscript{5}Yang (2012) develops a different approach to measuring misallocation on this margin.

\textsuperscript{6}To be clear, in the accounting part of the paper, I only interpret misallocation as related to changes in $TFPR$ not the distribution of $TFPQ$. The model would allow me to account for changes in productivity due to changes in both of those distributions.
sector as it affects productivity is the dispersion of leverage ratios across financial firms and the dispersion of plant borrowing rates echoing Gilchrist and Zakrajsek (2012).\(^7\) A calibrated version of my model can replicate the empirical findings of a 2% decline in productivity due to misallocation with a 50% increase in the marginal cost of leverage. This shock has a multiplier effect on bank capital, labor, and output with each of these falling closer to 10%.

The paper contributes to a growing empirical literature that studies firm dynamics over the business cycle. The classic work by Davis and Haltiwanger (1990) pointed to the tremendous value of micro-data for macro questions. Since then there has been a flourishing of work attempting to understand not only employment but also productivity (for the U.S. (Kehrig, 2011) and for France (Osotimehin, 2012)), pricing (Berger and Vavra, 2011), investment (Bachmann and Bayer, 2013), and capital reallocation (Eisfeldt and Rampini, 2006).

It is also useful to contrast the approach of this paper with much of the recent literature on financial frictions in the Great Recession. While there has been some literature emphasizing the effects on financing on businesses (Chodorow-Reich, 2013; Greenstone and Mas, 2012; Haltenhof et al., 2012), much of the work has focused on the role of financial frictions on the household side. For example, Mian and Sufi (2009, 2011, 2012) all focus on the role of the housing price collapse and its effect on consumer demand.\(^8\) This paper on the other hand focuses on the producer side and in particular, the supply effects of financial frictions. This makes it different from Bernanke (1983b) which also argued for demand effects through declines in investment.

## 2 Inferring Misallocation

To estimate the role of misallocation on industry productivity, I build on the accounting framework of HK. Industry output is a Dixit-Stiglitz aggregator of differentiated output from a continuum of plants

\[
Y = \left( \int_0^1 Y_i \frac{\psi - 1}{\psi} \, di \right)^{\frac{\psi}{\psi - 1}}
\]

\(^7\)Though this latter measure does not summarize the correlation between the rates and the productivity of the plants that pay them.

\(^8\)I am unaware of any work for the Depression that studies the direct effects of the financial collapse on household behavior. This would appear to me a rather glaring gap in the literature.
Plant $i$ operates a constant returns to scale production function.\footnote{It is also possible to posit a decreasing returns to scale production function as proposed in Lucas Jr. (1978) and used in Restuccia and Rogerson (2008) and Sandleris and Wright (2010). It is possible then to dispense with the differentiated products setup. The inverse of the returns to scale parameter would play a similar role to $\psi$ in my setting.}

\[
Y_i = A_i L_i^{\gamma_L} M_i^{\gamma_M} K_i^{1-\gamma_L - \gamma_M}
\]

I include materials and energy $M_i$ as a third factor of production for consistency with my empirical work. The plant solves

\[
\max_{L_i, K_i, M_i} p_i Y_i - (1 + \tau_i)(wL_i - rK_i - p_M M_i)
\]

subject to its demand curve. The term $\tau_i$ is a “wedge” that distorts the outcome from the efficient one. One can think of it as an abstract tax that increases the costs of a plant’s inputs. For exposition, I consider the case with only one wedge. It is possible to consider separate wedges on the different factors of production. These different wedges would then distort the plant’s relative choice between, say, capital and labor. The wedge I consider simply affects the scale of operations.

This wedge is taken as given from the point of the view of the plant. HK and others such as Restuccia and Rogerson (2008) think of these wedges as generated by government policies that restrict entry or favor certain enterprises, say. For now, I remain agnostic on where the wedges come from and prefer that they be thought of as purely an accounting device. Eventually I will offer evidence that the wedge is driven in part by working capital interest rates.

HK building on Foster et al. (2008) stress the difference between $TFPR$, which is productivity in terms of revenue and $A$, which is physical productivity (also known as $TFPQ$). The two are linked by the identity

\[
TFPR_i = p_i A_i
\]

With this demand structure, the only reason $TFPR$ should vary across plants is due to wedges.

It is important to understand why a frictionless market should have no variation in $TFPR$. The answer comes from the fact that for a monopolistically competitive plant facing a constant elasticity (CES) demand curve, the elasticity of price with respect to productivity is -1. So plants with higher productivity perfectly offset that productivity advantage by charge lower prices thereby leaving fixed $TFPR$. Note that a particularly simple interpretation of differences in revenue productivity is that
plant’s face different factor prices. In particular, the framework shows that plants facing higher factor prices will have higher TFPR.

2.1 Accounting for Misallocation

I now summarize the effects of misallocation for industry-level productivity. Let $L = \int L_i \, di, K = \int K_i \, di, M = \int M_i \, di$. Then define industry productivity in the usual manner as

$$TFP \equiv \frac{Y}{L^{\gamma_L} M^{\gamma_M} K^{1-\gamma_L-\gamma_M}}$$

Then HK show that

$$TFP = \left[ \int_0^1 \left( A_i \frac{TFPR_i}{TFPR} \right)^{\psi-1} \, di \right]^{\frac{1}{\psi-1}}$$

where $TFPR$ is the geometric average of $TFPR_i$. One thing to note is that fixing the distribution of $A_i$, TFP will be maximized when $TFPR_i = TFPR_j$ for all $i, j$.

To gain some intuition, it is useful to consider a special case with a particular assumption on the joint distribution of $TFPR, TFPR$.

**Proposition 2.1** Assume that $A_i, TFPR_i$ are jointly log normal with means $\log A, \mu_R$, variances $\sigma_A^2, \sigma_R^2$ and correlation $\rho_{QR}$, then

$$\log TFP = \log TFP^* - \frac{\psi}{2} \sigma_R^2 - (\psi - 1) \rho_{QR} \sigma_A \sigma_R$$

Equation 1 provides a straightforward decomposition of changes in productivity into changes in measures of misallocation.\(^{10}\) It identifies separately two key sources of efficiency losses: changes in the variance of $TFPR$ and the correlation between $TFPR$ and $TFPQ$.

The first term $\log TFP^*$ can be thought of as the efficient level of industry TFP taking as given the distribution of plant physical productivities. To build intuition, the first term can further be

\(^{10}\)It is useful to step back and compare this to other decompositions in the literature, e.g. Foster et al. (2008); Petrin et al. (2011). Those works have all written down accounting identities that decompose changes in productivity into broadly speaking within-plant changes and between-plant changes. The main interest there was in quantifying which of these terms dominated. While worthy exercises, those papers were not able to address the question of in what sense these shifts can be thought of as efficient since they all lacked a benchmark model.
written as
\[
\left[ E A_t^\psi - 1 \right]^{\frac{1}{\psi - 1}} = \log A + \frac{\psi - 1}{2} \sigma_A^2
\]

Here log \( A \) captures a “pure” technology shock that affects the mean of the productivity distribution. In reality, log \( A \) can also capture changes in the average plant productivity due to say labor hoarding. It will also capture any “cleansing effects” on the extensive margin of the Depression whereby the least physically productive plants are driven out (Caballero and Hammour, 1994). This selection process will also presumably have effects on the distribution of physical productivity, \( \sigma_A \). My measure of the efficient level of productivity will incorporate all of these changes along the extensive margin.

The term \( \sigma_R \) is a key measure of misallocation acting to decrease by productivity as emphasized by HK.\(^{11}\) Finally, the term \( \rho_{QR} \sigma_A \sigma_R \) is the dimension of misallocation stressed by Restuccia and Rogerson (2008). A higher correlation leads to greater efficiency losses all else equal because more productive plants face higher wedges on average. Finally, note that this focuses on within industry misallocation nor between industry misallocation. This will parallel my empirical work which will just focus on within industry misallocation. With data for a broader set of industries, it is trivial to extend this model to study aggregate productivity.

