ABSTRACT

Title of dissertation: PRODUCTIVITY DISPERSION, PLANT SIZE, AND MARKET STRUCTURE

Sasan Bakhtiari, Doctor of Philosophy, 2008

Dissertation directed by: Professor John Haltiwanger
Department of Economics

Ample evidence from micro data suggests that productivity at establishment level is dominated by idiosyncratic factors. The productivity differences across establishments are very large and persistent even with the narrowest definition of industries. There is an attempt to identify sources of frictions that cause such productivity dispersion and negatively affect the average productivity of industries. This dissertation contemplates a non-monotonic relationship between productivity and input size and studies its importance in shaping the relationship between productivity dispersion and the producer size, a fact that is presented along with supportive empirical results. The role of market structure is then elaborated in creating the observed behavior.

The US Census of Manufactures reveals significant productivity dispersion at any employment level. Moreover, this productivity dispersion falls with employment size within most manufacturing industries. This pattern is considerably strong for establishments in industries whose products are primarily locally traded. It will be shown that general approaches such as industry selection and simple statistical aggregation do not explain this pattern convincingly, while sector-specific factors such as market localization can mimic this behavior much more closely.

Based on these results, a market structure model is introduced that uses de-

mand size and market localization as constraining forces to generate a bell-shaped relationship between input size and productivity within a market and for locally traded goods. The non-monotonicity of the relationship is a clear departure from most economic models where input size of plants is monotonically increasing with their productivity in the long-run. Because of the bell-shaped relationship, the proposed model predicts significant long-run productivity dispersion at any level of input size. Also this dispersion decreases with input size, in the same way as is observed in the data.

The model is calibrated and then simulated using data on Ready-Mix Concrete. First, the relationship between productivity and input size in the data is of a similar bell-shaped form. The effect of market size is also shown to be consistent with model predictions. Second, simulated results produce productivity dispersions that fall with input size with almost the same slope as observed in the data. This, in turn, suggests that the difference between simulated and actual productivity dispersions, summarizing the effect of other frictions, is almost uniform across sizes. Finally the robustness of the results is demonstrated through various tests.

Throughout the discussion, a distinction is made between physical and revenue productivities and the theoretical implications of both measures are shown to be qualitatively the same.

PRODUCTIVITY DISPERSION, PLANT SIZE, AND MARKET STRUCTURE

by

Sasan Bakhtiari

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Professor John Haltiwanger, Chair/Advisor
Professor Curtis Grimm
Professor Ingmar Prucha
Professor John Shea
Professor Daniel Vincent

Dedication

I dedicate this dissertation to my family, especially my parents whose enduring support was an immense source of inspiration for me.

Acknowledgement

I am indebted to Prof. John Haltiwanger whose support and advice was invaluable to me. I am deeply grateful to Hyowook Chiang for providing me with the data. I also thank Dr. John Shea, Dr. Ingmar Prucha, Dr. Daniel Vincent, Dr. Curtis Grimm, Dr. Michael Pries, Dr. Arghya Ghosh, and the participants in brownbag seminars in University of Maryland and Center for Economic Studies for their insightful comments. Finally, I would like to express my gratitude to Jeremy Wu, Kristin Sandusky, Fredrik Andersson, and all the LEHD staff at the Census Bureau for their kind support.

Disclaimer

This document reports the results of research and analysis undertaken by the U.S. Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This document is released to inform interested parties of research and to encourage discussion. This research is a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grants SES-9978093 and SES-0427889 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant R01 AG018854-02, and the Alfred P. Sloan Foundation. The views expressed on statistical issues are those of the author and not necessarily those of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau supports external researchers' use of these data through the Research Data Centers (see www.ces.census.gov). For other questions regarding the data, please contact Jeremy S. Wu, Program Manager, U.S. Census Bureau, LEHD Program, Room 6H141, 4600 Silver Hill Road, Suitland, MD 20746, USA (Jeremey.S.Wu@census.gov, http://lehd.did.census.gov).

Contents

List of Tables

List of Figures

List of Abbreviations

AR	Auto-Regressive
ASM	Annual Survey of Manufactures
BLS	Bureau of Labor Statistics
CBSA	Core-Based Statistical Area
CFS	Commodity Flow Survey
CM	Census of Manufactures
CRTS	Constant Returns to Scale
LBD	Longitudinal Business DataBase
LEHD	Longitudinal Employment-Household Dynamics
LRD	Longitudinal Research Database
rLP	Revenue Labor Productivity
rTFP	Revenue Total Factor Productivity
PPN	Plant Permanent Number
SSEL	Standard Statistical Establishment List
SIC	Standard Industry Classification
TE	Total Employment

Chapter 1

Introduction

Ample evidence from micro data suggests that productivity, i.e. the amount of output produced per unit input, at the establishment level is dominated by idiosyncratic factors. Recent studies of productivity are increasingly dependent on models of heterogeneous producers and less reliant on representative firm models. The importance of studying heterogeneity among producers is underlined by a growing body of literature in economics that emphasizes the role of reallocation of resources from less productive units to more productive ones, both within and across industries, as a significant source of growth in aggregate productivity. Reducing trade barriers or deregulating certain aspects of industries, for instance, makes markets more competitive and, as a result, drives out less productive units out of the market and causes the average productivity to rise. These findings are in contrast to earlier models where technology advancement entirely accounted for productivity growth.

Speculations abound about why productivity differences should exist. One possibility is that remarkable differences are observed in productivity because observations are being pooled across a wide range of industries and products, so that at some level of industry disaggregation the differences would disappear. Many attempts have been made to test this presumption by narrowing down the definition of industry. The collective results strongly suggest a persisting dispersion of productivity even with the narrowest definition of industry and even within industries with observed homogeneity of products. The data also points to the fact that most of the observed heterogeneity is real and not a result of measurement error. Wages also display much dispersion and they are less likely to suffer from mismeasurement. The strong correlation between the dispersions of productivity and wages across industries also suggests that differences in productivity are real, not an artifact of noisy data.

The appropriate public policy response to productivity dispersion depends on correctly identifying the causes of productivity dispersion. In a perfectly competitive market, the most productive unit can always eliminate competition by undercutting others in price and driving them out of the market. The existence of large differences in productivity implies market imperfections or frictions that give the less productive units the possibility to continue producing by slowing down reallocation of resources or stopping it altogether. In addition, decreasing returns to scale in production also prevents infinite growth of the more productive producers, giving a productivity advantage to smaller less productive units. The outcome is an industry with lower average productivity that operates below optimum capacity. The magnitude of productivity dispersion provides a sense of the level of imperfections or frictions that plague a given industry. A combination of supply-side and demand-side frictions contribute to this effect. Understanding the nature of these frictions could help in devising policies that reduce barriers to creative-destruction and raise aggregate productivity. Even if a certain friction cannot be addressed by policy measures, understanding the extent of its effects still provides an estimate of how much average productivity could, theoretically, be improved.

To explain the observed long-run distribution of productivity and size, most economic models generate a monotonically increasing relationship between produc-

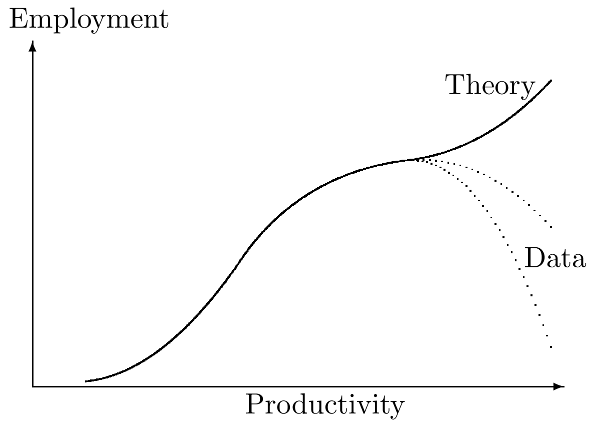


Figure 1.1: The relationship between productivity and employment as predicted by theory and observed in the data.

tivity and input size. In the absence of any shocks or uncertainties, frictions slow down growth, but more productive units still grow faster than other units and end up being larger in equilibrium. This monotonic size ranking of producers by their productivity is typical of most existing models. It is most commonly reasoned that if producer A is more productive than B, then A must be larger than B in the long run (Figure 1.1). This one-to-one relationship between input size and productivity generates zero long-run productivity dispersion at a given size level, or a limited amount of productivity dispersion when shocks and uncertainties are present. In view of such results, these models are more suitable for the analysis of mean productivity than for explaining the extent of productivity dispersion within plants of a given size.

The data, however, provides a contrasting picture of the productivity-employment relationship, where monotonicity of the relationship does not hold consistently at all levels of productivity. Within lower productivity plants, employment size mostly responds monotonically to productivity, so that higher productivity means higher size. This trend breaks down for high productivity plants, for which employment sizes are expected to be dominantly large but are mostly small to medium size (Figure 1.1-dashed lines). Restricting attention to old plants does not change the picture. This non-monotonicity can cause large productivity differences at lower levels of employment. Several reasons can be cited for why a highly-productive plant should choose a small employment level in the long-run. Prima facie, convex costs of labor and capital adjustment prevent rapid changes in size, making growth a gradual process. But modeling size growth as a process that depends simultaneously on production and demand conditions offers more interesting insight into a non-monotonic relationship between employment size and productivity.

This dissertation studies the importance of demand and supply conditions as constraints on size growth. I postulate that producers can grow rapidly in their employment only when favorable conditions hold in both supply technology and demand structure. Such a dual requirement for employment growth is reflected in the pattern of productivity dispersion, as producers with totally different productivities can still have the same long-run input sizes because they are subject to different demand conditions. This notion is a departure from the more conventional view that size growth is only a function of a firm's underlying productivity. Market localization is an aspect of demand structure that will be elaborated in this study by a theoretical model, as well as empirical tests, and shown to exhibit satisfactory explanatory power for the observed size distribution and patterns of productivity dispersion. Localized-market industries are actually a distinctive group of industries whose productivity dispersion falls at a considerable rate with employment size.

This dissertation also expands the understanding of productivity dispersion by examining its relationship with input size. Initial observations on the behavior of productivity dispersion have been the main motivation in contemplating a non-monotonic relationship between productivity and input size as an important source of productivity dispersion. In fact, data findings suggest that productivity dispersion is significant at every level of employment. Moreover, productivity dispersion falls with employment for many manufacturing industries. this finding is consistent with earlier work that noticed a similar relationship for wage dispersion among manufacturing plants. As will be shown in both theoretical and empirical results, a bell-shaped relationship between productivity and input size is capable of explaining such pattern convincingly.

In building up a case for the role of market structure, the roles of establishment-level dynamics and industry selection in shaping and trimming the distribution of productivity are also studied. Industry selection is shown not to play a considerable role in creating the observed pattern of productivity dispersion by employment size. After taking out young plants, thereby eliminating births and much of the dynamic volatility, the behavior of productivity dispersion by employment size is barely affected. I also find much persistence in productivity and employment size at the establishment level, especially within older plants. Manufacturing plants are fairly sluggish in adjusting their employment size, causing the distribution of productivity at each employment level to be mostly caused by a fixed subset of plants with a rather time-invariant distribution of size and productivity. Summarizing, longer-run factors such as demand structure are better suited to account for most of the observed behavior.

Simple statistical aggregation is one of the explanation with long-run implications that could generate a falling productivity dispersion by employment size. Larger plants are supposedly a collection of several smaller production units; consequently, they can average over a larger number of arriving productivity shocks and should perform closer to their mean productivity. However, the additivity of plant behavior is not supported either theoretically or empirically. The conditions under which production functions are additive are very restrictive and are rarely satisfied in practice. I will demonstrate the inability of statistical aggregation in explaining the observed pattern of productivity dispersion by running a counter-factual test, where larger plants are simulated by aggregating an appropriate number of smaller plants randomly chosen from the pool of all small plants. Comparing the results, this approach fails mostly because productivity dispersion falls much faster in the simulated plants than among the actual plants.

Recent works indeed suggest that market structure is a considerable force in shaping the distribution of productivity. Product substitutability, market size and trade possibilities have been show to account for a significant part of observed productivity differences within and across industries. A producer's profit depends on both its efficiency of production and its demand structure. The plants that survive are not necessarily those that are more productive, but also those that face more favorable demand conditions. This is true especially when markets are localized. I posit that when an industry faces localized markets, then market size becomes an important parameter in the determination of plant behavior. In response to limited demand, high productivity plants must choose low input levels to stay profitable. This behavior, in turn, creates a productivity-size relationship similar to the dashed line in Figure 1.1. As a result, the range of productivities at small-sized plants increases dramatically. Only when demand becomes large or inelastic enough can the top productivity plants grow large and make enough profit to justify their growth. This assertion is supported by the empirical observation that productivity dispersion falls at a considerable rate within localized market industries.

For my theoretical approach, I use a differentiated-product framework with localized demand markets. The main implication of such an approach is that it can endogenously create a bell-shaped relationship between a plant's productivity and its size defined in terms of a composite input factor. This relationship, in turn, produces more productivity dispersion for plants with lower input levels. Allowing for a distribution of markets with different sizes (say a segmentation of a national market) helps to make the productivity distribution denser at any given input size and to bring the results closer to reality.

To test my model, I use data on Ready-Mix Concrete, an industry that features both market localization and spatial product differentiation. In addition, the homogeneity of its output helps me reduce productivity variations caused by differences in taste and quality. Estimated moments from the data for the concrete industry show bell-shaped relationships between employment size and productivity that closely resemble model predictions. Using these moments, the model is then calibrated and simulated to create productivity dispersions at different levels of employment. The model is successful in simulating productivity dispersions that fall with employment size at roughly the same rate as is observed in the data. As a result, the gap between the actual and simulated dispersion curves is mostly uniform and can be attributed to an additive variation caused by a combination of other frictions affecting the industry.

To begin with, the next chapter reviews the related literature on productivity dispersion and wage dispersion. Earlier works on the differentiated-product model are discussed, and their relevance to my analysis is pointed out. The data and the productivity specifications that will lead me through my analysis are then discussed in Chapter 3.

In building the case for the role of market localization in forming productivity dispersions, Chapter 4 looks at patterns of productivity dispersion in overall manufacturing and across industries. Industry selection and statistical aggregation are both tested for their explanatory power, and both are rejected for failing to offer a convincing explanation of the observed behavior of productivity dispersion. The dynamics of plants contributing to the observed distribution are then investigated using transition matrices among productivity-size states, and particularly the high persistence of both productivity and employment size among old plants is demonstrated. I then shift attention to four-digit industries, and it is shown that the relation between productivity dispersion's and employment is very different across four-digit industries.

Market localization will be the focus for the rest of the dissertation. Using average shipment distance for four-digit industries, it is possible to classify industries roughly into localized versus non-localized, with localized-market industries being those whose market reach is most limited compared to the rest. Empirical observations and more rigorous tests show that productivity dispersion falls at a considerable rate within the localized-market industries and on average much faster than in the industries with non-localized markets. Further tests actually show that the degree of market localization is strongly related to the slope by which their productivity dispersion falls with employment.

Motivated by the results of Chapter 4, Chapter 5 introduces a differentiated-product model in which plants are subject to localized markets, i.e. any interaction among markets is ruled out by assuming that the cost of trading among markets is infinity. With differentiated products, plants will act as monopolists, making it easy to define the shape and the size of the demand curve for each producer. Market localization, in effect, creates a cap on how much output can be delivered by each plant, depending on the size of the corresponding market. Less productive plants will be unrestricted in deciding how much output they are going to produce. However, more productive plants with potentially larger output capacity than the cap, and the inability to improve their demand by trading with other markets, have to hire smaller input sizes to produce the output cap. This constraint keeps plants smaller and smaller as their productivities get higher. In this way, the relationship between productivity and input size will not be monotonic as in previous models, but will take a bell-shaped form. The bell-shaped relationship , in turn, generates significant productivity dispersion that varies by input size. The model is calibrated and simulated with Ready-Mix Concrete and the similarity of outcomes is demonstrated. Additional robustness tests are undertaken to demonstrate the insensitivity of the model implications to structural change in the model composition and market definition.

Finally, Chapter 6 summarizes the results of the study.

Chapter 2

Literature Review

A host of works have documented the existence and the extent of productivity dispersion within different industries. Bartelsman & Doms (2000) review several works and emphasize the role of individual producers in creating productivity dispersion. Based on empirical observations, they argue that most of the observed heterogeneity is real and not a result of measurement error. Relative productivities, wages and technology usage are shown to be highly correlated, pointing to the fact that heterogeneities among plants, and not measurement error, are driving the observed differences in productivities (Dunne, Foster, Haltiwanger & Troske 2000). More recently, Haltiwanger, Lane & Spletzer (2000) find significant wage and productivity dispersion among narrowly defined universes of restaurants and plumbing in Maryland, US. At the lowest possible level of disaggregation, Chew, Clark & Bresnahan (1990) look at plants belonging to the same multi-plant firm where the same technology and the same input is used to produce the same output. The difference between the most and least productive plants is an astounding 3:1 ratio[1]. They contemplate reasons why these plants do not converge in productivity as a result of the managerial and information networks that are supposed to exist within

[1]As case study, they choose a food manufacturing chain with plants spread around different cities in the US.

establishments belonging to the same firm. They hypothesize that most of the managerial decisions are decentralized and depend on the establishment-level quality of manager-job matches. In close relation, Abowd, Kramarz & Margolis (1999) show that productivity differences among firms can be attributed to heterogeneity in both technology and labor skills, so that the idiosyncratic factor of productivity can be partially accounted for by each.