3 Data

I only offer a very short summary of the data here. A much longer discussion of the source and some checks on the quality of the data can be found in Ziebarth (2013). The schedules come from the Census of Manufactures taken in 1929, 1931, 1933, and 1935. I use information for a select group of industries: manufactured ice, cement, macaroni, sugar refining, bone black,\(^{12}\) automobiles,\(^{13}\) radios, cane sugar, and agricultural implements. This gives a total sample of around 20,000 plants over all four years. While not a random sample of industries, this represents approximately 3% of the total number of schedules over all 4 years. Regarding the “size” of these industries as measured by number of digits in an appropriate SIC, these would be around 3 digits. Some like manufactured

\(^{11}\)Though do note that a 1% increase in \( TFPR \) does not necessarily have the same magnitude of an effect as 1% decrease in \( TFPQ \) since they are multiplied by different coefficients.

\(^{12}\)I thank Miguel Morin for sharing this data with me.

\(^{13}\)These data were originally collected by Bresnahan and Raff (1991) and is available as ICPSR 31761. Miguel Morin helped put these in a usable format.
ice and cement are closer to a 4 digit division.\footnote{There is an additional issue in the radio industry, which is that the Census only “created” this industry in 1931. Before 1931, radio producers were lumped into the broader industry of phonograph and other recording manufactures. See Scott and Ziebarth (2013) for a longer discussion on this issue.}

What is unique about the schedules relative to, say, the Annual Survey of Manufactures is the fact that they have a question about the physical output rather than solely revenue. I will exploit this fact for the cement industry where issues of multi-product production and quality differences are minimized. At the same time, these problems are why I focus simply on revenue productivity for the broad sample. The other limitation is the fact that there is only limited information on the capital stock or investment. Hence, I will focus on labor productivity. Again the cement industry will be an exception to this rule allowing for me to estimate total factor productivity.

4 Empirics

4.1 Results on Dispersion

I estimate revenue productivity for each plant in the following way. First of all, as in Basu and Fernald (1997), I measure revenue as gross output rather than value added. The next question is adjusting this for materials and labor used, in particular, how to estimate the production elasticities. There is of course a large literature on estimating production functions. I employ a cost shares type method to estimate the elasticities. Given the limitations of the data, this is the most attractive approach. For example, there is no information on investment, so it is infeasible to employ the method of Olley and Pakes (1996). The method of Levinsohn and Petrin (2003) makes little sense when estimating a gross output production function. Though a number of caveats apply, variants of this approach have been standard since Solow (1957). The complication relative to the standard case is that I do not observe total costs. So I rely on the fact that costs should be a constant multiple of revenue along with an assumption similar to HK that profits are distributed pro-rata to each of the factors. The assumption about distribution of profits is important. If instead all profits accrued to capital, then my estimates would be biased.

With this assumption, the labor elasticity $\gamma_L$ can be calculated as the fraction of the total wage bill in revenue, similarly for materials, and $\gamma_K = 1 - \gamma_L - \gamma_M$ as the residual (though I do not observe capital). To estimate the industry-specific production elasticities, I average each plant-level
estimate across all plants and years. This estimation method will be consistent with my theoretical specification to the extent that the working capital constraint I will impose affects each factor of production equally. I then calculate my measure of labor productivity $\log TFPR$ as

$$\log TFPR_i = \log p_i y_i - \gamma_L \log L_i - \gamma_M \log M_i$$

where $y_i p_i$ is revenue, $L_i$ is total wage earners, and $M_i$ is cost of total materials and energy used.

The census has an hours worked variable that is unfortunately missing for 1933, which is why I focus on a specification with only number of wage earners as the labor measure. The results are broadly similar if I use the hours worked information and restrict attention to 1929, 1931, and 1935. In addition, they are robust to use $\log \frac{y_i p_i}{L_i}$ as the measure of productivity.\(^{15}\)

One point to note is that $M_i$ is in nominal terms. There does not appear to exist a suitable price deflator for this variable at the industry-level. A similar issue faces Sandleris and Wright (2010) and they choose to not deflate the variable. I make the same assumption so any changes in the price of these inputs will show up as shifts in aggregate productivity. It again is important to emphasize that this will not affect differences in productivity between plants.\(^ {16}\) Besides lacking a deflator for materials inputs, an additional limitation is that I lack information on non-production workers. In 1929, 1933, and 1935, there is information about salaries paid for non-production workers. In principle, this piece of information would allow me to impute a labor input variable that included both types of workers along the lines of Baily et al. (1992) or Davis and Haltiwanger (1990). The problem is that I would have to drop 1931 since I have no information on salaries in that year. However, comparing the ratio of total salaries to total wages over the three years with data, I find little change in the mean ratio. Therefore, I simply assume that total workers (wage earners and salaried), which is equal to my labor input measure, are a constant multiple of total wage earners. I also inflate the total wage bill by the same multiple when I calculate production elasticities.

Finally, to reduce measurement error, I trim the 2% tails of the distribution for $TFPR, TFPQ$ \(^ {15}\) In results not reported here, I also consider a value added specification and using some information on hours worked for years other than 1933. Results are relatively unaffected.

\(^ {16}\) It will be useful to keep the possible scale of this bias in the overall productivity trends in mind. For reference, the overall price level falls by around 15% from 1929 to 1931 and then by another 7% from 1931 to 1935. Given an estimated materials share of around 1/2 and assuming that the prices of inputs track the general price index, then the growth of both $TFPQ$ and $TFPR$ will be overestimated by 7.5% from 1929 to 1931 and by an additional 3.5% from 1931 to 1935.
then recalculate production elasticities and plant productivities. This is a slightly broader cut than
the 1% trim than HK. The related results reported in Oberfield (2011) are very sensitive to how
the tails are handled. So I will offer the results from some other choices as well. Finally, I eliminate
any plants that report less than 50% of their revenue from the industry’s primary product.

I do not plot the raw distribution of $\log TFPR_i$ instead I demean by the industry-year average.
This ensures that changes in dispersion are not driven by differences between industries. After this
scaling, I choose to plot the dispersion across all plants with the results reported in Figure 3. \footnote{Kehrig (2011) first takes the variance industry by and industry and then reports the \textit{median} across industries. I do not favor this approach since changes are driven not just by changes in the industry but also by \textit{which particular industry} is the median. Still results are fairly similar if I use this median (or mean) approach.}
I find across the three choices for how much of the tails to trim that dispersion increases around
7.5% from 1929 to 1931, stays elevated through 1933 before falling back towards the 1929 value. I
can reject the null that the variance in 1929 is the same as that of 1931 with a $p$ value of less .01.
Note that this increase in dispersion is coming from a fanning out of both tails of the distribution.
Both the 10th and 90th percentiles are increasing in magnitude.

Compare this to Kehrig (2011) who finds most of the increase in dispersion is due to changes
in the composition of the least productive plants. This leads him to focus on entry and exit of
the least productive plants as the key mechanism. The fact that both percentiles are affected in
my setting suggests that this is not just entry and exit. Second comparing the magnitudes of the
effects relative to the evidence for the modern business cycle, I find an increase of only about half
as large as for the average peak-trough in the post-war economy. This may be due to the fact that
the sample here is skewed towards non-durables, in particular, ice. Kehrig (2011) found that this
type of industry showed less cyclical sensitivity.