However, the distribution of productivity is not static. The cross-sectional distribution of productivity is constantly altered as establishments enter or exit the market. Establishment-level productivity is also constantly churning as a result of idiosyncratic productivity shocks. Some prominent works that address dynamics of productivity in a theoretical context are those of Jovanovic (1982), Hopenhayn (1992), and Ericson & Pakes (1995). In a general setting, plants enter the market randomly drawing their productivities from the full range of a known distribution. Plants do not observe their exact productivity ex post and are hit by productivity shocks every period. Alternatively, they form a Bayesian estimate of how productive they are from the past string of noisy observations and make growth or exit decisions based on that estimate. More productive plants, or plants hit by a string of favorable shocks, stay and grow. On the other hand, less productive plants, or those hit by a string of unfavorable shocks, exit the market. These models differ in the persistence and mechanisms of productivity shocks. Jovanovic does not assume any shock persistence, while Hopenhayn argues that favorable or adverse shocks last for longer periods of time. Ericson & Pakes add more structure to the shock mechanism by attributing shocks to uncertain outcomes of technology investment.

Baily, Hulten, Campbell, Bresnahan & Caves (1992) were probably the first to use data to look at the dynamics of productivity and their implications for productivity dispersion. Together with Bartelsman & Dhrymes (1998), they provide a very helpful insight into the evolution of an industry from an empirical perspective.

Their papers find not only a wide range of productivity differences among four-digit manufacturing industries, but also a rather time-invariant distribution. Most notably, establishment-level productivities show considerable amount of persistence over time, contrasting Jovanovic's vision and supporting Hopenhayn's idea of correlated shocks. A high-productivity plant most probably stays high-productivity even after several years, and the same can be said about other levels of productivity.

The importance of productivity dispersion is demonstrated by Haltiwanger (1997) and Foster, Haltiwanger & Krizan (1998) who show that a considerable part of aggregate productivity growth within industries is caused by micro-level reallocation of resources from less productive producers to more productive ones. This result suggests that policies targeted at reducing excess productivity dispersion can be important. In a particular example, Olley & Pakes (1996) demonstrate that during the deregulation of telecommunication industry in 1980s, a large fraction of aggregate productivity growth in that sector occurred because of lower productivity plants getting out of the market and their resources being reallocated to other plants within the same sector.

This dissertation examines the relationship between productivity dispersion and employment size in an effort to identify frictions that produce matching behavior of productivity dispersion with what is observed in the data. Previous works has shown that plants of different sizes behave differently. Hall (1987) shows that larger plants grow more slowly and are less likely to fail, while Brown & Medoff (1989) show that larger plants offer higher wages on average. Also larger plants generally have their own research division and spend more on R&D activity (Acs & Audretsch 1991) and have more sophisticated management and organizational structures (Churchill & Lewis 1983, Greiner 1998). In the economic literature, however, there has not been much discussion about how productivity dispersion should be shaped by employment size. Davis & Haltiwanger (1991) provide some clue by showing that wage dispersion among manufacturing plants is significant at any employment level and falls with employment for both production and non-production workers. My studies show that a similar relationship holds between size and productivity dispersion.

In this study, market localization is used to build a non-monotonic relationship between productivity and input size that will be the source of sustained long-run productivity dispersion. In comparison, most economic models generate a monotonically increasing relationship between productivity and input size in the absence of shocks and uncertainties. For example, Bertola & Garibaldi (2001) and Bontemps, Robin & Van Den Berg (2000) use job matching and job search frictions, respectively, to produce the monotonic relationship between productivity and size. In presence of shocks and uncertainties, a limited distribution of productivity can be sustained at each size level. Good examples of such frictions are costs of labor adjustment (Hamermesh 1995) or costs of capital adjustment (Abel & Eberly 1996) that cause firms within some range of productivity not to make any adjustment decisions. However, as long as adjustment costs are independent of size and scale of operation, there is no reason to believe that such frictions would generate different productivity dispersions at different sizes.

From a different point of view, Churchill & Lewis (1983) and Greiner (1998) study costs of reorganization and restructuring as firms and their establishments go through several stages of growth. Fixed costs associated with transition between two stages of growth can divide firms into two groups: those who pay the fixed cost and grow beyond the barrier, and those who are unable to overcome the cost and whose growth ends there. Another way of thinking about this is to assume that plants must have accumulated a certain amount of "managerial capital" or must have achieved a certain level of "marketing skill" to be able to pass this barrier.

Only recently has it been known that market structure can play a major role in shaping productivity dispersion. Syverson (2003) noticed that product substitutability within an industry can affect the dispersion of productivity for that industry. As products become more substitutable, it gets harder for less productive plants to compete with more productive ones in price, forcing them to exit. With more differentiated products, the less productive units can still fill a niche in the market, and as a result, they get the chance to stay in the market and continue producing. Substitutability, in this context, is not limited to product diversity, but can also be caused by the degree of spatial differentiation, as is the case for the concrete industry (Syverson 2004). Due to the high transportation cost of concrete, customers make purchase decisions based on physical distances as well as prices. In this setting, the lower productivity plants can still survive because it is costlier for customers close to them to buy from farther plants, even if the prices are lower. The role of demand market in industry selection is also emphasized in Foster, Haltiwanger & Syverson (2008), who show that both profitability and productivity affect the selection process. The plants that survive are not necessarily those that are more productive, but also those that face more favorable demand conditions.

Several recent works have demonstrated that market size can also affect industry conduct. Melitz & Ottaviano (2005) show that enlarging market size or lowering trade costs reduce productivity dispersion of the operating firms by making competition more intense, thereby driving out the less productive firms out of the market. Some empirical evidence for the effect of market size is provided by Berry & Waldfogel (2003), who show that for restaurants and daily newspapers, both the average quality of service and the number of establishments rise with market size. Similarly, Asplund & Sandin (1999) show the same relation between number of establishments and market size for Swedish driving schools. The size of operating retail stores is also shown to be positively correlated with their market size (Campbell & Hopenhayn 2002).

For my theoretical model, I borrow from both Syverson (2003) and Melitz

& Ottaviano (2005) to build a differentiated-product model that can address both plant behavior and market size by offering flexibility in the definition of production function and consumption utility. This type of model is based on Dixit & Stiglitz (1977), who studied optimality of product diversification within a social welfare system. The model produces tractable solutions. Size and elasticity of demand can also be easily incorporated into the model.

Chapter 3

Productivity

Below, the Census of Manufactures (CM) is described in some detail as the main source of data. I supplement and enhance the CM with other datasets to provide complete sets of measures for production and market analysis. Measures of productivity that will be the basis for analysis throughout this dissertation are also defined below, and practical issues concerning those measures are discussed. In particular the measures of productivity used in this dissertation are "revenue" measures, computed using the deflated values of sales and cost-shares of input, instead of physical inputs and output.

3.1 Data

The main source of data in this dissertation is the US Center for Economic Studies' Census of Manufactures (CM). McGuckin & George A. Pascoe (1988) provide a detailed discussion of how the CM is composed and conducted. Briefly, the CM is conducted quinquennially in years ending with "2" and "7" and is the census of about 360,000 manufacturing plants in the United States.

The unit of observation in the CM is plant, defined as an individual physical location of production and identified by a Plant Permanent Number (PPN). This identifier is useful in building longitudinal links to study the dynamics of productivity and size.

The CM also provides information on plant observables and performance. Some of the reported variables are the total shipment value, employment for production and non-production workers and total hours worked, book values of machinery and structures and costs of materials and energy. For each plant the four-digit Standard Industry Classification (SIC), product class, and location (state-county) are also reported in the CM. The location information, especially, enables me to link each plant to its corresponding market for analyzing supply-demand relations. I use the real values for inputs and output constructed by Chiang (2005). Chiang uses the 4-digit deflators available from NBER/CES Productivity Database[1] and estimates real equipment and structure capital using a perpetual inventory model.[2].

In a later stage, I separate young and old plants and specifically look at the distribution of productivity within old plants. This will be useful in eliminating the effect of births on productivity dispersion and focussing on long-run behavior of plants. Jarmin & Miranda (2002) provide estimates for the age of plants in the CM using the US Center for Economic Studies' Longitudinal Business database (LBD). These estimates are linked to the CM and are used for age classification. A small fraction of assigned ages suffer from some estimation error, but age estimates are merely used here to distinguish old plants from young ones, thus minimizing the possibility of any major bias.

Some of the plants in the CM have missing or invalid state-county data. The Standard Statistical Establishment List (SSEL) is used to correct the geographical

[1]Refer to J.Bartelsman & Gray (1996) for more details.

[2]The perpetual inventory model assumes that capital evolves in the following form

$$K_t = (1 - \delta_t)K_{t-1} + I_t,$$

where K_t is the real capital stock (equipment or structure) at time t, and I_t is the real capital investment. δ_t is the depreciation rate of capital.

information for those plants.

I use Total Employment (TE) defined by Davis, Haltiwanger & Schuh (1996, Appendix A.3.1) as my main measure of employment size. The CM reports the number of Production Workers (PW) in four quarters and the annual number of Other Employees (OE) [3] for each plant. Total Employment for plant j is defined as

$$TE_j = OE_j + \frac{1}{4}\sum_{t=1}^{4} PW_j(t). \tag{3.1}$$

Here t is the index for quarters within a year. This measure corresponds to an average annual employment size rather than a point in time estimate.

3.2 Selection Criteria

A weighted subset of about 60,000 plants from the CM also appear in the Annual Survey of Manufactures (ASM). The weights are the reciprocal of the probability with which each plant is selected into the ASM. For most of the unweighted plants in the CM, the majority of whom are plants with small employments, all data other than employment size are imputed from administrative records. The quality of these imputes is in doubt and can adversely affect the accuracy of statistics for small plants. Since identifying those imputes is not completely obvious, I only use the ASM plants for my analysis to avoid serious errors.

Some plants in the data have excessive sizes and are believed to be administrative errors. For that reason, I exclude plants larger than 50000 employees and also those plants whose industry code is other than manufacturing from my analysis. Also, to limit myself to well-defined industries, I drop plants belonging to any four-digit SIC code ending in 9. These codes collect plants that could not be classified under any other detailed classification in the same two-digit or three-digit industry

[3] Every other employee that is on payroll in the pay period including March 12.

code.

I also exclude plants belonging to the two-digit SIC code 21 (Tobacco) from my analysis. Tobacco plants are disproportionately larger than their other manufacturing counterparts[4]. Because of that, tobacco does not seem to be an interesting industry for size analysis.

Finally, only plants located in the US 50 states are kept in my dataset[5]. This reduces the data set used for my analysis to 397 industries and a total of 202593 establishments belonging to the four census years included.

3.3 Measures of Productivity

I lead my study of productivity dispersion using revenue Total Factor Productivity (rTFP) based on a Cobb-Douglas production function. Since the CM lacks plant-level information on the prices of input and output, the productivity measures are computed using input cost shares and using deflated revenue as real output. Some recent literature has emphasized the distinction between revenue productivity and physical output productivity[6].

In particular, revenue measures of productivity are driven not only by the efficiency of production, but also by variations in input and output prices, and differences in product quality and taste across plants. While revenue productivity is generally intended to compare physical performance of different plants in the same industry, when applied to industries with large diversity of products, the revenue productivity is more a measure of revenue per unit input expenditure (Katayama et al. 2003). A partial solution to this issue is often to do analysis on industries with observed homogeneity of products, some of which are listed by Foster et al. (2008).

[4]Tobacco plants are about 6 times larger than an average manufacturing plant and about 7 times larger than a median manufacturing plant.

[5]Plants in other US territorial regions such as American Samoa,, Guam, Porto Rico, and Virgin Islands were excluded, while plants belonging to Alaska and Hawaii are still kept in the data.

[6]See Katayama, Lu & Tybout (2003) or Foster et al. (2008).

In general, to make the distinction clear, I use the term "revenue productivity" to address the computed measures of productivity using real values of inputs and outputs.

For a plant j belonging to industry i at time t, rTFP is defined as

$$rtfp_{ijt} = q_{ijt} - \alpha_i^h h_{ijt} - \alpha_i^{eq} k_{ijt}^{eq} - \alpha_i^{st} k_{ijt}^{st} - \alpha_i^e e_{ijt} - \alpha_i^m m_{ijt}, \tag{3.2}$$

where lower case letters here label variables in logs. Here q is the nominal output deflated by industry-specific price indices. h is labor input (total hours worked), and k^{eq} and k^{st} are the equipment and structures capital stocks, respectively. e is energy and m is material input. The α coefficients are computed for each industry sector using the cost share indices described by Chiang (2005). rTFP provides a detailed measure of productivity taking into account various productive factors. However, it produces relatively noisy estimates due to inaccuracies in, or in some cases unreported, data on capital stock or other input factors. This problem is more common with smaller plants, causing the level of measurement error to be nonuniform across sizes.

For robustness, I also compute revenue Labor Productivity (rLP), which follows the standard definition

$$rlp_{ijt} = q_{ijt} - h_{ijt}. \tag{3.3}$$

Again, lower case letters denote variables in logs. A small number of rLP estimates are missing, mostly due to unreported total hours worked (h_{ijt}). I impute those values by regressing the log of available hours worked on industry dummies, log of total employment, and log of output size. With the CM data, the regression model produces an R^2 of 0.957, rendering the level of imputation noise insignificant and making the imputes a practical addition. The imputed information, in turn, enables me to estimate some of the missing rTFPs too, those missing only total hours. No further effort was made to impute other missing rTFP's when estimates of capital stock or energy and material were missing. This is partly due to the noisier nature of reported capital, energy and material that causes most models to accumulate noise when imputing, casting doubt on how well the imputed values can be trusted.

I also need both productivity measures, rTFP and rLP, to be comparable over a range of industries and years. Hence, I construct and use residual productivities by regressing logs of rTFP and rLP on year and year by industry dummies.

In comparison to rTFP, rLP is less detailed in assessing the contribution of different inputs, but it is also less prone to measurement error because total hours worked is easier to measure. It is also reported by almost all plants and easily imputed for the missing ones. Over time, most plants show a strong correlation between their rTFP and rLP estimates. The qualitative similarity of the two measures enables me to argue that the observed behaviors are not a result of measurement error or selection bias on rTFP, but a reflection of an underlying real effect.

Chapter 4

Employment and Productivity Dispersion

Numerous supply-side and demand-side frictions are believed to be the reason behind the observed range of productivities in the data. Most of these frictions are sector specific. Subjecting disaggregate industries to different types or levels of frictions can then explain the differences in the dispersion of productivity among different industries.

In this chapter I look at the distribution of productivity within manufacturing and then within four-digit industries in manufacturing. The goal of my study is to investigate the existence of an empirical relationship between productivity dispersion and employment size. The relationships between average wage and size and between average productivity and size are already well documented, with both average wage and average productivity rising with employment size. On the other hand, existing literature is mostly silent as to whether there should be any long-run difference in the distribution of productivity across different levels of employment and if that relation should be influenced by certain frictions. Apart from models of industry selection, in which selection trims the distribution of productivity as plants entering the market survive and grow, there is little reason a priori to believe that the distribution of productivity must respond to the size and the scale of operation. However, Davis & Haltiwanger (1991) offer empirical evidence that wage dispersion among manufacturing plants responds to employment size. Particularly, they observed that wage dispersion is large at any employment level and falls with employment. These results hold for both production and non-production workers. With Dunne et al. (2000) showing a strong link between the dispersion of wages and that of productivity, it is reasonable to think that productivity dispersion will also fall with employment. In fact, initial observations presented in this chapter are consistent with this picture. I also show that simple explanations such as industry selection and statistical aggregation are inadequate in explaining the range of observed productivity dispersion and the associated behavior with employment size.

Further tests show strong persistence in productivity at the plant-level over considerable lengths of time, consistent with prior work. In addition, a high level of persistence in employment size is also observed. Furthermore, I find that the relationship between productivity and employment size does not seem to be monotonic even in the long run.

In light of these facts, I search for sector-specific factors to explain the behavior of productivity dispersion. In fact, at a more disaggregate level, industries can be totally different in how their productivity dispersion changes with employment. Though, productivity dispersion falls with employment for a majority of four-digit industries, for some industries productivity dispersion rises with employment. I explore the role of sector-specific factors, in affecting the slope by which productivity dispersion falls with employment size. Emphasis will be given to testing the role of market localization, and I find that localized-market industries show a distinctive picture where productivity dispersion falls much faster with employment than in other manufacturing industries.

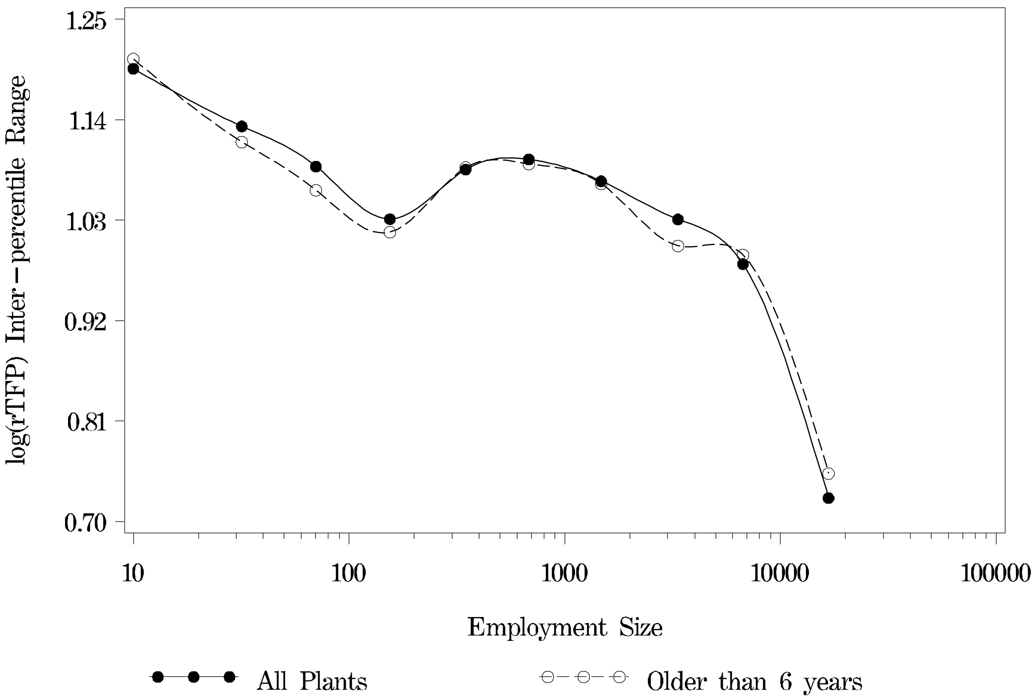
4.1 Pattern of Productivity Dispersion

To describe the relationship between productivity dispersion and employment, I start by breaking the range of employment size into classes. I use 95-5 inter-percentile range (denoted in the results by Δ) to compute productivity dispersion. Compared to standard deviation and other inter-percentile ranges, the 95-5 range keeps the fullest range of useful observations while effectively eliminating outliers, especially at lower employment levels. The range is computed for log revenue productivity within each class. The CM sample weights are used in all computations. I am borrowing my size classification from Baily, E.J.Bartelsman & J.Haltiwanger (1994) as listed in Table 4.1. This size classification has the benefit that it can effectively capture the behavior of productivity dispersion on the full range of employment sizes. The table also shows the average employment and summary statistics about revenue productivity within each class. Both rTFP and rLP are used as measures of productivity. Also, for better visualization, productivity dispersions are plotted in Figure 4.1.

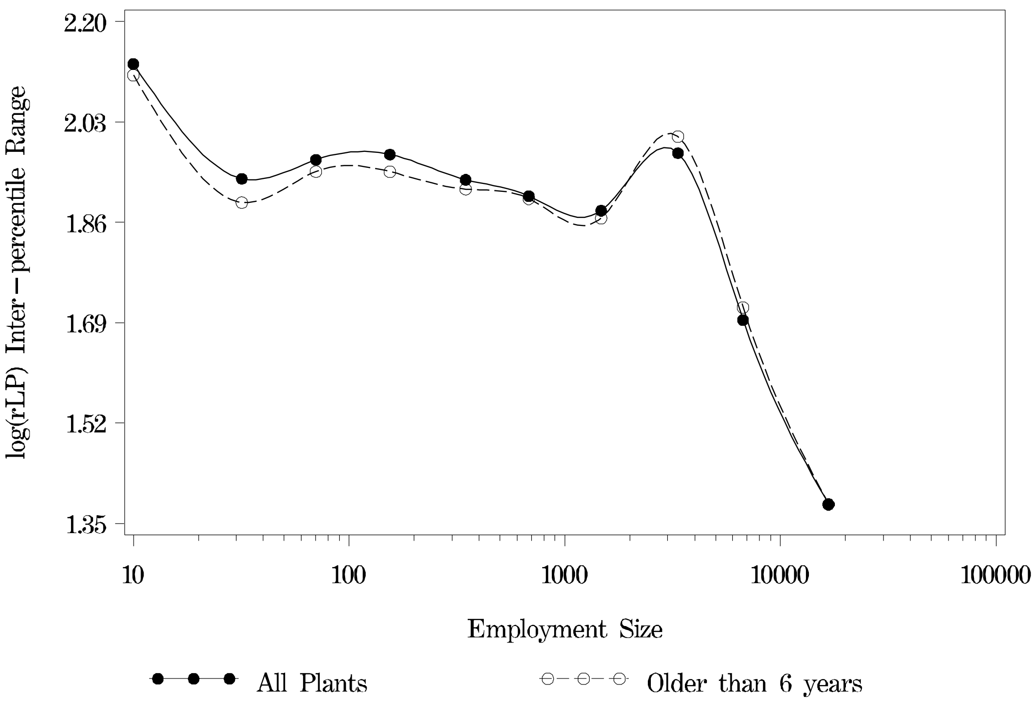
The trend of mean productivity with employment size follows prior findings that larger plants are more productive on average. Both measures of productivity mostly agree on this. At the same time, productivity dispersion seems to show a consistent pattern with employment size too. Two important points about the relationship between productivity dispersion and employment size can be noticed which are:

1. Productivity dispersion does not vanish at any employment level and remains markedly large at any given employment size. For employment class 1-19, the most productive and the least productive producers show a remarkable 12:1 productivity difference for rTFP [1]. This range is much larger with rLP. For

[1]Productivity range is actually computed for 1 and 99 percentiles of productivity to prevent outliers from inflating the range.



(a)



(b)

Figure 4.1: Inter-percentile range of log revenue productivities by employment size classes.

Employment	#Obs	$\overline{Emp}$	$\overline{\log(rTFP)}$	$\Delta\log(rTFP)$	$\overline{\log(rLP)}$	$\Delta\log(rLP)$
1-19	48034	9	0.009	1.196	-0.022	2.129
20-49	36918	32	-0.013	1.133	-0.070	1.934
50-99	32711	70	-0.009	1.089	-0.003	1.966
100-249	43487	155	-0.010	1.031	0.087	1.975
250-499	23911	348	0.017	1.086	0.162	1.932
500-999	11328	681	0.042	1.097	0.269	1.905
1000-2499	4671	1477	0.070	1.073	0.313	1.880
2500-4999	1032	3350	0.087	1.031	0.494	1.978
5000-9999	375	6706	0.038	0.982	0.394	1.695
10000+	125	16726	0.007	0.726	0.396	1.382

Table 4.1: Employment classes and summary statistics for each class.

plants with more than 10000 employees, the range of observed rTFP's reduced but is still a significant 4:1 ratio.

2. Both panels of Figure 4.1 imply that productivity dispersion gradually falls with employment at most levels. At largest levels of employment, productivity dispersion falls more rapidly.

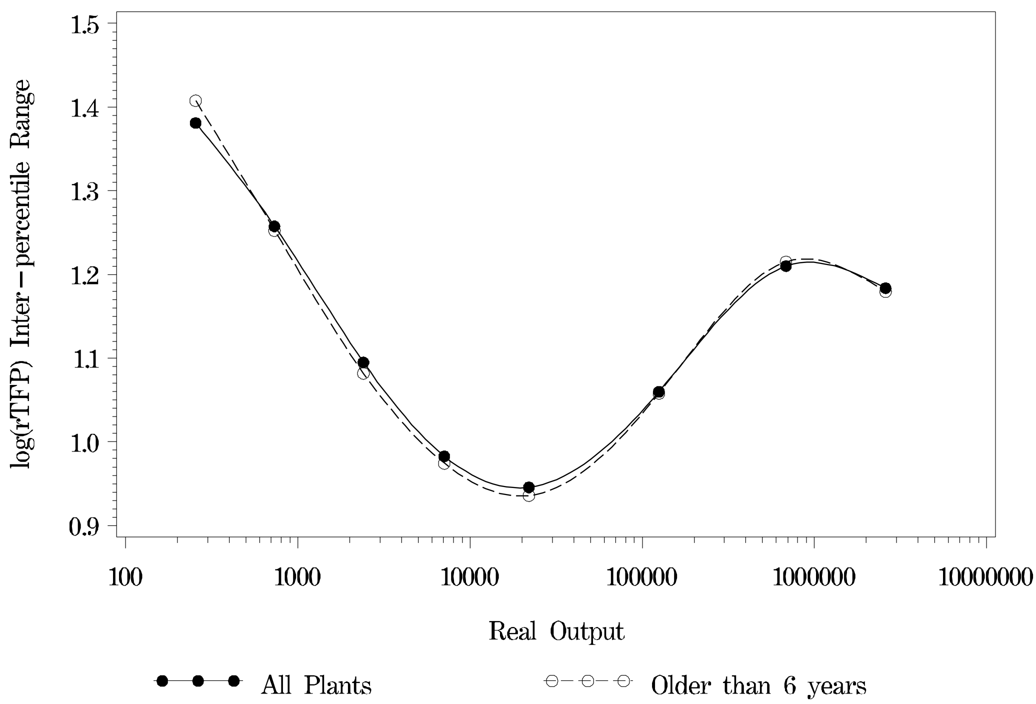
Above all, there is a consensus between both measures of productivity about how productivity dispersion behaves with employment size. At the same time, the observed relationship is consistent with the results for wage dispersion and employment size found in Davis & Haltiwanger (1991).

Employment is not the only way to measure plant size. Output size is also an indicator of the scale of operation at plants. Commonly, a plant's market share of output is seen as a measure of plant's influence on the market and the economy as a whole. Nevertheless, there are some difficulties interpreting results when using output size instead of input size. Due to lack of information on physical output, output is measured as the deflated value of shipments. Manufacturing plants are very diverse in their product mix and when pooling across all manufacturing industries, one must decide how much of the observed variation in productivity dispersion is related to actual size differences and how much is the distortion caused by product diversity. On the contrary, labor is relatively homogeneous across plants, so that employment size is a relatively consistent way of comparing different plants, even if those plants do not produce the same product[2]. Furthermore, Davis & Haltiwanger (1991) did not consider output size, so there is no prior expectation of how productivity dispersion should behave with output size.

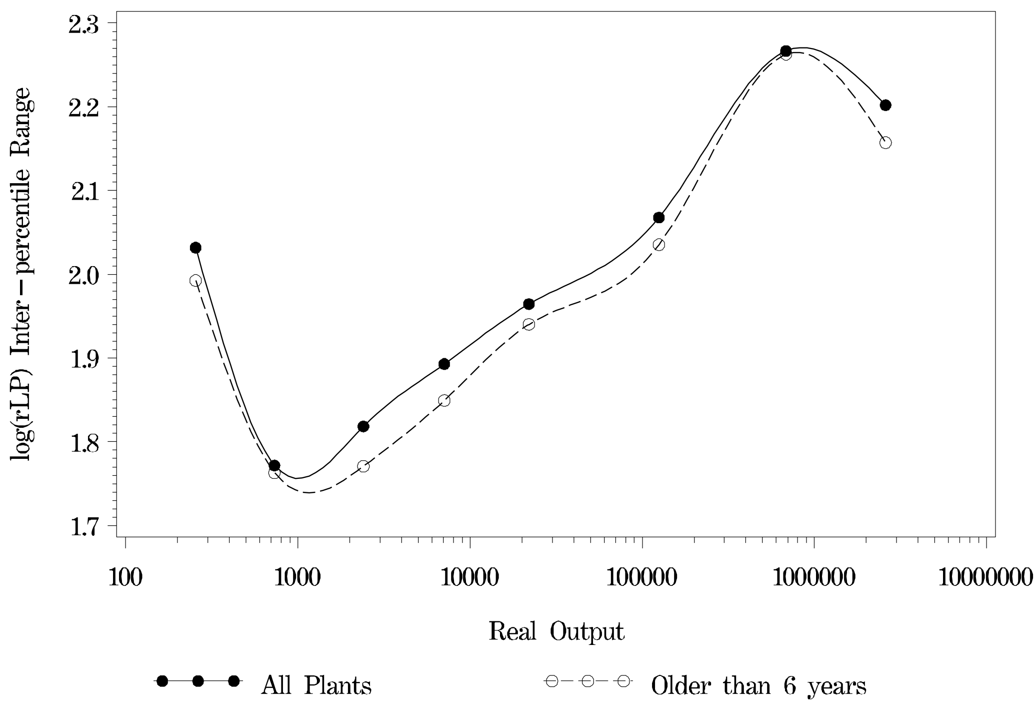
To explore the behavior of productivity dispersion by output size, I proceed as before by classifying the size of output and computing productivity dispersions using the 95-5 inter-percentile range of log revenue productivity. The results are listed in Table 4.2 and plotted in Figure 4.2. Again, the range of productivities observed at any output level is very large. But productivity dispersion does not seem to fall monotonically with output size. In fact productivity dispersion mostly increases with output. The interpretation of these results is also clouded by the fact that the shape of curves is rather different with rTFP and rLP. As explained in the previous chapter, rTFP takes into account the effect of different input factors that could capture the composition and texture of output in a better way than employment only. This detailed specification of production may explain the difference in the relationship between productivity dispersion and output size between the two measures of productivity.

Returning to the results with employment, the observed relationship for productivity dispersion triggers speculation about the type and extent of frictions that

[2]In practice, accumulation of plant-specific human capital can still differentiated among different groups of laborers.



(a)



(b)

Figure 4.2: Inter-percentile range of log revenue productivities by output size classes.

Output (×1000)	#Obs	$\overline{Output}$ (×1000)	$\overline{\log(rTFP)}$	$\Delta \log(rTFP)$	$\overline{\log}(rLP)$	$\Delta \log(rLP)$
0-0.5	21477	0.25	-0.035	1.381	-0.286	2.032
0.5-1	14517	0.73	-0.002	1.258	-0.148	1.772
1-5	49304	2.4	-0.001	1.095	-0.029	1.818
5-10	27841	7.1	-0.008	0.983	0.118	1.893
10-50	62493	22.0	0.015	0.946	0.296	1.964
50-100	24870	125.3	0.095	1.060	0.565	2.068
100-500	1218	685.3	0.164	1.210	0.819	2.267
500+	872	2596.3	0.115	1.184	0.842	2.202

Table 4.2: Output classes and summary statistics for each class.

can generate such a pattern. Some obvious answers lie in industry selection and statistical aggregation. In the next sections, I scrutinize each of these explanations and find that neither industry selection nor statistical aggregation can account for much of the decline in dispersion with size. Looking at the joint plant-level size and productivity dynamics, much persistence is observed in both, irregardless of how productive those units are. This leads me to consider longer-run explanations for the observed pattern. I will then look at the effect of sector-specific factors, especially market structure, in explaining the behavior of productivity dispersion.

4.2 The Role of Industry Selection

A natural way to explain the observed pattern of productivity dispersion is that declining dispersion with size is an outcome of industry evolution, as first described by Jovanovic (1982). In a Jovanovic type model, plants entering market feature the full range of possible productivities. As plants get older, the selection process claims the low productivity plants and forces them out of the market. As a result, the lower tail of the productivity distribution is trimmed out at larger employment levels, leaving a narrower range of productivities that survive at those sizes. Under such conditions, the pattern of productivity dispersion with employment size should resemble the observed one. It remains to be seen how much of a fall in productivity dispersion can actually be achieved by selection process alone.

To test the role of industry selection in shaping the behavior of productivity dispersion, I compare the existing picture to that achieved when restricting my attention to plants that are older than 6 years. This age restriction still provides me with a rich enough set of plants, about three-quarters of all plants. This restriction also eliminates births and dynamically volatile young plants, thus leaving me with a more stable set of plants[3]. At the same time, the effect of deaths on my analysis is not completely eliminated by setting this age limit, but it is certainly reduced significantly, as older plants are much less likely to fail than their younger counterparts (Evans 1987, Hall 1987)[4]. The effect of deaths on the distribution of productivity will come into more light in the next section where plant level dynamics are studied in more detail.

Figure 4.1 also shows productivity dispersion curves computed using only plants that are at least 6 years old (the dashed lines) along with those computed pooling over all plants. The discrepancy is clearly minimal, confirming that the observed behavior of productivity dispersion by employment is mostly unaffected by industry selection. This effect is shown to be the same when output size is used instead of input size. Figure 4.2 is the depiction of this latter result.

[3]By Davis et al. (1996) classification, such plants are "middle-aged" to "old" plants.

[4]Hall (1987) shows that plants older than 6 years were about 30% less likely to fail than plants younger than 6 years.

4.3 Testing Statistical Aggregation

In this section, I test the role of statistical aggregation in accounting for the observed slope by which productivity dispersion falls with employment. Declining productivity dispersion by size raises the possibility of an aggregation effect. With such a mechanism, larger plants are hit by a larger number of productivity shocks arriving at each production unit within the plant. Assuming that these shocks are imperfectly correlated and that plant's performance is an employment weighted mean of all these shocks, by the law of large numbers, larger plants should suffer from less variance in their productivities and perform closer to their mean productivity[5].

For this argument to be applicable in practice, it is required that plant behavior be additive, i.e. it should be possible to think of a large plant as collection of several smaller plants bound together. On theoretical grounds, Fisher (1993) shows that this additivity holds only under very restrictive conditions on the production process. To elaborate, assume output is produced using labor and capital as the only two factors of production. In addition assume that labor is homogeneous across plants, while capital is plant-specific. Leontief states the condition for possibility of capital aggregation as:

Theorem 1 (Leontief) *Let K be a vector of N variables $K_1, K_2, \ldots, K_N$. Suppose $f(.,.)$ is a function of K and L, continuously differentiable with respect to K with $\partial f(.,.)/\partial K_1 > 0$. Then the following two conditions are equivalent:*

1. *There exist functions $g(.)$ and $h(.,.)$ such that $f(K, L) = h(g(K), L)$.*

2. *$\frac{\partial f(K,L)/\partial K_i}{\partial f(K,L)/\partial K_1}$ is independent of L for $i = 2, \ldots, N$.*

For N plants, the above theorem says that their operation can be replicated by a larger plant if the marginal rates of substitution among each capital input

[5]To be rigorous, no finite number of production unit should dominate the total production within a plant so that the law of large numbers holds in this case.

used by different plants are independent from the level of total employed labor. A special production function with such a property is an additively separable function in capital and labor such as

$$f(K, L) = \psi(K) + \phi(L), \tag{4.1}$$

and with the assumption on homogeneity of labor it should further be that $\phi(L)$ is linear in L. Note that the role of intangibles such as human capital or worker-to-job match quality is totally ignored in this setting. With those heterogeneities present, the possibility of aggregation will be under even more strain.

In case technologies are different, Fisher finds the condition for possible aggregation of N plants with production functions $f_i(.,.),\ i = 1, \ldots, N$ as

$$\frac{\frac{\partial f_i}{\partial K \partial L}}{\frac{\partial f_i}{\partial K}\ \frac{\partial^2 f_i}{\partial L^2}} = g\left(\frac{\partial f_i}{\partial L}\right), \tag{4.2}$$

where function g is the same for all plants. Notice that the functional form of (4.1) satisfies this condition when $g \equiv 0$. Condition (4.2) is not generally very intuitive, but under constant returns to scale production it can be interpreted in an economic context.