4.2 A More Detailed Analysis for the Cement Industry

I consider a more detailed analysis for on particular industry: cement.\footnote{An earlier version of this paper also had a detailed analysis of the ice industry, which shares many of these useful features. For clarity, I have chosen here to just focus on cement since its productivity patterns most closely match the aggregate pattern.} By focusing on this one,
I can consider the robustness of simplifying assumptions made in the main body of the paper.
First of all, I chose in the text to focus on labor productivity due to data limitations. However,
for cement, I have reasonable measures of physical capital. Here capital is measured in “real”
units of production capacities rather than a dollar value. In particular, production capacities were transcribed from exhibits that formed a major part of an anti-trust case against the industry in 1941 (Cement Institute v. U.S.).\textsuperscript{19}

I now estimate both physical and revenue productivity using all three factors of production

\[
\log TFP_{i} = \log p_{i} y_{i} - \gamma_{L} \log L_{i} - \gamma_{M} \log M_{i} - (1 - \gamma_{L} - \gamma_{M}) \log K_{i}
\]

\[
\log TFPQ_{i} = \log y_{i} - \gamma_{L} \log L_{i} - \gamma_{M} \log M_{i} - (1 - \gamma_{L} - \gamma_{M}) \log K_{i}
\]

where I now observe capital \(K_{i}\) and \(y_{i}\) output directly. To quantify the effects of misallocation, define

\[
\log \hat{TFP}_{Q_{i}} = \frac{1}{N} \sum_{i} \log TFPQ_{it}
\]

\[
\log \hat{TFP}_{R_{i}} = \frac{1}{N} \sum_{i} \log TFP_{R_{it}}
\]

where \(TFPQ_{it}\) is inferred from \(TFPR_{it}\). Then \(\log \hat{TFP}_{Q_{i}}\) will be my estimate of the \(\log A_{t}\) term. Let \(\hat{TFP}_{R_{it}}, \hat{TFP}_{Q_{it}}\) denote the deviations of the log plant value from the industry-year log average. The standard deviation of these deviations are my estimates of \(\sigma_{At}, \sigma_{Rt}\). I also calculate the correlation between the deviations for a given year and industry, \(\rho_{QRt}\).\textsuperscript{20}

My estimate of the efficient level of productivity is given by

\[
\log TFP^{*} = \frac{1}{\psi - 1} \log \left( \frac{1}{N} \sum_{i} TFPQ_{it}^{\psi - 1} \right)
\]

Note I do not impose log normality in calculating the role of misallocation. To calculate the efficient level of productivity, I need to set a value for the elasticity of demand \(\psi\). For “illustrative” purposes, I will choose a value of \(\psi = 4\) in line with HK. I will also show how the fraction of the productivity declines explained by misallocation vary as \(\psi\) changes over a reasonable range of values from the literature.

\textsuperscript{19}Stephen Karlson transcribed these capacities from the originals and graciously shared the data.

\textsuperscript{20}In Ziebarth (2013), I show that these deviations from industry averages are correlated with exit decisions suggesting that these differences are economically meaningful and not simply measurement error. I also show that they display similar levels of autocorrelation to modern estimates also suggesting at least no greater degree of measurement error relative to modern data. Thirdly, I provide evidence that whatever the magnitude of (classical) measurement error is, it remains constant over the 4 censuses.
Before turning to the results, it is important to emphasize that in reporting the role of misallocation in productivity changes, I am deliberately not accounting for labor hoarding and changes in the composition of the labor force that can bias productivity measurements. I am solely interested in the difference between changes in the overall industry productivity and the efficient level. Changes in the efficient level, at least due to shifts in the mean, may be due to the usual explanations given for procyclical productivity. I take no particular stand on the quantitative role of those factors except to assume that they do not explain the total decline in productivity leaving a fraction to be explained. Of course, a full resolution of the productivity puzzle during the Depression will involve both of these channels.

Besides not accounting for these unobservable shifts, I will intentionally exaggerate the decline in productivity because of data limitations. Specifically, a major part of explaining the sharp mean productivity decline in my data is that my productivity estimates do not include hours worked, which fall sharply over the course of the Depression. However, the extent to which the declines in hours worked are similar across plants, this omission should not bias the results focusing on cross-sectional differences and, if anything, bias downwards the contribution of misallocation.

It also is useful at this point to address some possible worries, in particular, the role played the CES demand structure with the implication that the markup is constant across time and plants. What if, for example, the true pricing model was more “exotic” such as a limit pricing model where the most productive plant charges a price just below what the second most productive plant charges? The problem is that they use the demand structure to infer physical productivity of each individual plant since they do not observe physical output. Instead with an assumed value of $\psi$, they invert revenue to back out physical output and then measure $TFPQ$. It is clear that in the limiting pricing example, the method HK use could go astray. If the most productive plant is charging a price that only depends the productivity on the next most productive plant, then the inferred level of $TFPQ$ from $TFPR$ has almost nothing to do with its own productivity and everything to do with its nearest competitor. This would not affect my results for the cement industry since I use direct information on quantity produced without having to rely on the demand structure.
4.3 Shifts in Misallocation Measures

First, I document the sources of misallocation in this industry: dispersion in $TPFR$ and the correlation between $TFPR$ and $TFPQ$. To reemphasize, these results do not depend on a particular choice of $\psi$. In Figure 4, I display the statistics for cement. From 1929 to 1931, the standard deviation of $TFPR$ increases sharply and then increases a bit more between 1931 and 1933 before starting to fall back in 1935. This roughly matches the pattern for the broader sample. What is interesting is that $\rho_{RQ}$ does not follow exactly the same pattern. It increases only slightly between 1929 and 1931 and much of the overall increase is concentrated between 1933 and 1935. These results closely match the dynamics in the interbank market, which will play a central role in the theory that I develop. Though even at the end of the sample, both of these measures of misallocation are substantially elevated from the pre-Depression level while the dispersion in physical productivity is back to its original value suggesting some persistence in the initial shock.

As an aside, it is interesting to note that the pattern in $TFPR$ dispersion is not the same for dispersion in $TFPQ$. Modern work by Kehrig (2011) and Bloom et al. (2011) has found strongly countercyclical revenue productivity dispersion, but these authors have had to remain silent about what that meant for dispersion in physical productivity. In results not reported here, dispersion in $TFPQ$ shows much less cyclical variation than dispersion $TFPR$. These differences matter for how changes in dispersion measures potentially drive productivity over the business cycle.

4.4 Decomposing Productivity in the Cement Industry

Figure 5 displays for the cement industry three productivity measures: industry, efficient, and mean plant-productivity. Almost the complete fall in productivity between 1929 and 1931 is explained by misallocation with very little movement in the efficient level though the mean plant-level productivity does fall. This suggests that there are important shifts in the dispersion of physical productivity offsetting the decline in plant-level productivity. Figure 6 shows how the fraction of the observed productivity decline can be explained by increases in misallocation varies with $\psi$. Note that whether or not this fraction is increasing or not depends on the relative sizes of $\sigma_A, \sigma_R, \rho_{QR}$. The reason is that $\psi$ not only changes the losses from misallocation but also the efficient level of productivity. Higher values of $\psi$ make misallocation more costly all else equal, but they also make
dispersion in $\sigma_A$ more valuable for industry productivity.

Changes in misallocation play a substantial role for all the years over a wide range of values for $\psi$. As noted above, depending on the value of $\psi$ for 1931, misallocation can potentially explain the total decline in industry productivity. At a minimum, misallocation explains 50% of the decline. The fraction of decline between 1929 and 1933 averages around 20 to 30%, an economically meaningful. The smallest fraction is consistently in 1935 where this fraction hovers around 15% over the range of $\psi$ considered. For some perspective, it is useful to compare there results to those in Sandleris and Wright (2010). They study the 2001 crisis in Argentina and attempt a similar accounting exercise. They quantify the loss on TFP from greater misallocation across and between sectors at 10% of the total decline. My results for these industries suggest much higher losses from misallocation during the Great Depression.