Theorem 2 (Fisher) *In a two factor, constant returns case, production can be aggregated if and only if all technical difference is capital augmenting.*

In other words, if differences among technologies and intangibles can be approximated by a capital augmenting effect, still there is some hope of production aggregation. However, this condition limits the range of technical diversities that can be present and modeled. The matter gets even worse when the production function is not constant returns to scale. Heterogeneity of products and labor would make the aggregation conditions even more complicated and restrictive. The bottom line is that an assumption of additivity of plant operation imposes very restrictive conditions on how plants perform and produce. In practice, these conditions are hardly ever satisfied and simulation of larger plants by bundling small units is not expected to be realistic.

Bartelsman & Dhrymes (1998) also demonstrate the fallacy of aggregation in an empirical light. They consider a measure of aggregate productivity obtained by summing up contributions of input and output. They show that movements of this productivity measure show large deviations from the mean of individual plant productivity in the period 1974 to 1984. In particular, aggregate productivity shows constant growth over this period while mean productivity is actually falling for much of that duration. Adding to that, I further cement the impracticality of aggregation by testing a model of statistical aggregation. As will be shown, the results suggest that productivity dispersions that are a result of statistical aggregation fall much faster with size than what is seen in the data and hit the bottom quickly, suggesting that statistical aggregation is not responsible for the observed pattern.

I investigate statistical aggregation by performing a counter-factual experiment in which I bootstrap distributions of larger artificial plants by aggregating actual plants of 1-19 employees, representing the smallest units of production. 100 bootstrapped distributions are created for each employment class. 95-5 inter-percentile range of log-productivity is used again to measure dispersion for both actual and simulated plants[6]. Results for both measures of productivity are shown in Figure 4.3. As the plots show, statistical aggregation results in a very steep decline in productivity dispersion as employment size gets larger[7]. Statistical aggregation basically drives productivity dispersion to zero too fast to be a convincing explanation

[6]The percentiles from the bootstrapped distribution are shown to be asymptotic to the actual ones (Hall 1992).

[7]This is not surprising as pure statistical aggregation predicts that the standard deviation must fall at a rate $1/\sqrt{l_1}$, with l_1 being the number of labor units employed at a plant

for the observed slope[8].

Using statistical aggregation to explain productivity dispersion also poses another problem. As will be shown later, the behavior of productivity dispersion by employment very much different across industries with dispersion declining both at positive and negative slopes for different industries. Statistical aggregation does not seem to be able to account for such wide range of differences.

4.4 Persistence of Size and Productivity

The dispersion of productivity at any given level of employment was shown above to be large and negatively related to the employment level. Is this dispersion at a given employment level caused by transitory short-run adjustment costs in labor or capital? Or is it that there are quite a large number of plants that permanently differ in their productivity and at the same employment level, causing the range of productivity differences observed in the data?

It is certain that plant-level employment and productivity are not constant over time. Productivity of a plant changes over time as a result of technology change and exogenous shocks. Plants are commonly thought to adjust their input sizes according to their observed productivity level: plants receiving favorable productivity outcomes grow and move out of a particular employment level, and plants getting hit by unfavorable outcomes downsize to move to a lower employment level. As a result, size dynamics are expected to follow productivity dynamics closely. How fast entry and exit into an employment level happen depends to a great extent on how fast productivity changes.

In answer to the above question, dynamics of productivity have been stud-

[8]The counter-factual curve in Figure 4.3(d) does not converge to near zero as the number of aggregated units goes up. This is perhaps due to presence of other productive factors in residual rLP, hence leaving some covariation between productivities. Assuming uniform weighting of shocks, the limit dispersion will be ζ in this case, where ζ is the covariance between idiosyncratic rLP's.

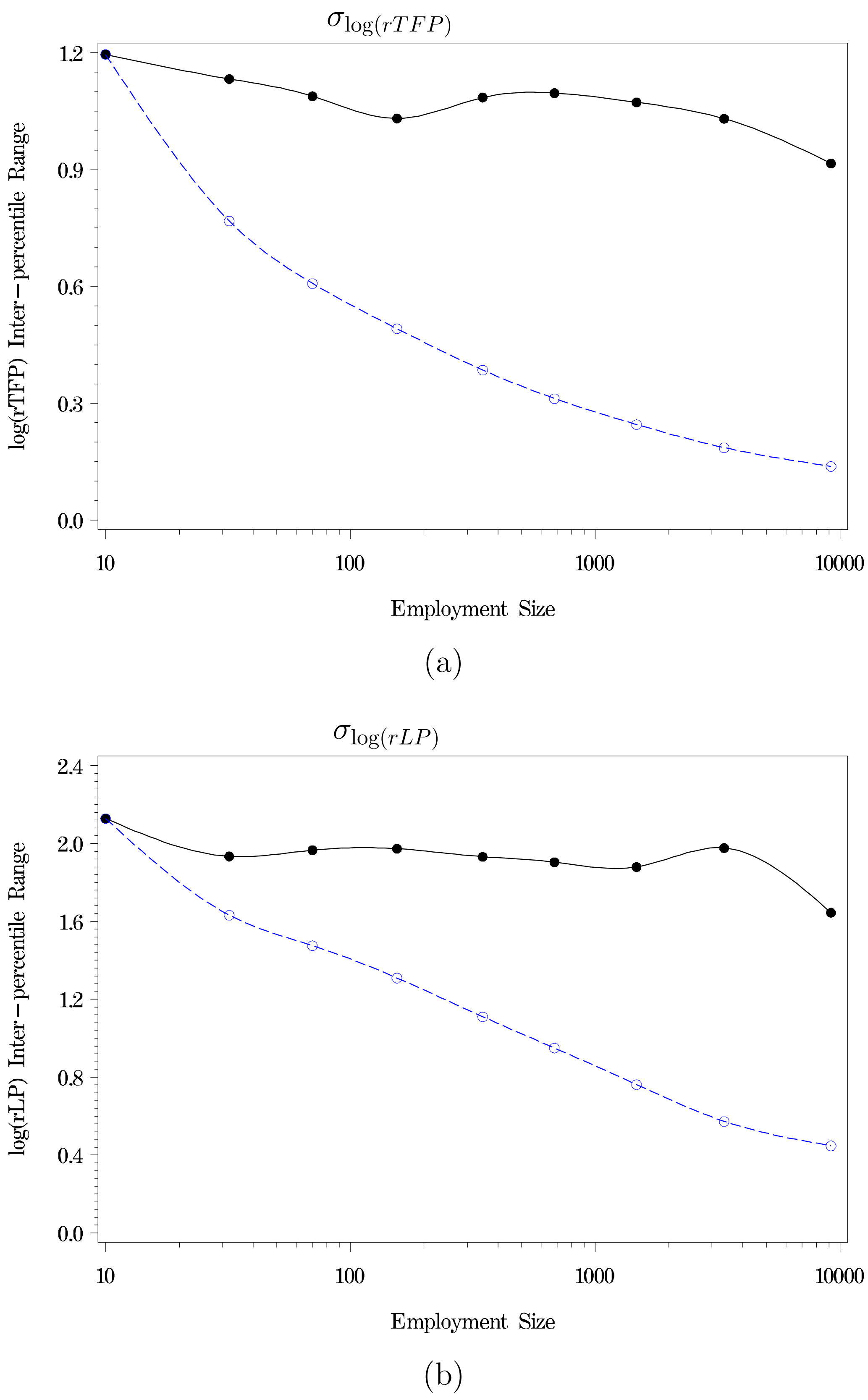


Figure 4.3: Comparing the actual dispersion of productivity by size (solid line) with the counter-factual one (dashed line).

ied by Baily et al. (1992) and Bartelsman & Dhrymes (1998) and shown to be a rather low frequency process pointing to slow and gradual changes of productivity. These works study the transition among the productivity ranks in the context of a discrete-time Markov process, with productivity rank normally defined as the quintile or decile to which a plant's productivity belongs. Given the estimated transition matrix, productivity rank reversals and jumps to non-neighboring productivity ranks are rarities, even in time periods as long as five or ten years. In a more recent work, Foster et al. (2008) use a simple one-lag auto-regressive model to estimate the persistence of productivity for continuing plants belonging to a subset of manufacturing industries. The novelty of their approach is that they show high persistence in physical productivity as well as in previously used revenue productivity. Their findings suggest a correlation coefficient of about 0.8 between plant-level physical productivities of two consecutive years. Additional results from Baily et al. (1992) show that the distribution of productivity is mostly time-invariant within four-digit manufacturing industries, indicative of a long-run distribution rather than a transient one.

It remains to see whether employment size is as persistent as productivity. My conjecture is that because of adjustment costs, size changes should be less frequent than productivity changes, so that employment should show more persistence than productivity. How much size persistence is really present? If employment is infrequently adjusted, then the distribution of productivity at any employment level is mostly composed of a steady subset of plants over fairly long periods of time. If employment size is adjusted rapidly, then a larger fraction of that distribution will consist of plants that are moving in and out. In turn, if most of the productivity distribution at a given employment level is caused by a steady set of plants, a long-run explanation for productivity dispersion and its relationship to employment size is warranted.

I explore persistence in employment and productivity by defining combined size-productivity states. Comparing dynamics of each variable is then easily achieved by looking at the implications of the combined dynamics along each dimension, be it productivity or size. For this experiment, I want to reduce the effect of industry selection among new plants, so that I can focus on the dynamic behavior of mature plants. Therefore, in what follows, I limit my set of plants to those that are at least 6 years old. This eliminate births and early age volatilities. Figure 4.1 suggests that the effect of such selection on the overall productivity dispersion is minimized.

Table 4.3 shows the distribution of plants that are older than 6 years in each census year, listed jointly by productivity and employment. Productivities are broken into three classes using quartiles of rTFP from all plants in the manufacturing sector[9]. Plants belonging to the top quartile are *high-productivity*, those belonging to the lower quartile are *low-productivity*, and those belonging to the middle half are *medium-productivity*. Likewise, employment is broken into three classes. However, different industries have different scales of employment. To make the scaling uniform, I divide employment sizes in each industry by the 90th percentile employment size in that industry[10]. Then I proceed by assigning the top quartile of this *normalized* employment as *large*, the bottom quartile as *small*, and the middle half as *medium-sized* plants.

A look at the distribution for different years confirms that the distribution of plants by employment and productivity is almost time-invariant. But more notably, the distribution does not suggest a monotonic relationship between employment and productivity. In most models, more productive plants are supposed to be larger in the long-run. Among old plants, this monotonicity seems to be present for

[9]As explained in the previous chapter, rTFPs are purged of their time and time by industry effects.

[10]Ideally, employment in each industry must be divided by the maximum employment size in that industry to make the scaling exactly uniform. But, in most industries, the maximum size is normally exceedingly large and an outlier. Dividing by the maximum employment size would cause serious distortion in scaling. Therefore I use 90th percentile employment size.

Size	Productivity	Distribution 1982	1987	1992	1997
	Low	0.109	0.098	0.088	0.089
Small	Medium	0.087	0.091	0.102	0.101
	High	0.053	0.059	0.059	0.056
	Low	0.138	0.149	0.154	0.155
Medium	Medium	0.247	0.242	0.234	0.238
	High	0.115	0.110	0.113	0.111
	Low	0.062	0.060	0.064	0.068
Large	Medium	0.126	0.129	0.127	0.119
	High	0.063	0.062	0.059	0.063
Sum of	Weights	114453	114713	134717	171206

Table 4.3: Distribution of plants by productivity and employment.

low- and medium-productivity plants. But monotonicity starts to break down for high-productivity plants, whose average size seems to be lower than their medium-productivity counterparts. Theoretically, I approach this paradox by emphasizing the role of demand structure, where growth in size is dependent on favorable demand conditions, which is potentially achieved through accumulation of enough intangible capital - e.g. managerial or marketing capital. Assuming that accumulation of such capital comes through a very slow, costly, or uncertain process, then even high-productivity plants may grow sluggishly. In the theoretical model of the next chapter, I will show the influence of market size by looking at market localization as a force that inevitably affects size growth in all plants whose products are traded locally.

To investigate the relative persistence of employment and productivity, I form discrete-time transition matrices for the above defined joint productivity-employment states between consecutive census years (five years apart). These matrices will feature transitions among 9 possible states plus exit. Entry is also included in the transition table and is defined as younger plants turning 6 years or older during the transition period. The availability of age estimates makes identifying entrants a trivial task. I use only the CM weighted sample to track plants longitudinally, because productivity measures are available and more reliable for those plants. But, care must be taken when forming the transition probabilities: The weighted panel changes every five years, so in two consecutive census years most of the smaller plants are replaced. To avoid including spurious exits, I use the full CM panel for the destination year. Plants are linked longitudinally by their PPN. Tables 4.4 to 4.6 show the estimated five-year transition probabilities starting from census years 1982, 1987, and 1992, respectively.

The exit rates mostly show a consistent pattern where the probability of exit falls with both the level of productivity and the level of employment in all the transition matrices. This pattern is in agreement with the findings Evans (1987) and Hall (1987). Dynamically, both productivity and employment size show much persistence. Looking at the transition probabilities, size persistence is implied by diagonal blocks having larger entries than non-diagonal blocks, especially among medium to high productivity plants. Also, size growth is slow and gradual as very few plants make the transition from small to large size or vice versa in a five-year period. The same thing can be said about productivity, where large productivity changes are not frequently seen in the data. At the same time, persistence of productivity is not perfect as the block matrices are far from being diagonally dominant[11]. This latter finding is consistent with the results of Baily et al. (1992) for manufacturing sector.

[11]Diagonal dominance means that the diagonal elements of the matrix are the strongest ones and for a square matrix $A = [a_{ij}]$ is formally defined $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for each diagonal element.

1982		1987									
		Small Size			Medium Size			Large Size			
Size	Productivity	Low	Medium	High	Low	Medium	High	Low	Medium	High	Exit
Small	Low	0.18	0.14	0.03	0.07	0.07	0.02	0	0.01	0	0.48
	Medium	0.16	0.24	0.05	0.10	0.18	0.03	0	0.01	0	0.23
	High	0.09	0.20	0.09	0.07	0.19	0.06	0.01	0.01	0	0.28
Medium	Low	0.12	0.05	0.02	0.21	0.24	0.06	0.04	0.05	0.01	0.20
	Medium	0.07	0.07	0.02	0.15	0.34	0.07	0.03	0.07	0.02	0.16
	High	0.06	0.07	0.06	0.10	0.26	0.17	0.02	0.05	0.05	0.16
Large	Low	0.01	0.02	0.01	0.16	0.14	0.04	0.21	0.19	0.06	0.16
	Medium	0.01	0.01	0	0.07	0.18	0.04	0.11	0.36	0.11	0.11
	High	0.01	0	0	0.06	0.11	0.08	0.08	0.20	0.31	0.15
	Entries	0.16	0.13	0.05	0.12	0.23	0.08	0.06	0.12	0.05	0

Table 4.4: Transition matrix of productivity-size from 1982 to 1987.

1987		1992									
		Small Size			Medium Size			Large Size			
Size	Productivity	Low	Medium	High	Low	Medium	High	Low	Medium	High	Exit
Small	Low	0.15	0.05	0.02	0.16	0.10	0.06	0.01	0.02	0	0.43
	Medium	0.10	0.11	0.02	0.19	0.21	0.07	0.01	0.02	0.01	0.26
	High	0.15	0.06	0.05	0.12	0.12	0.14	0.02	0.01	0.02	0.31
Medium	Low	0.04	0.03	0.01	0.22	0.15	0.10	0.09	0.06	0.02	0.28
	Medium	0.02	0.02	0.01	0.18	0.23	0.12	0.07	0.12	0.04	0.19
	High	0.01	0.04	0.03	0.15	0.17	0.21	0.04	0.08	0.10	0.17
Large	Low	0.02	0.01	0	0.09	0.11	0.05	0.27	0.19	0.07	0.19
	Medium	0	0	0	0.07	0.09	0.05	0.14	0.39	0.15	0.11
	High	0	0	0	0.05	0.06	0.06	0.10	0.22	0.39	0.12
	Entries	0.06	0.09	0.01	0.15	0.17	0.11	0.12	0.17	0.11	0

Table 4.5: Transition matrix of productivity-size from 1987 to 1992.

1992		1997									
		Small Size			Medium Size			Large Size			
Size	Productivity	Low	Medium	High	Low	Medium	High	Low	Medium	High	Exit
Small	Low	0.11	0.12	0.02	0.13	0.11	0.02	0.01	0.01	0	0.47
	Medium	0.10	0.18	0.03	0.12	0.16	0.05	0.02	0.01	0	0.33
	High	0.12	0.20	0.05	0.12	0.13	0.10	0.01	0.01	0.01	0.25
Medium	Low	0.04	0.08	0.01	0.18	0.18	0.08	0.08	0.05	0.02	0.28
	Medium	0.03	0.04	0.01	0.20	0.23	0.08	0.08	0.11	0.04	0.18
	High	0.05	0.05	0.02	0.18	0.17	0.13	0.06	0.08	0.07	0.19
Large	Low	0.01	0	0	0.10	0.08	0.04	0.31	0.19	0.06	0.21
	Medium	0.01	0.01	0	0.07	0.09	0.03	0.21	0.34	0.11	0.13
	High	0	0	0	0.07	0.06	0.05	0.17	0.19	0.31	0.15
	Entries	0.08	0.08	0.02	0.21	0.20	0.10	0.19	0.08	0.04	0

Table 4.6: Transition matrix of productivity-size from 1992 to 1997.

A more detailed comparison is possible when I invoke properties of discrete-time Markov chain to compute $T[j|i] = [T_{ij}]$, i.e. the expected number of years a plant spends in a size-productivity state j before exit conditional on having started from state i [12]. The advantage of using matrix T is that it provides an intuitive way of understanding the speed and path of change for productivity or employment or both together by expressing quantities in the better interpretable form of time lengths. A shortcoming of this method is that it does not distinguish between the time spent in one long spell or several shorter spells. But, at the same time, this method overcomes short-period transitory shocks and helps to interpret the dynamics of size and productivity in a more long-term context. The shortcoming is also alleviated by the fact that transition matrices show the dynamics of size and productivity to be slow, reducing the possibility of many spells. With this issues in mind, the computed matrix for this example is

$$T[j|i] = \begin{pmatrix} 6.45 & 1.23 & 0.32 & 2.26 & 2.48 & 1.02 & 0.93 & 1.34 & 0.64 \\ 1.55 & 6.87 & 0.44 & 3.01 & 3.79 & 1.46 & 1.33 & 1.89 & 0.95 \\ 1.54 & 1.75 & 5.60 & 2.80 & 3.56 & 1.73 & 1.34 & 1.86 & 1.01 \\ 1.20 & 1.14 & 0.33 & 8.53 & 4.06 & 1.75 & 2.02 & 2.62 & 1.30 \\ 1.12 & 1.15 & 0.35 & 3.71 & 9.88 & 1.99 & 2.25 & 3.34 & 1.66 \\ 1.14 & 1.23 & 0.49 & 3.51 & 4.52 & 7.44 & 2.10 & 3.12 & 1.93 \\ 0.86 & 0.83 & 0.25 & 3.34 & 3.97 & 1.71 & 8.81 & 4.44 & 2.15 \\ 0.86 & 0.85 & 0.25 & 3.33 & 4.39 & 1.85 & 3.56 & 11.06 & 2.91 \\ 0.80 & 0.77 & 0.24 & 3.15 & 4.00 & 1.97 & 3.29 & 5.10 & 9.31 \end{pmatrix}. \tag{4.3}$$

The order of states is exactly as in the transition matrices (Tables 4.4-4.6). For instance, the row 4, column 3 element provides the expected number of periods a

[12] It can be shown that $T = 5 \times (I - S)^{-1}$, where I is the identity matrix, and S is the transition matrix without the entry row and exit column. The multiplier 5 changes the results into annual. I use the mean average of the three transition matrices for years 1982, 1987, and 1992 as S. More details on the derivation of this formula can be found in Kemeny & Snell (1983).

plant starting as a low-productivity medium-sized will spend as a high-productivity small plant before exiting the market, which is about 4 months from the data. The persistence of each productivity-employment state readily manifests itself in the relatively large diagonal elements in (4.3), though the matrix is not diagonally dominant. For a more detailed analysis, I look at persistence of size and productivity separately. To look at persistence of size, I add up the expected length of time a plant stays small, medium-sized, or large by each row of (4.3) to get

$$T[Size|i] = \begin{array}{c} \underline{Size} \\ Small \\ \\ Medium \\ \\ Large \end{array} \begin{pmatrix} Small & Medium & Large \\ 7.99 & 5.77 & 2.92 \\ 8.86 & 8.27 & 4.17 \\ 8.89 & 8.09 & 4.21 \\ \cdots & \cdots & \cdots \\ 2.66 & 14.33 & 5.94 \\ 2.62 & 15.58 & 7.25 \\ 2.86 & 15.47 & 7.15 \\ \cdots & \cdots & \cdots \\ 1.95 & 9.01 & 15.40 \\ 1.96 & 9.57 & 17.53 \\ 1.80 & 9.12 & 17.69 \end{pmatrix}. \tag{4.4}$$

Note that, as before, the first three rows are small plants of different productivities, the next three are medium-sized, and the last three are large plants. For instance, row 4 says that a low-productivity medium-sized plant will spend medium-sized for an expected period of 14 years prior to exit, while it is expected to be small or large for about 3 and 6 years, respectively. The total expected lifetime for this plant is about 23 years, adding up the three numbers. With this notion, persistence of size seems to be much stronger at larger employments, where plants spend as large-sized for an average of 17, which is almost 60% of their expected lifetime. But even small plants spend eight to nine years with their small size, which is slightly less than 50% of their expected lifetime. Remember that only old plants are kept in the data, therefore the implications refer to long-run behavior of plants and not the industry selection.

For comparison, I will look at the dynamics of productivity in the same way. Adding the expected number of periods by productivity class yields

$$T[Prod.|i] = \begin{array}{c} \underline{Prod.} \\ Low \\ \\ Medium \\ \\ High \end{array} \begin{pmatrix} Low & Medium & High \\ 9.64 & 5.06 & 1.98 \\ 11.75 & 7.81 & 3.38 \\ 13.00 & 9.25 & 4.11 \\ \cdots & \cdots & \cdots \\ 5.89 & 12.55 & 2.85 \\ 7.09 & 14.36 & 3.99 \\ 7.75 & 16.30 & 5.01 \\ \cdots & \cdots & \cdots \\ 5.67 & 7.17 & 8.34 \\ 6.75 & 8.87 & 9.85 \\ 7.24 & 9.86 & 11.51 \end{pmatrix} \tag{4.5}$$

In this case, for easier visualization, I rearranged states so that now the first three rows are low-productivity plants of different sizes, the next three are medium-productivity, and the last three are high productivity plants. Sizes are sorted from small to large within each productivity class. For example row 4 here says that a small medium-productivity plant spends about 12 years in the same productivity class, while its productivity will be low or high for expected periods of 6 and 3 years, respectively.

The level of persistence in productivity demonstrated in (4.5) is seemingly high and mostly comparable to the level of persistence observed in employment size. High-productivity plants show less persistence in productivity than in their size as they spend less than 50% of their expected lifetime as high-productivity. The role is reversed in low-productivity plants, which now spend more than 50% of their expected lifetime as low-productivity.

My overall conclusion is that both productivity and employment size demonstrate fairly high levels of persistence within plants older than 6 years, underlining the fact that, in the short to medium run, both productivity and employment can be thought of as almost constant. In light of this evidence, the distribution of productivity observed at each employment level can be regarded as created by a fixed set of plants in fairly long periods of time and treated as resulting from long-run behavior. The next sections of this chapter rely on this conclusion and examine sector-specific factors that shape the distribution of productivity at different levels of employment. In particular, I study whether market localization can generate long-run plant-level behavior consistent with the distribution already seen in Table 4.3.

4.5 A Cross-Industry Analysis

So far, I have focussed on the behavior of productivity dispersion in manufacturing as an aggregate industry. However, given cross-industry differences in technology and size, it would be interesting to know if there are differences in the behavior of productivity dispersion across four-digit industries, and if these differences can be related to industry characterizations. To measure how productivity dispersion relates to the level of employment within a four-digit industry, I construct the following sector-specific statistics. For industry i I define:

$$r_i = \Delta_{Large}/\Delta_{Small}, \tag{4.6}$$

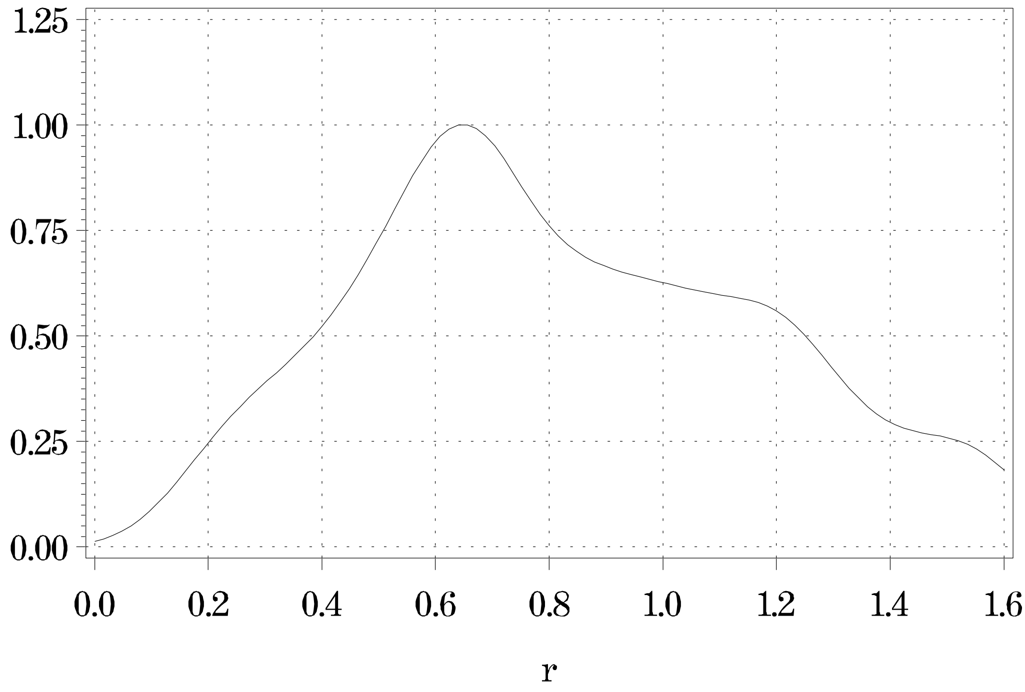
where, as before, Δ represents 95-5 inter-percentile range of productivity, for either rTFP or rLP. *Large* refers to plants in the upper employment decile of their industry. Similarly, *Small* refers to plants in the lower employment decile of their industry. Very small values of r_i correspond to industries for which larger plants exhibit much less productivity dispersion than small plants. As r_i grows towards 1, productivity dispersion is expected to level out across small and large plants.

I compute r_i for 397 four-digit manufacturing industries, using both rTFP and rLP measures. For more comparability, the productivity measures are purged of year effects within each four-digit industry. To visualize the shape and extent of heterogeneity of r among industries, I compute the KDE estimate of the obtained r_i values. The estimated distributions with rTFP and rLP are plotted in Figure 4.4. As pictures show, industries are very diverse in how their productivity dispersion relates to the level of employment, with r ratios ranging from close to zero to about 1.6. About 75% of all industries have an r ratio less than 1, indicating that falling dispersion by size is rather common among manufacturing industries but not quite universal.

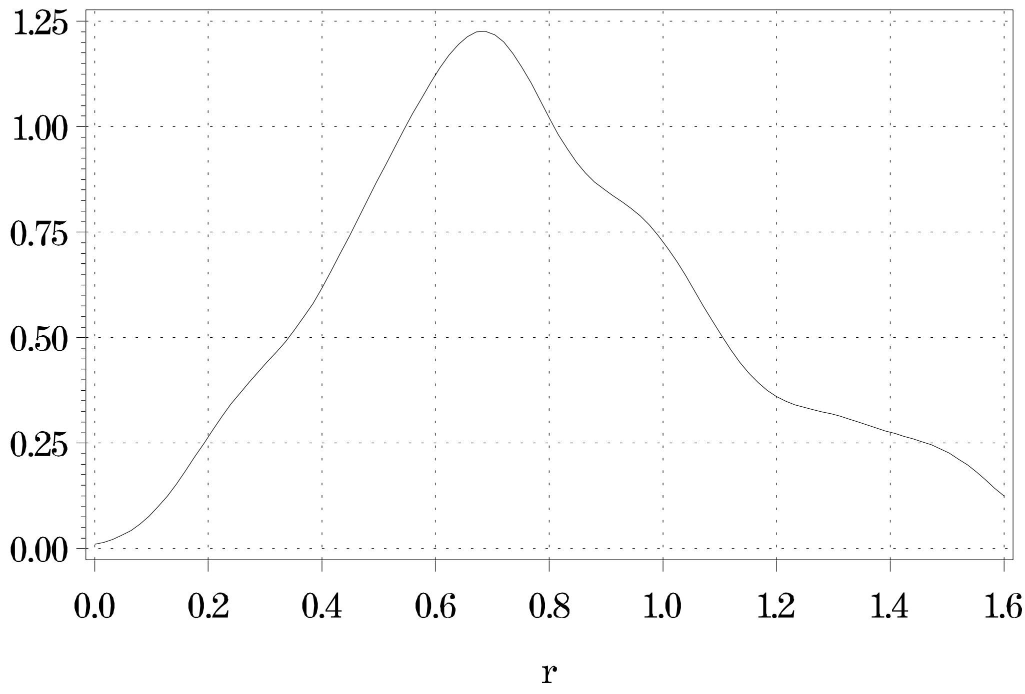
4.6 Role of Market Localization

Different sector-specific factors affect large and small plants in different ways, potentially leading to long-run differences in the productivity dispersions of large and small plants. In order to explain the pattern of productivity dispersion discussed earlier in this chapter, I look at several sector-specific factors, and especially market localization.

As described in the model below, market localization implies the possibility of high-productivity small plants due to limited demand. These plants will contribute to an increased productivity dispersion at smaller employment sizes. The idea that



(a) Using rTFP



(b) Using rLP

Figure 4.4: KDE plot of r_i.

small market size may limit employment is echoed in the work of Campbell & Hopenhayn (2002) in the case of retail industry. In contrast, it is reasonable to think that industries with access to national or international markets, through trade or as a result of low transportation costs, can benefit from serving a wide array of demand markets, so that they are less constrained in choosing their employment and output size. For example, the Honda LLC plant in Lincoln, Alabama, USA (with population of 4577) targets a national market rather than the local one; therefore, its size and scale of operation does not depend on its local market at all.

I estimate a model specifying the r ratio as a function of several sector-specific factors that are potentially important determinants of the slope of the relationship between productivity dispersion and the level of employment. The model has the general form

$$r_i = \beta_0 + D_i B + X_i C + \epsilon_i, \tag{4.7}$$

where D_i is a vector of dummies that classify the degree of market localization in industry i, and X_i is a vector of other industry-specific controls which will be described below in more detail.

To measure market localization for each sector, I use the US Transportation Department's 1997 Commodity Flow Survey (CFS). The survey provides information on value, tonnage, ton-miles shipped and average distance shipped for disaggregate commodities. Two different measures of market localization are extracted from the survey: Average shipment distance in miles (DISTANCE) and value-per-ton shipped in dollars (VALUE/TON). Commodities that are shipped shorter distances on average are likely to be sold in local markets. VALUE/TON gives a cruder measure of market localization; in the presence of transportation costs, commodities whose value-per-ton is lower are more likely to be shipped locally. The commodity descriptions are matched to Standard Industry Codes (SIC) as closely as possible using US Dept. of Labor descriptions[13].

To avoid possible match quality problems, each measure of market localization is broken into classes. With shipment distance (DISTANCE) the classes are 0-100 (D_{1i}), 100-300 (D_{2i}), with 300+ as the control group. When using value per ton (VALUE/TON) as measure of localization, the classes are 0-500 (D_{1i}), 500-2000 (D_{2i}), with 2000+ as control group.

Another influence on an industry's market localization is the amount of exposure to international trade. More export intensive industries have access to larger markets. In this situation, lower productivity plants specialize in domestic markets, while larger more productive plants expand their operation in response to trade possibilities (Melitz 2003). I control for EXPINT defined as the ratio of the value of industry exports to its output. Larger import penetration also signals greater trade exposure. The variable IMPPEN is the ratio of the value of industry imports to the sum of imports and the industry output. Both of these data are described in more detail by Feenstra (1996) and Feenstra (1997). Both of these variables are included as industry control variables X_i's.

Other industry specific controls include variables that may affect small and large plants differently, and therefore may increase (or decrease) the dispersion gap between small and large plants. While I cannot control for all possible factors, I include a variety of variables in X_i that might be important. Due to availability of data, 1987 data is used for all controls. CLUSTER is an index that measures how much the industry is geographically concentrated. Clustering is often associated with knowledge or technology spillovers. In clustered industries, the technology and experience of larger plants can quickly diffuse to smaller plants who are getting free ride on this pool of knowledge. This effect can level out productivity and its dispersion across plants irregardless of size. At the same time, by being able to

[13]Available at `http://www.osha.gov/pls/imis/sic_manual.html` .

offer higher wages, larger plants can steal high-skill workers from smaller plants, adding to the volatility of productivity at small plants. Depending on which one dominates, we can see a positive or negative coefficient. I use the Ellison & Glaeser (1997) index of concentration, which combines both natural advantage and spillover effects of each geographical location into one measure[14]. As a robustness check, the R&D intensity of industries is also included. The variable RANDD measures the ratio of R&D expenditure to total industry output. The data is taken from the 1987 NSF report. The idea is that in R&D intensive industries, plants are more vulnerable to knowledge spillovers but can also depart as a result of their investing behavior in new technologies. In this way, RANDD should be able to account for some of the clustering effect.

When industries produce more diversified products, larger plants normally increase their scale of operation not only by increasing their output, but by introducing new varieties. Conversely, smaller plants specialize in just one variety, leaving them more vulnerable to demand and taste shocks. Also in such industries, plants can respond to demand shocks by changing their production variety. In this process, introduction of each new product can be thought as a new entry with its sunk cost (Bernard, Redding & Schott 2006), better afforded by larger plants. The index DIVINDX measures the product diversity within an industry. I use the diversity index defined by Gollop & Monahan (1991). This index has the advantage that it accounts for diversity not only by looking at the number of different products in an industry, but by how different the products are and how unequal the distribution of products is across production lines. The PPC product code from the CM is used here to distinguish different products in each industry.

Fixed costs of operation can also shape the distribution of productivity by directly controlling the cutoff productivity - i.e. the lowest productivity plant that

[14] They list a table of clustering measure by 4 digit SIC code in their NBER working paper

Variable	Mean	Std.Dev.	Min.	Median	Max.
rLP r	1.004	1.316	0.147	0.763	22.33
rTFP r	1.173	1.090	0.047	0.888	9.488
DISTANCE (miles)	479.9	288.6	31	393	1090
VALUE/TON ($\times$\$1000)	12.3	28.5	0.006	3.023	218.3
EXPINT	0.096	0.128	0.0001	0.050	1.052
IMPPEN	0.150	0.158	2e-5	0.101	0.886
FIXEDCOST	0.273	0.113	0.043	0.251	0.818
SUNKCOST	0.003	0.007	3e-6	0.0008	0.073
CLUSTER	0.051	0.071	-0.013	0.027	0.480
DIVINDX	0.490	0.320	0	0.569	1.046
RANDD	2.619	2.518	0.400	1.200	7.500

Table 4.7: Summary statistics on regressor variables.

can survive and continue producing. I use the variable FIXEDCOST, defined in the same way as Syverson (2003) as the ratio of non-production employment to total employment. Non-production workers pose an overhead cost to the plant that is paid every period. Sunk costs also may vary by market size and affect decisions to introduce new product varieties. SUNKCOST here is measured in the same way as in Sutton (1991). This measure is the output share of the median plant[15] multiplied by the ratio of capital to output in an industry, with the median plant representing the minimum efficient scale of production.