5 A Model of Financial Frictions and Productivity

“As the crisis of 1929-33 ha[s] proved, the [interbank] system has contributed to intensify cyclical maladjustments.” Palyi (1939)

5.1 Empirical Motivation

The obvious next question is what drives these changes in misallocation. Compare the U.S. experience during the Great Depression to Canada’s. Between 1929 and 1933, Canadian GNP falls by around 40% relative to a 2% trend.\textsuperscript{21} while the U.S. experiences a nearly identical decline of around 38% over the same time. Comparing TFP changes, Canada’s aggregate productivity falls by around 15% while U.S. productivity falls by 20% relative to trend. (See Figure 8.) If not in the context of the Depression, this difference of 5 percentage points would seem quite large. But the differences between the two are much starker if one considers productivity excluding the government and agricultural sectors. With this measure, the decline in U.S. TFP is nearly double that of Canada’s between 1929 and 1933. What is striking is the divergence in productivity after 1933. While U.S. aggregate and private non-agricultural productivity begins to recover in 1935, Canadian TFP stubbornly remains well below trend. So what explains this difference in the behavior or

\textsuperscript{21}All numbers are from Amaral and MacGee (2002).
productivity? One of the most striking differences is the fact that Canada did not suffer a banking crisis like the U.S. No Canadian banks failed during the Great Depression.\textsuperscript{22}

This comparison is only by way of motivation. I go further in offering two related empirical claims to motivate my model based on financial frictions. First, there is some evidence that the dynamics of the dispersion of interest rates track those of misallocation.\textsuperscript{23} Figure 7 shows a sharp increase in the dispersion of lending rates measured both by the standard deviation and interquartile range. These changes are not driven by outliers in 1933 as I have trimmed the high and low observations for each year. This pattern is also present with the higher frequency though less geographically representative data in Smiley (1981). It is also important to emphasize that this is only one dimension of misallocation with the correlation between $TFPR$ and $TFPQ$ serving as another source. This correlation in the plant-level data is not captured in this figure. My model will offer a natural explanation for why the correlation and dispersion should be linked.

The second observation motivating my model is the collapse of the interbank funding market driven by cascading failures through these correspondent networks. Before discussing this particular claim, it is useful to step back and discuss the U.S. banking market as a whole at this time. In particular, this is period when branching was severely restricted. Some authors such as Bodenhorn (1995) have argued that the U.S. still did not have a fully integrated capital market by the Depression. There was no interstate branching and in the majority of states, banks could not open up other branches even within the state. The lack of branching made banks very dependent on functioning interbank markets to fund liquidity shortfalls. With branching, resources can move at least relatively freely between branches of the same bank. Without branching, banks have to rely on interbank markets for liquidity. This led to a situation where geographically-concentrated agricultural shocks had a tendency to weaken financial conditions in that region with a drought or flood setting off a spiral of declining income and bank capital (Landon-Lane et al., 2011).

Of course, each region was not completely on its own especially after the creation of the Federal Reserve system, which allowed for the rediscounting of paper by eligible banks. Now by 1929, there were still many banks that were not a part of the Federal Reserve system and, hence, had no direct access to the discount window. Correspondent networks filled this gaping hole by discounting paper

\textsuperscript{22}The economy did suffer a debt crisis brought on by an unexpected deflation (Amaral and MacGee, 2002).
\textsuperscript{23}An earlier version of the paper found a positive correlation between state-level lending rates and $TFPR$ for the cement and ice industries consistent with accounting framework.
provided by banks who were not members of the Federal Reserve system. So-called correspondent banks then turned around and rediscounted this paper at the relevant regional Fed. Writing at the time, Palyi (1939) aptly summarized the tight link between unit banking and the correspondent networks.

Unit banking as it exists in this country cannot be maintained unless an institutional setup is provided to keep the units in close contact with the money market...Unless we develop a similar system, which has never existed in this country, we have to rely on correspondent relationships as the only other alternative to permit the survival of unit banking, which would soon deteriorate without contact with the “fresh air” of the open money market.

This system of correspondents allowed small local banks the opportunity to tap a broader set of funds. The ability to tap this broader pool was crucial in mitigating differences in borrowing rates across regions.

However, Richardson (2007) highlights the fragility of this system. Correspondent banks would credit checks upon receipt, and respondent banks would do so as soon as the check was dropped in the mail but before the funds were ever drawn from the check-writing bank. This was fragility was laid bare during the Great Depression. Using archival evidence, Richardson (2007) showed that a large fraction of bank suspensions during 1931 were due to failures of correspondent banks. These failures in the correspondent networks translate into declines in gross flows in the interbank market. This is despite the fact that demand for these short-term funds was very high as the country was swept by panicked depositors looking for their cash. In 1933, the total amount of interbank deposits for all Federal Reserve member banks reached a low point of $3.4 billion in June of 1933, a decline of 23% from its peak. After the nadir, interbank deposits rapidly regained their pre-Depression highs by the end of 1934. These three observations regarding the link between TFPR and lending rates, the dispersion of lending rates, and the collapse of interbank markets highlight the chain of reasoning that I will attempt to capture in my theoretical model.

5.2 Plants

I now present a simple static model to explain the losses in productivity due to misallocation stemming from financial frictions exacerbated by the banking collapse. Before turning to the model, it is useful to distinguish my setup from existing literature. Like Gilchrist and Zakrajsek
(2012), I emphasize dispersion in borrowing costs as a main driver of misallocation. Unlike those authors, I do not focus on an uncertainty shock as the underlying driver of this misallocation. Similar to Khan and Thomas (2013), I study a model with misallocation on the intensive margin. Unlike them, I have no real frictions in the form of investment irreversibilities. Finally as in Buera and Moll (2013), I have an extensive margin of entry and exit. Unlike those authors, I have no dynamic savings problem for the plants.

There is a continuum of geographically separated regions (“islands”) and on each island, there is a continuum of plants. On island \( i \), all plants are the same and produce variety \( i \). Final output is a Dixit-Stiglitz aggregator of all the types

\[ Y = \left[ \int_0^1 y_i \frac{\psi - 1}{\psi} \, di \right]^{\psi - 1} \]

with \( \psi > 1 \). For now, I will take final demand \( Y \) as exogenous. Plants operate a linear production technology in labor alone

\[ y_i = A_i L_i \]

It would not be difficult to extend the analysis to a constant returns to scale production function that included other factors of production as in the empirical analysis.

At the beginning of the period, each island receives a productivity draw \( A_i \) from a log normal distribution with variance \( \sigma_A^2 \) and mean \( \mu_A - \frac{1}{2} \sigma_A^2 \) where \( \mu_A \) is an aggregate technology shock. When \( \mu_A = 0 \), then the mean of \( A_i \) equals 1. Shocks are i.i.d. across islands and independent over time as well. This assumption eliminates the need to keep track of the history of shocks across islands.

Plants have a working capital constraint that requires them to borrow up front to pay their wage bill, e.g., Barth and Ramey (2001). Working capital constraints have also been emphasized by the literature examining sudden stops in emerging markets as a potential key source of amplification (Neumeyer and Perri, 2005). Plants can fund that wage bill by borrowing from only banks on their

\footnote{There is a slight complication here. Because I will require a plant to obtain credit from a bank on its own island, to insure that price taking of interest rates is a reasonable assumption, I need many plants on a given island. The issue with this assumption is that it then makes the assumption of monopolistic competition suspect. To justify this assumption rigorously, I would have to add another layer of product differentiation within a given island. To ease exposition, I assume this issue away.}
own island at gross rate $r_{fi}$. Hence, they maximize

$$p_i y_i - r_{fi} w L_i$$

where they face demand curve derived from the profit maximization problem of the final goods producer,

$$\frac{p_i}{P} = \left( \frac{y_i}{Y} \right)^{-\frac{1}{\psi}}$$

and the production function $y_i = A_i L_i$. Plants make their labor choice after the productivity shocks have been realized. This assumption about the local nature of plant finance is very natural and in part motivated Bernanke (1983b) to argue for the presence of non-monetary effects of bank failures. Banks in his view embodied information about the quality of various local projects.