Table 4.7 lists summary statistics for the dependent and independent variables.

The primary coefficients of interest are those on the dummies D_1 and D_2, cor-

[15]I use the mean output value of plants belonging to the 49th to 51st employment percentile of an industry and divide it by total output in that industry.

responding to industries with more localized markets. From earlier discussion, it is expected that industries subject to localized markets will display a steeper decline of productivity dispersion with employment size. This hypothesis can be put to test by observing the ranking among the estimated coefficients for the D_1 and D_2 dummies. Remember that D_1 indicated the most localized industries, while D_2 indicated industries whose market reach is farther. The control group is all the industries whose markets are more or less globalized. Let B_1 and B_2 be the estimated coefficients on D_1 and D_2 in each regression, respectively. Table 4.8 lists the estimated model for market localization with and without other industry controls. The results are estimated using both rTFP and rLP as measures of productivity. The first column is the estimated model without any sector-specific controls, using DISTANCE as the measure of localization. Column two shows the estimated coefficients for the full model. The role of market localization reflects itself in negative estimated values for B_1 and B_2. The slope by which productivity dispersion falls with employment is found to be steeper when market reach of the industry is less than 300 miles.

Table 4.9 lists the same model but using VALUE/TON as the measure of market localization. The model is still able to produce supportive results, though the results are mixed. In this case, estimates are statistically less significant, and some of the estimated values for B_2 are positive. This is probably because VALUE/TON is not as accurate as DISTANCE in specifying the extent of market reach.

I am also interested in inspecting the values of $B_1 - B_2$ to see if a monotonic rank ordering exists for different degrees of market localization, i.e. if more localization means a steeper negative slope of the productivity dispersion curve. With $B_1 < 0$ and $B_2 < 0$, a negative value for $B_1 - B_2$ suggests such a monotonic rank ordering. Table 4.10 lists the estimated values and their statistics for each of the two localization measures. The estimated differences are all negative, with most of them statistically significant. Together, these two measures of market localization

	Dependent variable : r_{rLP}		Dependent variable : r_{rTFP}	
	(1)	(2)	(1)	(2)
B_1	-0.377 (0.222)**	-0.278 (0.231)	-0.396 (0.172)**	-0.529 (0.179)**
B_2	-0.348 (0.170)**	-0.212 (0.177)	-0.255 (0.132)**	-0.323 (0.179)**
EXPINT		-0.259 (0.629)		-0.323 (0.137)
IMPPEN		1.707 (0.454)**		0.748 (0.364)**
FIXEDCOST		-0.363 (0.698)		1.498 (0.542)**
SUNKCOST		-1.591 (9.302)		4.875 (7.219)
CLUSTER		0.277 (0.996)		-0.652 (0.772)
DIVINDX		-0.184 (0.219)		-0.229 (0.171)
RANDD		0.017 (0.033)		-0.073 (0.026)**
R^2	0.016	0.060	0.020	0.061

(*) significance with $P < 0.1$. (**) significance with $P < 0.05$.
One-tailed test used for estimated coefficients for D_1 and D_2, two-tailed test used for the rest.

Table 4.8: Table of coefficients for model (4.6) using average shipment distance as measure of market localization. Standard deviations appear in the parenthesis.

	Dependent variable : r_{rLP}		Dependent variable : r_{rTFP}	
	(1)	(2)	(1)	(2)
B_1	-0.279	-0.089	-0.351	-0.428
	(0.199)*	(0.210)	(0.154)**	(0.165)**
B_2	0.107	0.406	-0.047	-0.103
	(0.160)	(0.176)**	(0.124)	(0.140)
EXPINT		-0.302		0.133
		(0.633)		(0.499)
IMPPEN		1.989		0.766
		(0.460)**		(0.376)**
FIXEDCOST		-0.681		1.237
		(0.688)		(0.541)**
SUNKCOST		-4.165		5.085
		(9.314)		(7.316)
CLUSTER		0.723		-0.713
		(1.005)		(0.790)
DIVINDX		-0.173		-0.217
		(0.216)		(0.171)
RANDD		0.056		-0.063
		(0.034)		(0.027)**
R^2	0.008	0.072	0.014	0.049

(*) significance with $P < 0.1$. (**) significance with $P < 0.05$.
One-tailed test used for estimated coefficients for D_1 and D_2, two-tailed test used for the rest.

Table 4.9: Table of coefficients for model (4.6) using value per ton shipped as measure of market localization. Standard deviations appear in the parenthesis.

Localization Measure	Coef.	LP r model (1)	LP r model (2)	TFP r model (1)	TFP r model (2)
DISTANCE	$B_1 - B_2$	-0.029	-0.067	-0.141	-0.207
		(0.255)	(0.255)	(0.198)	(0.198)**
VALUE/TON	$B_1 - B_2$	-0.387	-0.494	-0.304	-0.325
		(0.223)**	(0.223)**	(0.173)**	(0.176)**

(*) significance with $P < 0.1$. (**) significance with $P < 0.05$.

Table 4.10: Difference in estimated coefficients to test if the degree of market localization ranks r_i.

draw a picture where markets are ranked based on their level of trade: Among plants operating in markets with less possibility of trade, the fall in productivity dispersion from small to large plants is much deeper.

In addition to DISTANCE and VALUE/TON, I control in some specifications for EXPINT and IMPPEN, both measuring trade exposure. Higher trade exposure means larger markets and less constraint on the conduct of plants. In this case, the gap between productivity dispersions of small and large plants should get narrower. The estimated coefficients on IMPPEN, especially, are positive and significant, suggesting that the slope of productivity dispersion becomes less steep as intensity of trade increases.

Other industry controls produce less significant estimates, however, some interesting results are displayed. Coefficients on fixed and sunk costs are more significant with r_{rTFP} as dependent variable. The estimated positive coefficients in this case show that within industries that face higher entry or overhead costs, productivity dispersion falls more slowly with employment. In support of this result note that with higher sunk and fixed costs, only the more efficient plants will have the incentive to enter the market, causing the dispersion of productivity to be almost the same even after industry selection has taken place.

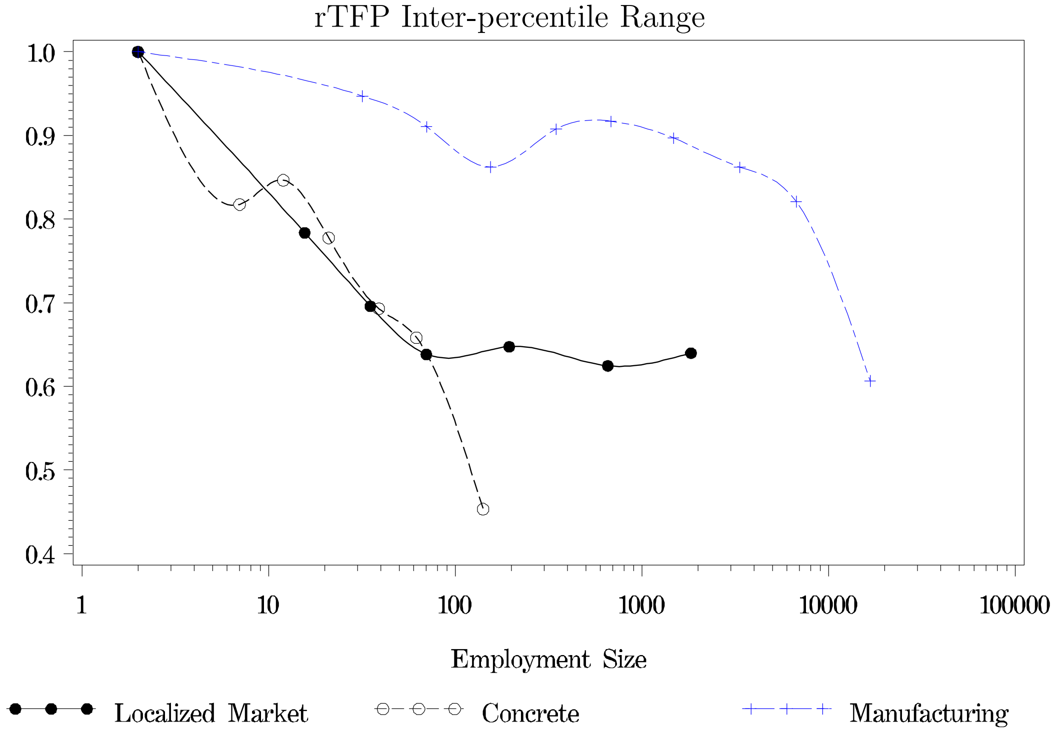
The estimated coefficient for CLUSTER is not significant and changes sign when using rLP or rTFP. However, RANDD quantified some of the effects of industry concentration. The significant coefficient estimated on RAND is corresponding to r_{rTFP} as dependent variable and is negative. It can be argued that in research intensive industries, the large plants generally invest in several new technologies at the same time, or acquire them by buying smaller firms, causing them to be more productive on average. Smaller plants, on the other hand, take more risks by investing in only one technology and can be very diverse in their productivities as a result. This difference in behavior can increase the slope by which productivity dispersion declines from small to large plants, which seems to overtake the industry clustering effect.

Finally, the estimated coefficient on DIVINDX is negative and significant with r_{rTFP} as dependent variable. This is the case where higher product diversity differentiates among small and large plants, giving the high productivity plants the chance to diversify their output in response to demand shocks, therefore, avoid large "revenue" productivity shocks. Smaller plants that generally specialize in production of one variety will be more affected by demand shocks and will be more dispersed. The negative estimated coefficient on this variable signifies this fact.

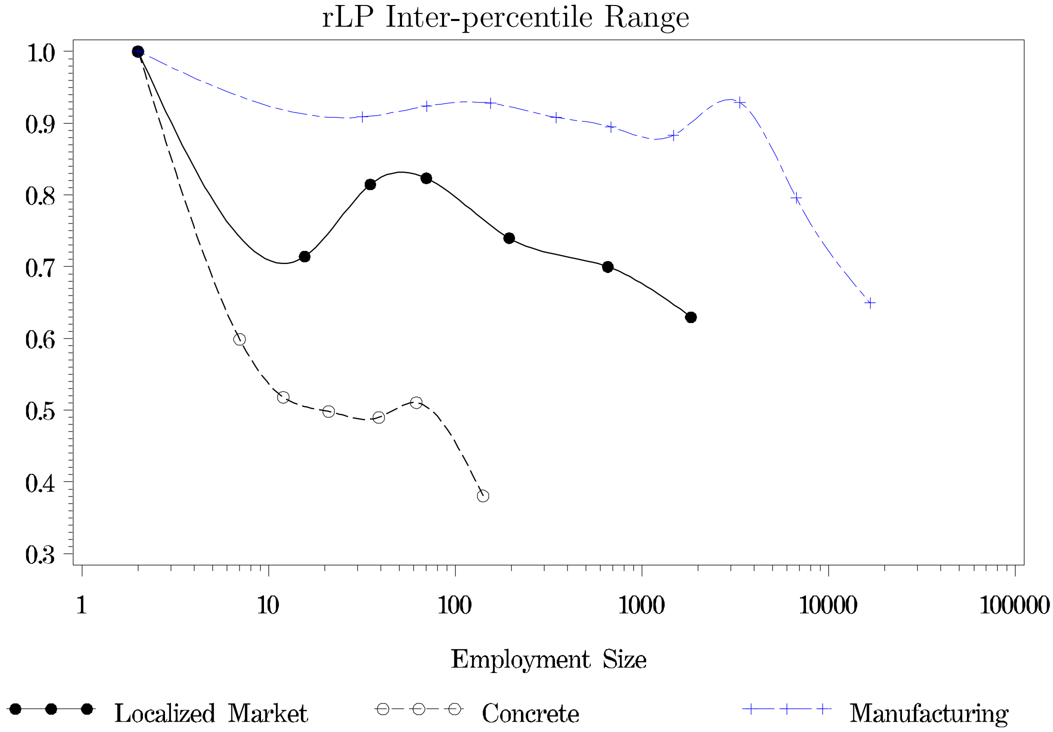
To visualize this "localization effect" I recompute the array of productivity dispersions by employment size within different groups of plants. Using CFS, I choose my localized-market industries as those who shipped their products on average no more than 100 miles away. 18 industries are selected this way whose list can be found in Appendix A. Figure 4.5 illustrates the behavior of productivity dispersion in localized-market industries using both rTFP and rLP. For plotting purposes, plants are again grouped into employment classes, and the 95-5 inter-percentile range was used to measure the range of productivities in each employment class while eliminating outliers[16]. For comparison, the productivity dispersion curve for aggregate manufacturing is also plotted along. All curves are normalized to start

[16]Employment size classes are again 1-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000-2499, 2500-4999, 5000-9999 and 10000+.

from 1, so that slopes can be compared. As pictures show, productivity dispersion falls more monotonically and at a considerable rate with localized-market industries. Again both plots with rTFP and rLP agree on this picture. Also, for future reference, Ready-Mix Concrete (SIC 3273) is singled out and shown on the plots as a particularly interesting industry and the benchmark for the empirical tests of my forthcoming model.



(a)



(b)

Figure 4.5: The behavior of productivity dispersion by employment size for localized-market industries.

Chapter 5

A Model of Market Localization

From the analysis of previous chapter, market localization is correlated with the magnitude of declining productivity dispersion by employment size. This chapter introduces a theoretical framework that can support long-run productivity dispersion at any input size, in which productivity dispersion falls with input size as observed in the data.

5.1 The Theoretical Model

The theoretical framework used here is the same as in Syverson (2003) and Melitz & Ottaviano (2005). The theory is based on the differentiated product model developed by Dixit & Stiglitz (1977). Plants are assumed to operate in localized markets and to have monopoly power over their demand. The model has several advantages for my analysis. First, market size and the elasticity of demand can be incorporated into the model easily through the choice of utility function. Also, the model is static and therefore tractable. Since I am interested in productivity dispersion as a long-run equilibrium phenomenon, not as a transient process, a long-run model is both simpler and more appropriate than a model with dynamics.

I add to the model a general single input constant or decreasing returns production function. This production function creates a connection between productivity, output and input size. The analysis will be more complicated than the regular framework, where measured productivity is summarized in production cost. The payoff is that I obtain strong results. Most importantly, the relationship between plant productivity and its input size is bell-shaped. This result, in turn, offers an explanation for a declining pattern of productivity dispersion by input size.

5.1.1 Consumers

A market is composed of L identical consumers. There is a continuum of producers, each producing a distinct variety of product indexed by j. The set of available products in each market is J, which is a subset of total possible varieties J^*. Let N be the measure of set J [1]. As in Syverson (2003), the representative consumer's utility function is

$$\begin{aligned} U &= y + \alpha \int_J q_j^c dj - \frac{1}{2}\eta \left(\int_J q_j^c dj \right)^2 - \frac{1}{2}\gamma \int_J (q_j^c)^2 dj \\ &= y + \alpha \int_J q_j^c dj - \frac{1}{2} \left(\eta + \frac{\gamma}{N} \right) \left(\int_J q_j^c dj \right)^2 - \frac{1}{2}\gamma \int_J (q_j^c - \bar{q})^2 dj, \end{aligned} \tag{5.1}$$

where y is the numeraire consumption, q_j^c is the consumption of each variety by the representative consumer, and $\bar{q}^c = \frac{1}{N}\int_J q_j^c dj$. The utility function parameters α, η, and γ are all non-negative. The utility function has a general quadratic form with parameters α and η determining the consumption of each variety relative to numeraire and γ determining the degree of distinction between varieties. For $\gamma = 0$, there is no variety distinction and the consumer cares only about the aggregate consumption. For higher values of γ, the consumer gains utility by smoothing consumption across different varieties. Compared to other utility functions in the differentiated products literature, the utility function of (5.1) has the advantage

[1] Equivalently, N will be a measure of plants operating in the market, which will be determined endogenously by the equilibrium conditions discussed later.

that it generates a linear demand curve whose elasticity varies by market size, and hence is more suitable for my analysis.

Utility maximization yields the following inverse demand curve per consumer for each variety

$$p_j = \alpha - \eta N\bar{q} - \gamma q_j^c. \tag{5.2}$$

Total demand for each variety is $q_j = Lq_j^c$, so each plant faces the following demand curve

$$p_j = \frac{\alpha\gamma + \eta N\bar{p}}{\gamma + \eta N} - \frac{\gamma}{L} q_j, \tag{5.3}$$

where $\bar{p} = \frac{1}{N}\int_J p_j dj$. Note that (5.2) requires that all prices be bounded above by α; consequently, we will have $\bar{p} \leq \alpha$. The ratio L/γ in (5.3) will have a direct controlling effect on the elasticity of demand and many of the theoretical results discussed in the next sections.

5.1.2 Producers

Plants produce distinct products facing the demand curve (5.3). Upon entry, each plant makes a random draw of its productivity ϕ_j from a known cumulative distribution $G(\phi)$ with support $\phi \in [0, \phi_M]$. There is also a continuous distribution of markets with different sizes. Plants incorporate randomly in a particular market. The cost of trading with other markets is assumed to be infinity; thus plants do not face the possibility of accessing other markets to broaden their demand.

Plants use a single composite input factor x_j for production and choose their input size optimally to maximize profit[2]. The rental price of input w is assumed exogenous and constant within each market and also across markets. Section 5.5.1 examines the effect of price variations across markets and shows that most of the model implications are robust to even a large amount of price variation. Plants

[2] x_j summarizes the contributions of several production factors in a single quantity.

produce according to a production function of the form

$$q_j = \phi_j x_j^{\nu}, \tag{5.4}$$

where ν is assumed to be a fixed value in the range (0,1]. A revenue-based measure of productivity will be useful later to compare the theoretical results to their empirical counterparts and is defined as

$$\theta_j = \frac{p_j q_j}{x_j}, \tag{5.5}$$

Each plant faces a profit function of the form

$$\pi_j = p_j(q_j)q_j - wx_j - f, \tag{5.6}$$

where f is the fixed cost of operation, which is the same for all plants and all markets. Plants are profit maximizing, and the quantity of output that maximizes profit for each plant is the solution to the first order condition

$$\frac{2\gamma}{L}q_j + \frac{w}{\nu\phi_j^{1/\nu}}q_j^{\frac{1}{\nu}-1} = \frac{\alpha\gamma + \eta N\bar{p}}{\gamma + \eta N}. \tag{5.7}$$

The solution to (5.7) is not trivial in general due to non-linearity. However, the set of possibilities can be narrowed down to simplify further analysis.

Proposition 1 *There exists a unique positive solution to* (5.7).

All proofs are in Appendix B. In the coming sections, I assume that q_j is the unique optimal output produced by each plant j.

5.1.3 Free Entry Equilibrium

In equilibrium, plants must be indifferent between entering the market or staying out. If the fixed cost of entering the market is f_E, then the equilibrium requires that the expected profit be equal to this entry cost to prevent an influx of new entry. In presence of fixed costs, only plants operating above a certain cutoff productivity ϕ_* will be profitable and will stay in the market. Put formally, it must be that

$$\int_{\phi_*}^{\phi_M} \pi_j(\phi) dG(\phi) = f_E. \tag{5.8}$$

Plants operating at the cutoff productivity ϕ_* are earning zero profit, i.e.,

$$\pi_j(\phi_*) = 0. \tag{5.9}$$

The equilibrium conditions (5.8) and (5.9) together with (5.6) and (5.7) determine an implicit relation between cutoff productivity ϕ_* and model parameters. As one observation, note that ϕ_* is always less than ϕ_M, since $\phi_* = \phi_M$ is a clear contradiction to (5.8) when $f_E > 0$.

Finally, N can be determined endogenously when ϕ_* is known. Finding a closed form solution for N has proved to be difficult, although this does not limit my ability to assess the model's implications for productivity dispersion. Later, in Section 5.4, I look at a CRTS production function under which I have closed form solutions for all endogenous variables in terms of parameters. For that reason, I defer further discussion of N to that section.

5.1.4 Analytical Results

In this section I seek to describe the plant behavior within a market under the assumption that markets are localized. Comparative statics are also presented that define the distribution of input size and productivity within each market and across markets. For now, without loss of generality, I focus on a single market.

Taking the partial derivatives of (5.7) with respect to ϕ_j provides the first general result

$$\frac{\partial q_j}{\partial \phi_j} = \frac{Lwq_j^{1/\nu-1}}{2\gamma\nu^2\phi_j^{1/\nu-1} + (1-\nu)Lw\phi_j q_j^{1/\nu-2}} > 0, \tag{5.10}$$

Proposition 2 *More productive plants produce more. However, there is an upper limit on output size that increases with L.*

The existence of an output cap is a direct result of prices having to be non-negative and bounded by α in (5.3).

Lemma 1 *Revenue productivity is increasing in physical productivity, that is $d\theta_j/d\phi_j > 0$.*

In the absence of demand and productivity shocks, revenue productivity is a monotonic and one-to-one transformation of physical productivity. This transformation consists of a scaling (non-uniform unless $\nu = 1$) and a shift. This result proves useful, because any qualitative model implications with respect to physical productivity can be immediately generalized to revenue productivity too. For this reason, in the coming propositions, I will refer to both measures of productivity simply as "productivity".

Continuing with the analysis, I combine (5.4) and (5.7) to get the following relation between input and output size:

$$x_j = \left(\frac{q_j}{\phi_j}\right)^{1/\nu} = \frac{\nu}{w}\left(\frac{\alpha\gamma + \eta N\bar{p}}{\gamma + \eta N} - \frac{2\gamma}{L}q_j\right)q_j. \tag{5.11}$$

Taking partial derivatives in (5.11) with respect to ϕ_j and using (5.10), it can be shown that

$$\begin{cases} \partial x_j/\partial\phi_j \geq 0 & \text{if } q_j \leq \frac{L(\alpha\gamma+\eta N\bar{p})}{4\gamma(\gamma+\eta N)}, \\ \partial x_j/\partial\phi_j < 0 & \text{Otherwise.} \end{cases} \tag{5.12}$$

Proposition 3 *Under the localized market assumption and when $\gamma > 0$, the relationship between input size and productivity is bell-shaped* [3].

Proposition 3 follows from continuity of the solutions plus (5.12). This result is a major departure from standard models, for it asserts that the input size need not grow monotonically with productivity. In my model, input size inside a market goes up only to the extent permitted by demand limitations. The maximum value of input size provides some measure of "size opportunities" in that particular market. Using (5.11) and (5.12) this maximum value is

$$x_{\max} = \frac{L\nu}{8w\gamma}\frac{\alpha\gamma + \eta N\bar{p}}{\gamma + \eta N}. \tag{5.13}$$

Taken at face value, (5.13) suggest that the peak of the bell-shaped curve should get higher and move to the right as markets get larger, though the endogeneity of N and $\bar{p}$ require some caution when making such statements. The simulation results in Sections 5.4.2 and 5.5.2, nevertheless, are clearly consistent with this assertion.

Given Proposition 3, the bell-shaped relationship between input size and pro-

[3]This characteristic is not a result of using the utility function (5.1) that yields a linear demand curve. It is easy to show that the unimodality of the curve holds under much weaker assumptions. To show this, let $p(q_j)$ be a general demand curve with $p'(.)$ and $p''(.)$ its two first derivatives with respect to q_j. Writing the first order conditions, and after some algebra, gives the equation that specifies input size as

$$x_j(q_j) = \frac{\nu}{w}q_j\Big(p'(q_j)q_j + p(q_j)\Big).$$

The slope of this curve is

$$x'_j(q_j) = \frac{\nu}{w}\Big(p''(q_j)q_j^2 + 3p'(q_j)q_j + p(q_j)\Big).$$

Assuming the boundedness of the demand function and its derivatives, it is obvious that $x'_j(0) > 0$. Further assuming that $p'(.) < 0$ and $p''(.)$ is bounded above by a small enough value (possibly positive), then $x'_j(q_j)$ becomes negative at some output level and stays negative afterwards. Because of the monotonicity of the relationship between productivity and output, the same deduction equally applies to the relationship between productivity and input size.

The assumptions on $p(.)$ state that the unimodality property is preserved if demand is not discontinuous at any point and its elasticity does not increase too fast. With discontinuities or large increases in demand elasticity, demand size expands very fast with small changes in price, so that the plants operating at an incrementally lower price enjoy a surge in demand and are inclined to have larger sizes as a result. Under such conditions, the size-productivity curve might exhibit positive monotonicity or might have more than one maximum.

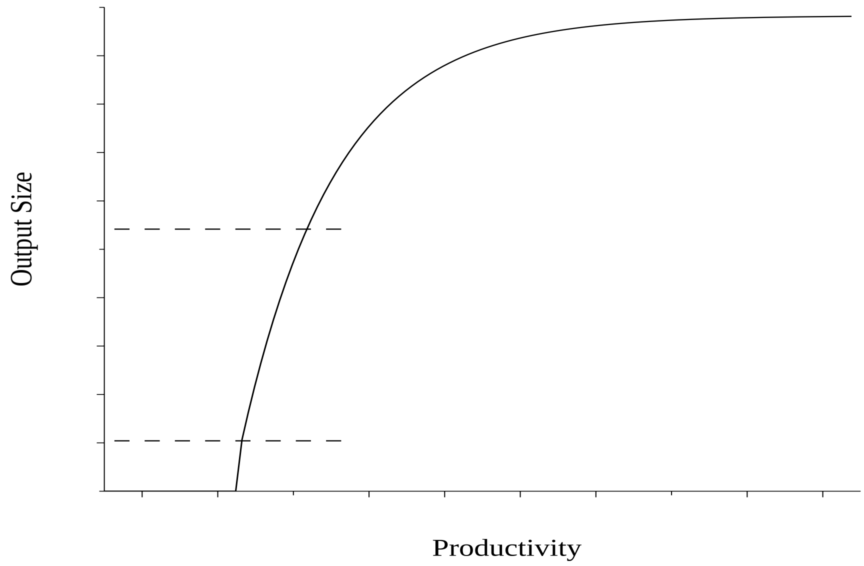
ductivity is no surprise. With output bounded from above, more productive plants are able to produce the limit output by hiring smaller inputs. The more productive they get, the less amount of input they need to produce that output (Figure 5.1). It is also useful to repeat that the localization of markets is essential to Proposition 3. This assumption makes it impossible for the more productive plants to improve their demand by trading with other markets, so that market size becomes a parameter in determining the plant performance. The following proposition characterizes the behavior of plants when their markets become globalized.

Proposition 4 *As $L \to \infty$, the relationship between input size and productivity converges to a monotonic increasing relationship.*

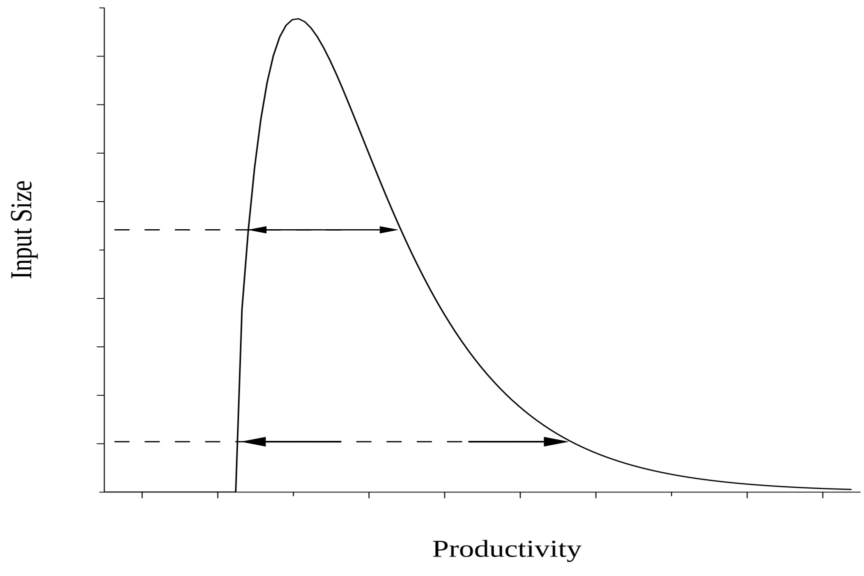
Figure 5.2 shows the limit behavior of plants when L goes to infinity for both constant and decreasing returns to scale production functions. With decreasing returns in the production function ($\nu < 1$), the relationship between input size and productivity converges to a strictly increasing exponential relation. With constant returns production function ($\nu = 1$), since the marginal productivity is not affected by size, the input sizes all go to infinity in the limit, rendering a not so strictly monotonic relation. In both cases, market size does not play a role in the plant performance anymore. Note that as $\gamma \to 0$, the implication is not the same as Proposition 4. With $\gamma = 0$, the products are perfectly substitutable, and the most productive plant in market can offer the lowest price and take over the whole market. Because of that, we shall have $\bar{p} = 0$ and $N = 0$ and a degenerate distribution of size and productivity in the market [4].

The bell-shaped relationship between productivity and input size in localized markets is what allows this model to generate higher productivity dispersion at lower levels of input size. For any given market, the gap between productivity differences closes as the level of input goes higher, and at the maximum input level of (5.13) the

[4]Notice that with a continuous measure, the measure of a singleton is zero.



(a)



(b)

Figure 5.1: The bell-shaped relation between (a) output and productivity and (b) input and productivity. The arrows demonstrate the range of productivity dispersion in small and large plants.

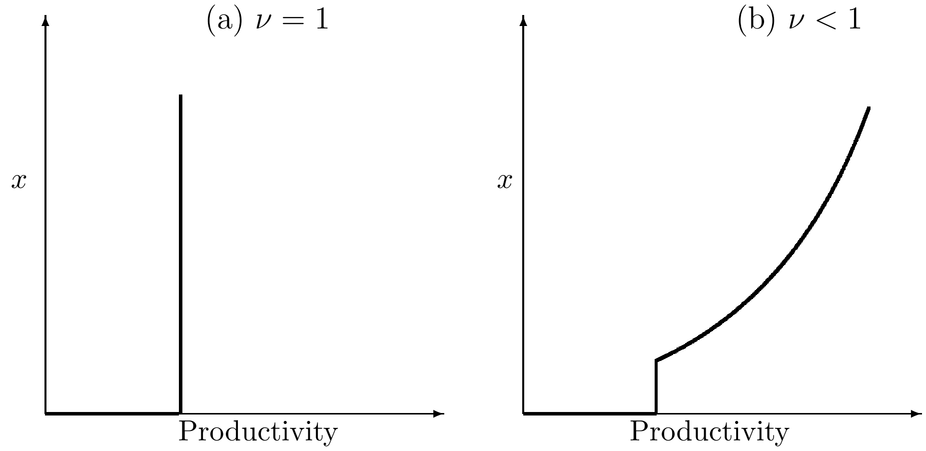


Figure 5.2: The limit behavior of plants when $L \to \infty$.

dispersion goes to zero (Figure 5.1). Later, in Section 5.4, a continuum of markets of different sizes is used to generate a dense productivity distribution at any given input size whose dispersion still falls with the level of input[5]. On the other hand, at any given output size within a market the productivity dispersion is zero. This is a direct consequence of the monotonicity of the relationship between output size and productivity. However, when markets of different sizes are present, productivity dispersion by output size can still be generated.

The cutoff productivity ϕ_* is another variable in this model that affects the distribution of productivity within a market, especially for plants with lower input sizes. Therefore, it is useful to know how the cutoff productivity varies by market size. Let q_* and x_* be the output and input size for the plant operating at cutoff productivity ϕ_*. With some algebra, it can be shown that

$$\frac{\partial \phi_*}{\partial L} = \frac{\frac{\gamma}{L^2}\left(\frac{\gamma}{L}\frac{2\nu-1}{\nu}q_* + \frac{f}{\nu q_*}\right)\int_{\phi_*}^{\phi_M} q(q-q_*)dG(\phi)}{\frac{wx_*}{\nu^2\phi_*}\left(\frac{f}{q_*^2} + (2\nu-1)\frac{\gamma}{L}\right)\int_{\phi_*}^{\phi_M} q dG(\phi)}. \tag{5.14}$$

[5]Unimodality of the bell relationship is not essential in this discussion. As long as the range of productivities at the top of the curve is lower than its base, a declining productivity dispersion by input size can still be produced.

Proposition 5 *When* $\nu \geq 0.5$*, then* $\partial\phi_*/\partial L > 0$*.*

Proposition 5 is in line with the findings of Melitz & Ottaviano (2005), who find that in larger markets tougher competition drives out the less productive plants and raises average productivity in the market. It can also be shown that $\partial\phi_*/\partial\gamma < 0$, consistent with the finding of Syverson (2003) that more product substitutability, or equivalently smaller γ, leads to higher cutoff productivities.

Summarizing, (5.12) and (5.14) together provide two instruments by which distributions of productivity across markets can be analyzed and compared. The generated patterns of productivity dispersion in the later sections of this paper will be a direct application of both of these findings. As the last comment, it should also be emphasized that the restriction $1 \geq \nu \geq 0.5$ merely helps to resolve ambiguous signs on the comparative statics. The model outcome is by no means limited to decreasing or constant returns to scale. In fact, because of the continuity of all relations, the propositions and model implications still hold in some neighborhood of $\nu = 1$, which also includes regions of increasing returns to scale. In practice, there is little reason to believe that industries' return to scale is much beyond 1.

5.2 Data Considerations

5.2.1 Data on Plant Performance

I test the model implications by looking at the Ready-Mix Concrete industry (SIC 3273) as an industry that shows high degrees of market localization and (geographic) product differentiation, two conditions required by the model to generate a falling productivity dispersion by input size. Due to high costs of transportation, concrete plants do not ship their output very far compared to other manufacturing industries, therefore they qualify as localized market[6]. The physical output is mostly homogeneous. As a result, the magnitude of revenue variation due to quality or taste differences is largely minimized, leaving mostly physical productivity to drive differences in revenue productivity across plants. Finally, concrete is a highly differentiated industry not by product variety, but by spatial diversity (Syverson 2004). Due to high costs of transportation, customers make purchase decisions not only based on efficiency of production but also based on physical distance. As a result of this diversity, a wide range of productivities are present in the data for my analysis.

I am including data on concrete plants from the 1982, 1987, 1992, and 1997 CM panels. I use the weighted CM subsamples for my analysis and estimations, which provides me with a total of 3970 sample concrete plants. I also use the total number of concrete plants operating in a certain region from the complete list of CM records. For descriptive results, to make productivities comparable over a range of years within the concrete industry, I use residuals from regressing productivity values on year dummies. I then readjust the mean value of residual productivities so that it is equal to the original mean.

5.2.2 Demand Market

Due to availability of detailed data and required crosswalks, I use Core-Based Statistical Areas (CBSA) as markets for concrete plants. A CBSA is a functional region around an urban center. The CBSA system includes a mix of micro- and metropolitan areas in the United States, providing me with a sufficiently large range of market sizes[7]. Economic activity is mostly concentrated within a CBSA, making

[6]The US Bureau of Transportation Statistics' Commodity Flow Survey reports that concrete plants shipped their products to an average radius of 64 miles in 1993 and 82 miles in 1997.