The optimal choice of labor solves

$$r_{fi} w = Y^{\frac{1}{\psi}} \left( \frac{\psi - 1}{\psi} \right) A_i^{\psi - 1} L_i^{-\frac{1}{\psi}}$$

where I have normalized the final price index $P$ to 1 without loss of generality. Dividing the optimal choices for firms on island $i, j$ yields an expression useful for later

$$\frac{r_{fi}}{r_{fj}} = \left( \frac{A_i}{A_j} \right)^{\frac{\psi - 1}{\psi}} \left( \frac{L_i}{L_j} \right)^{-\frac{1}{\psi}}$$

which shows that the ratio of labor use across islands is related to the ratio of both productivities and interest rates across islands. Now I calculate $TFPR_i$ as

$$TFPR_i \equiv p_i A_i = A_i \left( \frac{y_i}{Y} \right)^{-\frac{1}{\psi}} = \frac{w r_{fi}}{Y^{\frac{1}{\psi}} \left( \frac{\psi - 1}{\psi} \right)}$$

An interesting extension would be to allow multiple factors of production with differences in the pledgability of the different inputs. In particular, one might think that it is easier to use materials purchased with a loan as collateral than labor inputs. In this case, there might be a wedge between the rates a plant can borrow for materials versus labor. Shocks to the financial shock may then have asymmetric effects on the different interest rates and generate different dynamics for labor versus materials wedges.
This implies that ratio of $TFPR$ for any pair of islands $i,j$ is given by

$$\frac{TFPR_i}{TFPR_j} = \frac{r_{fi}}{r_{fj}}$$

Plants on islands with higher interest rates have higher levels of $TFPR$ reflecting the fact that they are constrained from fully expanding output to the efficient scale though for now, certain firms are not simply excluded from the credit market altogether. In fact,

$$\sigma_R = \frac{w}{Y^{\frac{1}{\sigma}} (2^{1-\frac{1}{\sigma}})} \sigma_{rf}$$

where $\sigma_{rf}$ denotes the standard deviation of the log of gross working capital interest rates. Therefore, any shock that increases the interest rate spread on working capital loans will increase the dispersion of $TFPR$. This is not a surprising finding as $r_{fi}$ is a direct analog of $1 + \tau_i$ in the accounting framework. I now turn to the banking sector to close the model and determine the working capital interest rates.

### 5.3 Banks

There is a continuum of banks on each island. Abstracting from differences across the island, each bank starts out with net worth $N$. One could think of this just as well as “lending capacity” to include both capital and deposits. Again given the i.i.d. structure of productivity shocks, there is no reason for any particular island to build up excess capital. To finance working capital loans $H_i$, banks use their net worth and interbank loans $b_i$ ($b_i \geq 0$ implies a bank is borrowing funds). The assumption that commercial banks are mainly in the business of short term loans is in line with much of the evidence from the time. White (2000) notes that commercial banks by “law and tradition” specialized in short term loans to fund materials and labor outlays. This was distinct from the role played by investment banks in facilitating long-term investment.

So the flow of funds constraint is given by

$$H_i = b_i + N$$

As noted above, banks can only fund projects on their own island and can only obtain additional
funds from other banks through the interbank market at $r_b$. This assumption of a quasi-fixed supply of funds is a common one in the literature. To solve the model in closed form, I will assume a particular form for interbank borrowing

$$\frac{b_i}{N} = \frac{r_{fi} - r_b}{\omega}$$

where $\omega$ is the marginal cost of leverage and indexed the degree of frictions in interbank lending. If $\omega \to 0$, then $r_{fi} \to r_b$ for all $i$ and there is no dispersion in $TFPR$ across islands. In this limiting case, banks can take advantage of any arbitrage opportunity by leveraging up with no cost. Hence, there cannot be any interest rate differentials across islands. One could could think about this borrowing rule as the optimal choice of bank that faces a quadratic cost of additional interbank borrowing or as a linear approximation to a more general cost function.

There are many ways to motivate a limit on leverage a bank and this sort of borrowing rule. For example, banks could face an incentive problem where the market limits their leverage because of fears that the bank will abscond with assets (Gertler and Kiyotaki, 2010). Limitations on leverage can also be motivated by problems due to costly state verification (Bernanke et al., 1999) or margin value-at-risk requirements (Brunnermeier and Pedersen, 2009). I take no stand on what the proper micro-foundation is. What is different about my assumption is that a particular bank’s leverage is a smooth increasing function of the differential between the rate it can charge firms on its island and its borrowing costs from the interbank market. Compare this to the case of Buera and Moll (2013) in which they study a hard constraint on leverage.

I will approximate the solution around a particular level of productivity $\bar{A}$ such that if $A_i \leq \bar{A}$, banks on those island will be net interbank lenders and vice versa. For $A_i = \bar{A}$, then it must be the case that

$$r_{fi} = r_b$$

Setting $b_i = N - wL_i$ to 0 pins down the productivity cutoff $\bar{A}$

$$N = w\bar{L} = w^{1-\psi}L\bar{A}^{\psi-1}r_b^{-\psi}$$

where $\bar{L}$ is labor demand for this type of plant. Relative to Khan and Thomas (2013) or Gilchrist
et al. (2010), this is a much less elaborate financing problem. Most importantly, there are no
dynamic components to the problem since all borrowing is 1 period debt and there is no (explicit)
possibility of default, strategic or otherwise. This is clearly a large cost but the advantage is that
it is trivial to add an extensive margin of entry and exit.

5.4 Solving the Model

Loan market clearing implies that the total amount of working capital loans on a given island has
to be equal to the amount of bank net worth on the island plus interbank borrowing.

\[ wL_i = N + b_i \Rightarrow wL_i = N \left(1 + \frac{r_{fi} - r_b}{\omega}\right) \]  

(4)

This relationship delivers a positive relationship between plant borrowing rates and total wage bill
taking everything else as given. Ceteris paribus more productive firms pay higher interest rates, a
slightly counterintuitive result. If the quality of plants were with regards to the risks of bankruptcy,
then intuition would suggest that the lower quality plants with a higher probability of failure would
face higher working capital interest rates to compensate banks for the risk. However, in this model
as in Gertler and Kiyotaki (2010) with a quasi-fixed supply of loans and no internal supply of funds,
more productive plants have a higher demand for loans, and, hence they must pay higher interest
rates.

Interbank clearing requires that

\[ \int_0^\infty b_i dF(\log A_i) = 0 \]

Subbing for the expression for optimal interbank lending,

\[ \int_0^\infty N \left(\frac{r_{fi} - r_b}{\omega}\right) dF(\log A_i) = 0 \]  

(5)

where \( F \) is the cdf of the normal distribution function with mean \( \mu_A - \frac{1}{2} \sigma_A^2 \) and standard deviation
\( \sigma_A \). Then I have

**Proposition 5.1** \( \bar{r}_f = r_b \) where \( \bar{r}_f = \int_0^\infty r_{fi} dF(\log A_i) \) is the average working capital interest rate.
It is useful to note here that this model will generate a distribution of leverage ratios across banks though the overall banking sector will not be leveraged. The lack of overall leverage follows from the aggregate loan market clearing constraint. Again this assumption can be relaxed in a richer model where banks can also raise funds through the retail deposit market before the productivity shocks are realized.

Equation 4 will be the key equation since it pins down the relationship between $r_{fi}$ and $A_i$. Using the labor demand expression to solve for $wL_i$,

$$w^{1-\psi}LA_i^{\psi-1}r_{fi}^{-\psi} = N \left( 1 + \frac{r_{fi} - r_b}{\omega} \right)$$

Denote the implicit relationship between $r_{fi}$ and $A_i$ by $r_{fi} = g(A_i)$. It is useful to note that $g$ is a strictly increasing function as demand for loans is strictly increased in $A_i$. Differentiating with respect to $A_i$, I have

$$\frac{\partial \log r_{fi}}{\partial \log A_i} = \frac{\psi - 1}{\psi + \frac{w^{\psi-1}N}{L\omega}A_i^{1-\psi}r_{fi}^{\psi+1}}$$ (6)

Therefore,

**Proposition 5.2** To a first order approximation about $\log A_i = \log \bar{A}$, $\log R_{fi}$ will also be normally distributed with mean $\log R_b$ and variance

$$\sigma_R = \sigma_A \cdot \rho_{QR}$$

This first-order approximation also implies that

$$\rho_{QR} = \left. \frac{\partial \log r_{fi}}{\partial \log A_i} \right|_{\log A_i = \log \bar{A}} = \frac{\psi - 1}{\frac{N\psi}{\omega} + \psi}$$

The expression shows that as frictions in the interbank market go to 0 i.e. $\omega \to 0$, then the standard deviation of $TFPR$ goes to 0 along with the correlation between the $TFPR$ and $TFPQ$.

The model here is slightly counterfactual as it draws a very tight one-to-one link between $\sigma_R$ and $\rho_{QR}$ though it is clear from the data that the relationship is not nearly so simple. This tight link depends on the fact that $\sigma_A$ is taken as exogenous. Any (normal) log-linear approximation will have this problem as both $\rho_{QR}$ and $\sigma_R$ are linear functions of the derivative of relationship between
$R_{fi}$ and $A_i$. Also, my approximation ignores non-linear behavior in the tails of the productivity distribution brought on by the fact that leverage scales with the difference between $r_{fi}$ and $r_b$ rather than the ratio of $r_{fi}$ to $r_b$.

To solve the model, note that aggregating loan market clearing across all islands along with interbank clearing implies $w = \frac{N}{L}$ and labor clearing will be trivially satisfied since $L$ is exogenous. Then I simply need to solve for $r_b, \bar{A}$ to fully characterize the equilibrium. Under my approximation, I have that

$$\log r_{fi} = \log r_b + \rho_{QR}(\log A_i - \log \bar{A})$$

Recall that

$$\tilde{r}_f = \int_{0}^{\infty} r_{fi}d\Phi(\log A_i) = r_b \quad (7)$$

Substituting my expression for $\log r_{fi}$ into equation 7, I find that

**Proposition 5.3** The values of $\bar{A}, r_b$ are determined by the intersection of the following two curves

$$\log \bar{A} = \frac{1}{2} \frac{\sigma_A^2(\psi - 1)}{\omega} (8)$$

$$\log \bar{A} = \frac{\psi}{\psi - 1} \log r_b N - \frac{\psi}{\psi - 1} \log L \quad (9)$$

where the second equation is a rewritten expression for the cutoff $\bar{A}$.²⁶

These equations characterize equilibrium in the interbank market. Equation 8 determines a negative relationship between $r_b$ and $\bar{A}$. An increase in $r_b$ implies a higher supply of interbank loans from banks with marginal projects. This leads to a lower productivity threshold as marginal firms are able to take advantage of the increased supply. Equation 9 on the other hand determines a positive relationship between the two. When $r_b$ increases, this increases the threshold productivity since only relatively more productive firms are able to compensate banks for foregoing higher returns in the interbank market. Note that both equations only depend on the product of the interbank rate and bank capital, $r_b N$. Therefore,

**Proposition 5.4** Shocks to $N$ do not affect the distribution of TFPR nor TFP.

²⁶Note that the cutoff $\bar{A}$ is a direct function of $\rho_{QR}$, $\log \bar{A} = \frac{1}{2} \sigma_A^2 \rho_{QR}$. 25
Decreases in \( N \) are completely compensated by increases in \( r_b \) and vice versa. It is also clear from this that a shift in the aggregate leverage ratio of the financial sector does not play a role. In fact for my setup, this ratio is fixed to a value of 1.

Under the linear approximation, calculating the effects on productivity is trivial since as before I have

\[
\log TFP = \log TFP^* - \frac{\psi}{2} \sigma_R^2 - \rho_{QR}(\psi - 1)\sigma_A \sigma_R
\]

It is clear that a shock to \( \omega \) only affects equation 8 shifting it rightwards. This leads to an increase in \( \rho_{RQ} \) and, thereby, an increase in \( \sigma_R \) and a loss in productivity. Hence,

**Proposition 5.5** Increases in the marginal cost of leverage \( \omega \) increase both the variation in \( TFPR \), \( \sigma_R \), and the correlation \( \rho_{QR} \) thereby reducing \( \log TFP \).

What these results have shown is that higher marginal costs of leverage can drive declines in productivity. What matters here is the differential effects of these financial frictions across different types of plants rather than effects common to all.\(^{27}\)

### 5.5 Changes on the Extensive Margin

I now show how to endogenize movements in the first two moments of the physical productivity distribution, \( \log A, \sigma_Q \) as well as changes in \( \sigma_R, \rho_{QR} \). It is important to distinguish the endogenous variables \( \log A, \sigma_Q \), which are observed, from \( \mu_A, \sigma_A \), which are exogenous unobserved parameters. In the model with only an intensive margin, in effect, \( \sigma_Q = \sigma_A \). I assume that if firms are sufficiently unproductive, they are excluded from production by banks who prefer to lend into the interbank market. Previously, my assumption on the bank’s lending policy implied that even if \( r_{fi} < r_b \), banks would still lend some fraction of their resources to firms on their island. Now if \( r_{fi} < r_b \), then \( b_i = -N \) as in Buera and Moll (2013). This generates an extensive margin of entry and exit.

\(^{27}\)One thing interesting to note about the model is that it naturally generates procyclical productivity independent of shifts in the marginal costs of leverage. Recall that I took final demand \( Y \) to be exogenous and that I normalized total labor supply to

\[
L = Y \left( \frac{\psi}{\psi - 1} \right)^{\psi}
\]

Note that \( \log \bar{A} \) depends negatively on \( L \). Therefore, a demand shock that increases \( Y \) drives down \( \log \bar{A} \) and with it, \( \rho_{QR} \) and \( \sigma_R \). This will increase overall productivity. Similarly if I had included an aggregate productivity shock, this would enter positively into \( L \) and hence, supply shocks would be multiplied endogenously in this framework. So even abstracting from business cycles driven by financial factors such as changed in \( \omega \), my model delivers procyclical labor productivity. This is due to a cleansing effect in that demand is higher for the fixed supply of working capital.
Hence, firms with productivity $A_i < \bar{A}$ do not produce because they have no resources to hire labor. There is an implicit relationship between $r_b, \bar{A}$ given by

$$Nr_b = \bar{A}^{\psi-1}L$$

(10)

where I have used the fact that $w = \frac{N}{L}$ assuming banks do not collect deposits.

Now interbank clearing implies that

$$\frac{\tilde{r}_f - r_b}{\omega - r_b} = \Phi \left( \frac{\psi - 1}{\psi} \log \bar{A} \right)$$

(11)

where $\tilde{r}_f = \int_{A}^{\infty} r_{fi}dF(\log A_i)$ as before. The difference now is that there is a wedge between $\tilde{r}_f$ and $r_b$ related to the fact that there is a supply of capital from banks on bad islands. For firms with productivity $A_i \geq \bar{A}$, loan market clearing implies that

$$w^{1-\psi}L \bar{A}^{\psi-1}r_{fi}^{-\psi} = N \left( 1 + \frac{r_{fi} - r_b}{\omega} \right)$$

I will approximate the solution to this expression like before with a log linear relationship between $A_i, r_{fi}$.

Rather than approximating around $\bar{A}$ as before, I approximate around $\hat{A}$ implicitly defined as

$$w^{1-\psi}L \hat{A}^{\psi-1}r_{fi}^{-\psi} = N \left( 1 + \frac{\tilde{r}_f - r_b}{\omega} \right)$$

(12)

This is the level of productivity for which the interest rate charged on the island is equal to the average interest rate across islands. As before, I can now write under this approximation,

$$r_{fi} = \tilde{r}_f \left( \frac{A_i}{\bar{A}} \right)^{\rho_{QR}}$$

where $\rho_{QR}$ is the derivative of the loan market clearing expression evaluated at $\hat{A}$

$$\rho_{QR} = \frac{\psi - 1}{\psi + \frac{A^{1-\psi}(\bar{N}\tilde{r}_f)^{\psi+1}}{L^{\psi+1}\omega}}$$
Substituting this expression into the definition of $\hat{r}_f$, I have

$$\hat{A} = \left[ \int_{A}^{\infty} A_i^{\rho_{QR}} d\Phi(\log A_i) \right]^{1/\rho_{QR}}$$

(13)

I now have unknowns $\hat{R}_f, \hat{A}, \bar{A}, r_b$ and equations 13, 12, 11, and 10 as well as the definition of $\rho_{QR}$.

For illustrative purposes, I will consider the case where $\sigma_A = 1$ and $N = L$. Then the system of equations defining the solution is given by

$$\frac{\hat{r}_f - r_b}{\omega - r_b} = \Phi \left( \frac{\psi - 1}{\psi} \log \hat{A} \right)$$

$$1 + \frac{\hat{r}_f - r_b}{\omega} = \hat{A}^{\psi-1} \hat{r}_f^\psi$$

$$\log \hat{A} = \frac{1}{2} \rho_{QR} + \frac{1}{\rho_{QR}} \log \Phi \left( \rho_{QR} - \frac{\psi - 1}{\psi} \log r_b \right)$$

$$\rho_{QR} = \frac{\psi - 1}{\psi + \hat{A}^{1-\psi} \phi_{f}^{\psi+1}}$$

$$r_b = \hat{A}^{\psi-1} \phi_{f}$$

In contrast to the basic case, these equations cannot be written solely in terms of $Nr_b$ and $N\hat{r}_f$.

This implies that shocks to $N$ do have effects on productivity. The key difference here is that capital levels affect the extensive margin and changes to that margin are not neutral with regards to misallocation.

What is interesting under this specification is that there are endogenous movements not only in $\rho_{QR}, \sigma_R$ but also $\sigma_Q$ and $\log A$. In particular, using the formulas for the first and second moments of truncated normals, it will now be the case when $\mu_A = \frac{1}{2}$ and $\sigma_A = 1^{28}$ that

$$\log A = E[\log A_i | A_i > \bar{A}] = \frac{\phi(\log \hat{A})}{1 - \Phi(\log \hat{A})}$$

$$\sigma_Q^2 = Var(\log A_i | A_i > \bar{A}) = \left[ 1 - \left( \frac{\phi(\log \hat{A})}{1 - \Phi(\log \hat{A})} \right)^2 \right]$$

These formulas are very similar to those derived in Buera and Moll (2013). It will still be that $\sigma_R \propto \sigma_{rf}$. However, unlike the case of no intensive margin where both $TFPR$ and $TFPQ$ were log

---

28Recall that this implies that the mean of the distribution for $\log A_i$ is 0.
normally distributed and, hence, \( \log TFP \) had a closed form, in the case of an extensive margin, \( TFPR \) will follow a truncated normal and no closed form for \( \log TFP \) exists.

The effects of a shock \( \omega \) on \( \sigma_R, \rho_{QR} \) are the same as the model with no extensive margin conditional on a fixed \( \bar{A} \). However, the cutoff \( \bar{A} \) falls with \( \omega \) with knock-on effects for \( \sigma_A \) and \( \log A \). The key is that \( r_b \) is decreasing in \( \omega \). This reflects decreased demand for funds through the interbank market. This leads banks with marginal projects to lend to the firms on their island instead of lending to the interbank market. Business cycles driven by this type of shock would not experience “cleansing” during recessions (Caballero and Hammour, 1994). Instead similar to Kehrig (2011), there would be a “sullying” effect of recessions as decreased competition for a fixed factor allows marginal plants to remain in the market.

5.6 Simulating the Effects of \( \omega \)

In this section, I calibrate the model and consider the comparative statics for \( \omega \) to examine this issue. The model is very tightly parameterized especially after making the simplifying assumptions about the underlying distribution of productivities. The result is that I only need to pick values for \( \psi, \omega, N \). I fix \( \psi = 4 \), a value common in the literature, and normalize \( N = 1 \). I simulate the model in Figure 9 varying \( \omega \) over a range. For this calibration, increases in \( \omega \) lead to increases in \( \rho_{QR} \) and \( \sigma_R \). Note that the overall effect on productivity from these shifts are somewhat mitigated by the fact that \( \sigma_Q \) also increases. On the other hand, the mean plant-level productivity also falls dragging down the overall industry average. In the model with no extensive, a 50% increase in \( \omega \) leads to an approximately 2.5% decline in productivity, roughly in line with my empirical estimates for cement. This will have further effects on labor use leading to a much greater decline in output than that of productivity.

6 Empirical Evidence on the Role of Financial Frictions

One testable prediction of my model and the shock that it emphasizes is that while the dispersion in \( TFPR \) is countercyclical, the dispersion in bank leverage ratios is procyclical. Using data disaggregated to the state-level, Figure 2 plots this crucial variable. It in fact moves quite closely with the dynamics of productivity showing a sharp fall between 1929 and 1933 and then a rapid
recovery from then onwards. There is some fragmentary evidence from modern data on the question of the cyclicality of the dispersion of leverage ratios. He et al. (2011) document that during the most recent crisis, hedge funds and the broker/dealer sector shed assets while the commercial banking sector increased asset holdings over this time. This is consistent with the idea that the variance in leverage across financial intermediaries declined over this time assuming hedge funds started with higher leverage ratios.

But what about other measures of the health of the banking sector? My model, while making a strong prediction about the procyclicality of the dispersion of leverage, also makes equally strong predictions that these productivity changes need not be correlated with other banking measures such as leverage ratios or the level of bank capital. The declines in the banking sector over the first four years of the Depression are all too clear. At an aggregate level between 1929 and 1933, the number of banks dropped from around 24,000 to 14,000, a 54% decline. Capital for all banks in the Federal Reserve system fell by 33% over a similar time frame and only recovered around 5% from 1933 to 1935. Banks in 1929 were leveraged 9.5 to 1 and over the course of the first half of the Depression, they deleveraged with the leverage ratio falling to around 8 in 1933. Hence, when asset prices collapse, the deleveraging process has multiplier effects on total loans, which fell by 66%, two and half times the decline in capital.

These declines in banking variables can surely match the downturn in productivity. Figure 2 confirms that fact with data from Comptroller of the Currency (1937). While not very disaggregated, state borders were meaningful economic boundaries due to prohibitions on interstate banking. The drops in capital, leverage, number of banks, and so on all line up well with the decline in productivity. However, each of these variables struggles to match other dynamics of aggregate productivity. For example, overall leverage is declining well in advance of the drop in productivity so that by 1929, it is already quite close to its nadir in 1933. On the other hand for bank capital, while there is a precipitous drop between 1929 and 1933, there is no real recovery afterwards. In fact, bank capital was still 13% below its pre-Depression peak in 1941. These facts provide additional evidence for the particular shock that I emphasize.\footnote{All the statistics reported in this paragraph come from Board of Governors of the Federal Reserve System (1941).}

\footnote{This decline is even smaller in real terms when the deflation occurring at this time is taken into consideration.}

\footnote{An earlier version of the paper examined the relationship between state-level bank leverage and TFPR for the ice and cement industries. When interbank markets are closed, there should be a positive relationship between bank leverage and TFPR. To understand this implication, consider the limiting case where interbank markets are perfect.}
7 A Competing Explanation: Demand Dispersion

A different explanation for the changing dispersion in revenue productivity could simply be a change in demand dispersion. In particular, assume instead that final output is given by

\[ Y = \left( \int_0^1 (Z_i Y_i)^{\psi-1} \right)^{\frac{\psi}{\psi-1}} \]

where \( Z_i \) can be interpreted as a demand shock for plant \( i \). This seems like a potentially attractive theory. It is well documented that the regional incidence of the Depression is not uniform across the country with some regions hit much worse (Rosenbloom and Sundstrom, 1999). Putting aside where those differences come from and why they would rise and fall over the Depression, such a theory would seem to explain both facts regarding dispersion and the correlation between productivity measures (at least for the cement industry). Areas that had relatively high demand would have high revenue productivity and at the same time, due to labor hoarding would appear to have high physical productivity as well.

The issue is that demand shocks do not translate into differences in revenue productivity. Returning to the model and the expression defining relative labor demand across islands \( i, j \), with demand shocks, this becomes

\[ \frac{r_{fi}}{r_{fj}} = \left( \frac{Z_i}{Z_j} \right) \left( \frac{A_i}{A_j} \right)^{\psi-1} \left( \frac{L_i}{L_j} \right)^{-\frac{1}{\psi}} \]

This still reduces to

\[ \frac{TFPR_i}{TFPR_j} = \frac{r_{fi}}{r_{fj}} \]

The point is that \( TFPR \) will now equalize demand-adjusted productivity in the sense that areas with high demand will on average have higher prices. However, the result is still the same that only differences in factor prices matter. So this cannot explain why dispersion changes.

What if markups were not constant but depended on demand? Then the expression for the

then there should be no correlation between the two variables as there is no variation in \( TFPR \). However, during the crisis period, the banks with the best projects are restricted from leveraging up further thereby generating a positive correlation. There was basically no relationship between the variables in 1929 and strongly positive one in 1931 and then a smaller (imprecisely estimated) effect in 1935.
ratio of $TFPR$ across the islands would be

$$\frac{TFPR_i}{TFPR_j} = \frac{r_{fi} \mu(Z_i)}{r_{fj} \mu(Z_j)}$$

where $\mu(Z)$ is the markup as a function of local demand $Z$. In this case, an increase in the dispersion of $Z$ could lead to an increase in the dispersion of $TFPR$. The difficulty with this explanation is that given the evidence for countercyclical markups (Floetotto and Jaimovich, 2008), it would fail to explain the rise in the correlation between $TFPR$ and $TFPQ$. In the best case, it would predict no relationship between the two. In the worst case, it would predict a negative relationship.

To see this, assume that island $i$ has high demand, then, all else equal, this would lead to lower markups and, hence, lower $TFPR$. Now if no assumptions are made, there is no reason to expect any relationship between $TFPR$ and $TFPQ$. This is in sharp contrast to the strong positive (and increasing relationship) between the two. However, if one went further and assumed that $TFPQ$ was at least partly driven by labor hoarding, then one would expect a negative relationship between the two. Areas with high demand should have less labor hoarding and, hence, higher measured $TFPQ$ while, under the assumption of countercyclical markups, $TFPR$ would be lower.

The issue is that this explanation of rising demand dispersion could still not explain why the correlation between $TFPR$ and $TFPQ$ becomes more positive over time, at least, in the cement industry. If the relationships between markups and labor hoarding with demand are roughly linear, then the correlation should be unchanged. To explain the increase in the correlation would require an additional assumption of a convex relationship between the markup and demand. Then greater dispersion would lead to an increase in the correlation due to the misspecified linear model. So this explanation is not inconsistent, but it does require a number of auxiliary assumptions, which are either in conflict with the evidence (procyclical markups) or have no clear theoretical basis (convex markup-demand relationship).

A related explanation emphasizes the geographic variation in bank failures rather than demand dispersion. As emphasized by Wicker (1996), while the breadth of the banking crises during the Depression were unprecedented, the incidence was not uniform across different regions. For example, in the first banking crisis of 1930, bank failures were basically restricted to the St. Louis Federal Reserve region and in New York City due to the collapse of the Bank of the United States.
In the beginning of 1931, the epicenter of the banking crisis moved to Chicago before spreading to the entire country. This broad scope of the crisis continued in 1933. It is possible to argue that this geographic variation in banking outcomes driving variation in demand in turn generated the variation in revenue productivity across regions.

The problem with this view like demand dispersion is that it has a hard time explaining the increase in the correlation between $TFPR$ and $TFPQ$. It is trivial to explain an increase in the dispersion of $TFPR$, but without auxiliary assumptions, the theory has no hope of explaining the relationship between $TFPR$ and $TFPQ$. In fact, there should be no correlation between $TFPR$ and $TFPQ$. Adding an assumption of labor hoarding would only make the relationship worst since presumably in areas with tighter credit and higher interest rates, output should be lower and, hence, measured $TFPQ$. I would argue that the joint behavior of $\rho_{QR}, \sigma_R$ is a compelling argument for an aggregate shock view rather than a direct shock to the dispersion of local economic conditions.

8 Conclusion

I have proposed a new hypothesis to explain part of the sharp decline in aggregate TFP during the Great Depression. It identifies an increase in the misallocation of resources across plants as playing a crucial role. To develop this hypothesis, I collected a plant-level dataset for a number of industries. First of all, I found that a key indicator of misallocation, dispersion in revenue productivity, increased substantially. I then quantified the fraction of the productivity decline explained by misallocation for one particular industry, cement. This is about 15% between 1929 and 1935 and higher in intervening years. I find that two sources of misallocation, dispersion in $TFPR$ and correlation between $TFPR$ and $TFPQ$, both show sharp increases and contribute to the overall productivity decline. Shocks to the financial sector can provide a plausible explanation for this productivity decline.

In the end, while many policy proposals are often heavily influenced by the current interpretation of the Great Depression, this is actually a case where extrapolating from the Depression to the current Great Recession might not provide useful. The explanation given here is very institutionally rooted in the history of the banking system in the 1930s. Again the comparison to Canada in the 1930s is useful. Canada did not have branching restrictions like the U.S. did (Bordo et al., 1994,
2011). Instead of the fragmented U.S. banking system, the banking market in Canada was heavily concentrated with a few large banks owning branches in nearly every small town. These wide branch networks allowed for the movement of resources within a bank itself. By insulating a sizable fraction of resource flows from the vagaries of interbank markets, this limited the potential damage on productivity from a collapse in interbank funding.\footnote{In ongoing work for the Depression (Loualiche and Ziebarth, 2013), I examine source of resource reallocation: “markets” inside of the firm. Preliminary results suggest that these markets helped to move resources to their most productive areas substituting for a lack of external financing.}

The U.S. banking sector of today shares many features in common with the banking system of Canada then and today. This may be why during the most recent financial crisis, productivity did not really decline like the Depression. The closest parallel to the American Great Depression instead is the modern emerging market crises. As noted above, these crises are often associated with large falls in aggregate productivity, and in general, these economies usually have much more localized finance similar to the U.S. of the 1930s. This is all to emphasize that not all financial shocks are created equal and, in many ways, what was experienced in the Great Depression is unlikely to be experienced ever again. In Twain’s words, “History does not repeat itself though it does rhyme.”
References


Bordo, M., A. Redish, and H. Rockoff (2011). Why didn’t Canada have a banking crisis in 2008 (or in 1930, or 1907, or ...)? NBER WP 17312.


Greenstone, M. and A. Mas (2012). Do credit market shocks affect the real economy? Quasi-experimental evidence from the great recession and 'normal' economic times. Unpublished, MIT.


Richardson, G. (2007). The check is in the mail: Correspondent clearing and the collapse of the banking system, 1930 to 1933. *Journal of Economic History* 67, 643–671.


Figure 1: From Cole and Ohanian (2000) and Amaral and MacGee (2002). Deviation from a linear trend.
Figure 2: Data for all banks disaggregated to the state-level from various volumes of the Report of the Comptroller of the Currency. All variables have been normalized to their 1926 value. Leverage is measured as the ratio of total assets to bank capital. The denominator includes capital, surplus, and undivided profits.
Figure 3: Dynamics of $\sigma_R$ across all industries. Revenue productivity for each plant has first been demeaned by the industry-year average. I report changes in the log value relative to 1929.
Figure 4: Dynamics of $\sigma_R$, $\sigma_Q$, and $\rho_{QR}$ in the cement industry. Recall that productivity here is total factor productivity using the capacity variable as a measure of the capital stock. Changes in correlation is in levels. For standard deviations in logs.
Figure 5: Productivity dynamics of cement industry for particular value of $\psi = 4$. 
Figure 6: Fraction of productivity decline from 1929 explained by changes in misallocation for cement varying $\psi$. 
Figure 7: Summary statistics for state-level net lending rates from Bodenhorn (1995) normalized to 1929 value. The largest and smallest observation have been trimmed from each year.
Figure 8: Comparing US GNP and TFP to that of Canada during the Great Depression.
Figure 9: Effect of marginal cost of leverage $\omega$ on steady state values with extensive margin. Initial values have been normalized to 1. For each simulation, $N = 1$. 