[7]US Office of Management and Budget's definition of a metropolitan area is an urban area with population of at least 50,000. Micropolitan areas are those with population between 10,000 and 50,000.

it a suitable candidate for market analysis, though the degree of market isolation can still depend on the physical proximity of CBSA's.

The demand for concrete in a particular is measured as the population of construction workers (SIC 15– to 17–) aggregated to the CBSA level. Syverson (2004) discusses the suitability of such a definition by arguing that the construction industry is the main consumer of ready-mix concrete, while costs of concrete is a small share of construction costs. This makes the demand measure reasonably with respect to productivity shocks to concrete. Construction employment is taken from the County Business Patterns aggregated to the CBSA level and matched by CBSA-year[8].

There are 667 markets that match to my existing plants. The minimum market size is 48 construction workers (Yazoo, Mississippi) and maximum market size is 327,397 construction workers (New York, New York). More detailed statistics for this market definition can be found in Table 5.7 where I will compare different definitions of market sizes.

5.3 Plant Behavior in Localized Markets

The theoretical model predicts that market localization will affect the behavior of productivity dispersion. Under market localization, the relationship between productivity and input size was shown to be bell-shaped, a fact that will be put to test in the data. In addition, I will also estimate the relationship between the number of operating plants and market size, something left mostly unexplored above because of model complexity. These estimates will provide me with moments that I will use in Section 5.4 to pinpoint the model parameters and simulate results.

[8]The employment data for some of the counties is suppressed to protect confidentiality of the data. I follow Syverson's method to impute those data. Basically, since the number of employers in several different size groups is being reported, I will multiply the number by mid point of the size range and sum up to generate the impute.

Because physical productivity estimates are unavailable from the data, most of the remaining discussions rely on revenue productivity θ. Both measures rTFP and rLP are used when appropriate for robustness.

In the coming empirical results, instead of measuring a composite input, I will measure the input size of plants by their total employment (TE) as defined by Davis et al. (1996, Appendix A.3.1). Employment is easily observed for each plant and has reasonably low measurement error compared to estimates of a composite input. In defense of this shift, note that if the relative intensity of productive factors is assumed constant within an industry, the optimal choice of each input factor will be a constant proportion of employment size, so that the composite input will be a linear function of employment. This enables me to treat the production function (5.4) as if it depended on labor only.

5.3.1 Employment and Productivity Relationship: The Outline

In this section, I seek a basic description of the relationship between productivity and employment using data on concrete plants, and I further investigate the effect of market size on the shape of the relationship. At this stage I impose as little structure as possible, relying on visual investigation of plant concentration along the employment and productivity axes. These observations are helpful in motivating the more structured estimation results that will follow. What I am showing here is the region where plants are mostly concentrated. This is done by detecting the edges of the scatter plot and plotting them in the form of an outline[9]. Edge detection is a popularly used method used in machine vision to recognize objects

[9]Due to the Census Bureau's requirements to protect the confidentiality of individual data, it is not permissible to show the scatter plot in its raw form without enough safeguards.

in a picture[10]. In its simplest form, it is implemented by detecting areas where the intensity of pixels changes abruptly, which I use to detect the points forming the edges in a scatter plot where the less populated area borders the more populated area [11]. The estimated edges are still rugged and noisy because of outlier effects. To further suppress individually recognizable information, I smooth the edges by limiting the first derivative of edge curvature and then passing the edge points through an averaging filter[12]. The final result is illustrated in Figure 5.3. Plots with both rTFP and rLP clearly show a bell-shaped area of concentration for concrete plants. The effect of market size is also put to test by detecting the scatter area for plants belonging not to the full range of market sizes but to the range of market sizes up to 3000 construction workers. Figure 5.3 shows that the detected scatter area for plants belonging to smaller markets still resembles a bell-shape (a smaller one) and covers the lower section of the whole scatter area. These are the plants tied to more limited markets, hence they are smaller in size. This is consistent with the general expectation and with the model's predictions in particular. Similar to the model, the data shows that the bell-shaped area grows upwards as a result of expanding markets.

With this preliminary evidence about the behavior of productivity and employment in the concrete industry, the next section will use a more structured approach

[10]Ziou & Tabbone (1998) offer an extensive introduction to popularly used methods and discuss other practical issues concerning edge detection.

[11]My approach works in this way: I first divide the productivity-employment plane into a 350×250 grid-map and flag the existence of any plant in each cell area as the intensity. Then I use Sobel's mask to estimate partial derivatives of pixel intensity in x and y (productivity and employment) directions and use the following as a measure of total intensity change

$$\Delta = \sqrt{\Delta_x^2 + \Delta_y^2},$$

where Δ_x and Δ_y are the estimates of partial derivatives. Edges are detected by picking points at which the absolute value of the estimated derivative is larger than some threshold.

[12]The shape of my averaging filter is

$$\tilde{x}_i = 0.3x_{i-1} + 0.4x_i + 0.3x_{i+1},$$

where $\tilde{x}_i$ is the filtered value and i indexes productivity points when sorted in ascending order.

to estimate a smooth bell-shaped relationship in the data and to investigate the effect of market size in more detail.

5.3.2 Employment and Productivity Relationship: Semi-Parametric Model

I will estimate a smooth relationship between productivity and employment in the data by fitting a semi-parametric model. The relationship between productivity and employment is characterized precisely in the theoretical model. Hence, I will use a polynomial of pre-determined degree in logs of variables to approximate that relationship. However, the effects of time and market size are more obscure in the model and will be approximated non-parametrically by fitting thin-plate splines (Moussa & Cheema 1992)[13]. The general form of my model is

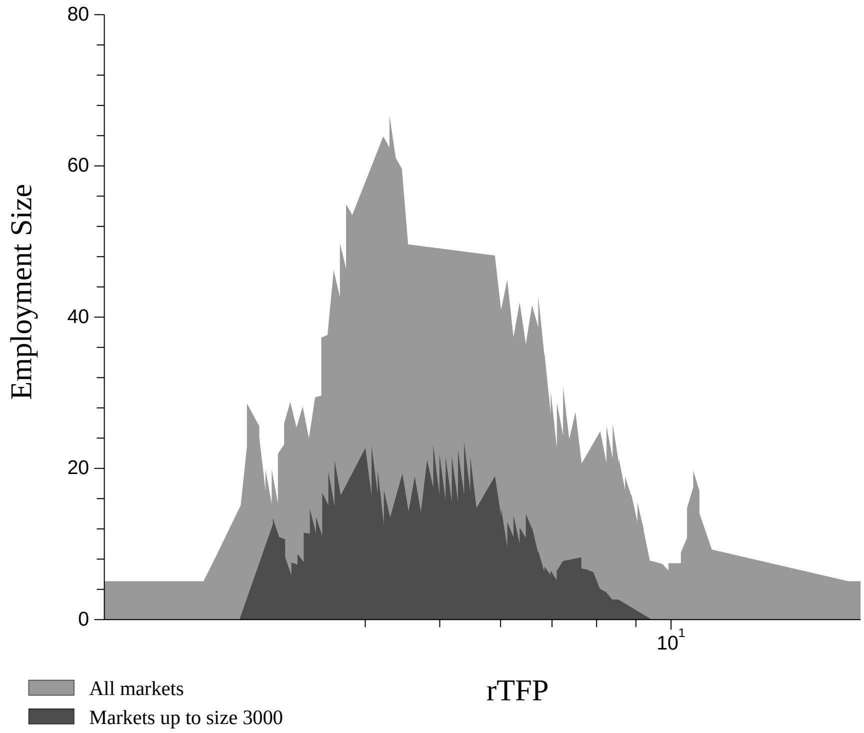
$$\log(l_{jt}) = \sum_{p=0}^{P} \alpha_p \log(\theta_{jt})^p + h(L_j, t) + \epsilon_{jt}, \tag{5.15}$$

where l_{jt} and θ_{jt} are respectively the employment size and the revenue productivity for plant j at time t. L_j is the market size for plant j. To minimize the computational burden and to reduce running time down to a reasonable length, market size is classified by its log being rounded to the nearest 0.5. P is the degree of the polynomial term used in the model.

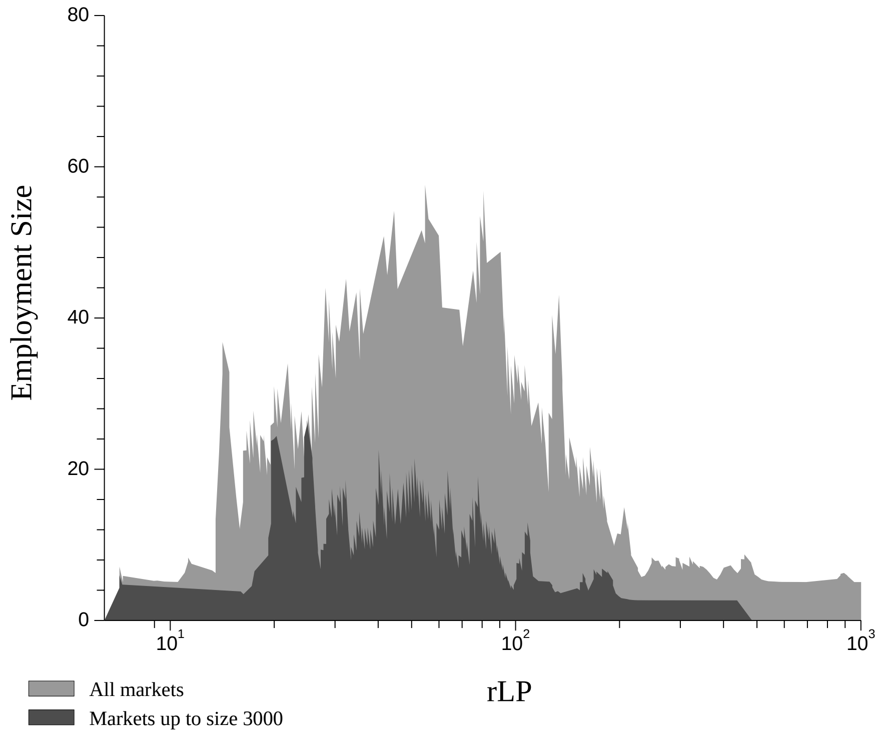
The estimates are computed using a penalized least-squares function that minimizes the following function with respect to α_p's and a proper choice of the function $h(.,.)$

$$S_\lambda = \frac{1}{s_{jt}} \sum_j \epsilon_{jt}^2 + \lambda J_2\big(h(L,t)\big). \tag{5.16}$$

[13] Moussa & Cheema (1992) survey sets of different basic functions that can be used for this purpose and describe the fitting method. Splines are preferred because they can produce a better and smoother fit using lower orders when compared to polynomials. The fit is done by approximating the non-parametric part of the relationship by a linear sum of basic functions up to a finite order.



(a)



(b)

Figure 5.3: The outline of the productivity-employment relationship in the data.

$J_2\big(h(.,.)\big)$ is a measure for the roughness of the fit and is normally defined as the integral of the square of the second derivative of h with respect to its arguments. More formally

$$J_2\big(h(L,t)\big) = \int_0^\infty \int_0^\infty \left[\left(\frac{\partial^2 h}{\partial L^2} \right)^2 + 2\left(\frac{\partial^2 h}{\partial L \partial t} \right)^2 + \left(\frac{\partial^2 h}{\partial t^2} \right)^2 \right] dLdt \qquad (5.17)$$

In practice, approximations are used to compute the above integral, so that computational complexity is kept within reasonable bounds. λ is the penalty parameter, whose choice is a trade-off between accuracy of the fit and its smoothness. s is the number of observations used. My actual choice of value for λ proved not to be very crucial as the estimation result remains practically unchanged for values of λ within a wide range from 0.1 to 10. I report results when I set λ equal to 1.

The choice of polynomial degree in model (5.15), however, seems critical. A small value of P will not capture enough curvature, and high values of P will add in noise and cause instability of estimates. In an experimental stage, I added polynomial powers one by one, until the estimates started to become unstable. The most stable predictions are achieved when $P = 4$.

To demonstrate the estimation results, output was produced for three representative market sizes: 1000, 10000, and 100000. Care was taken that the sizes form a geometrical series, so that the results provide a clue as to whether the market-size effect is linear or non-linear in the data. It is useful to recall that in the theoretical model market size affects the shape and peak of the bell non-linearly.

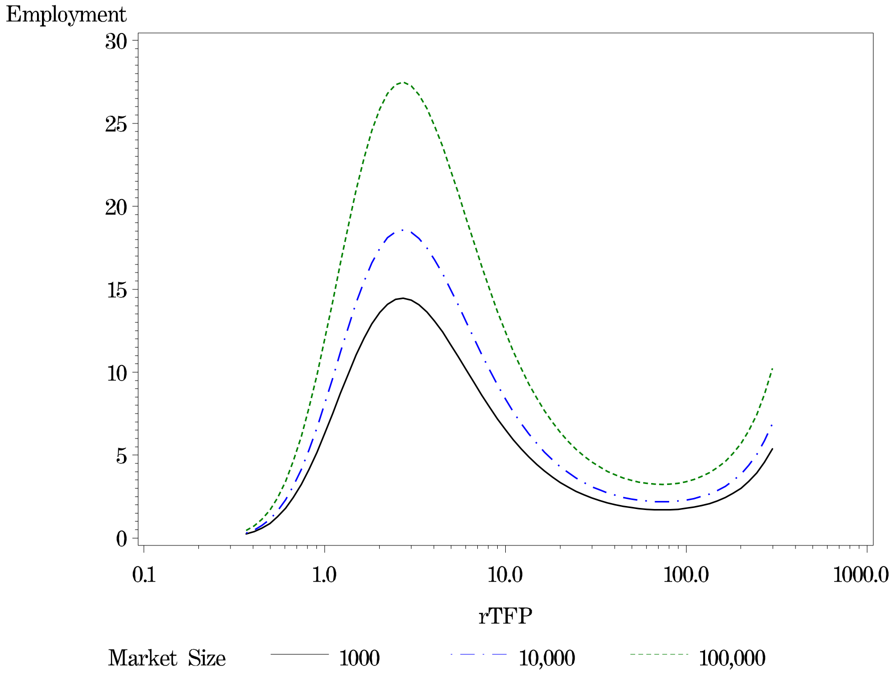
The estimated curves are shown in Figure 5.4. The plots are in agreement with expectations. First, the relationship between productivity and employment within the concrete industry is of a bell-shaped form. Second, the effect of market size is shown to be consistent with model prediction, where plants are on average larger and the width of the bell is wider in larger markets. Third, scaling market size affects the results non-linearly: Going from market size 1,000 to 10,000 increases the peak size by 28%, while going from market size 10,000 to 100,000 causes a 48% increase (using the plots with rTFP). This non-uniformity of scaling will be revisited in the later simulations of the theoretical model where it will be shown that doubling market size again causes the peak plant size to more than double. It is also noteworthy that both measures of productivity are mostly similar in their predictions, raising confidence in the estimated shape and behavior of the productivity-employment relationship.

5.3.3 Number of Plants per Market

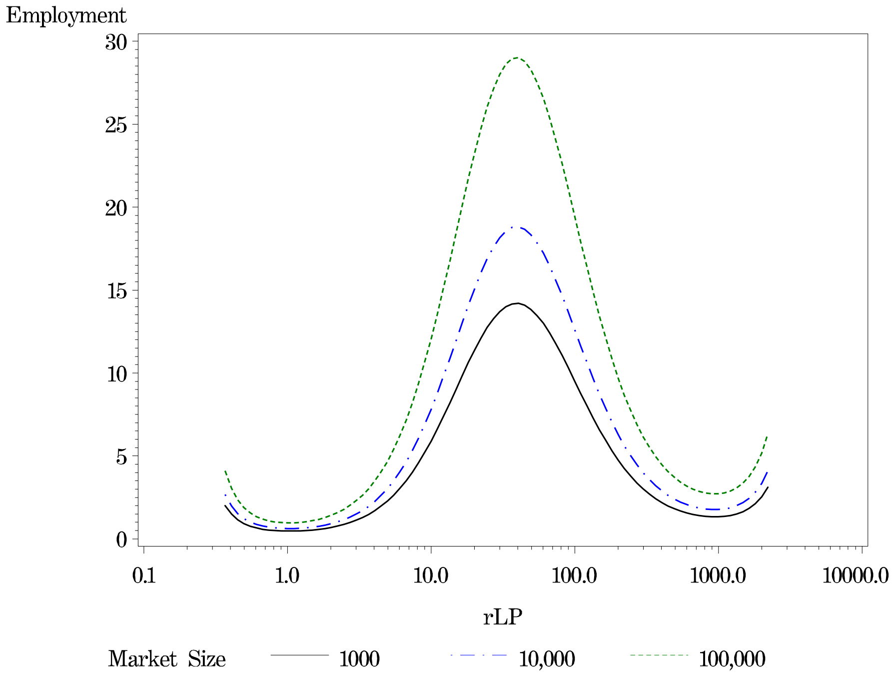
The empirical relationship between the number of plants and market size is another moment that will be needed in Section 5.4 to estimate the complete set of model parameters. So this section is dedicated to the empirical estimation of such a relationship. The theoretical model of Section 5.1 does not provide an analytic result about the relationship between N should and market size L. In practice, larger markets offer larger demand and should have the capacity to accommodate more production plants. This fact seems especially likely under decreasing returns production, where the production function intrinsically favors a large number of smaller operators. Asplund & Sandin (1999) offer evidence for a positive effect of market size on N.

In the data, each plant can be associated with a market size through its geographical link. The total number of plants operating in that market can be found from the complete CM panel. A penalized least squares method is again used to produce a smooth non-parametric relation between the number of plants and market size in the data. The relation is of the form

$$n_m = v(\log(L_m)) + \zeta_m. \tag{5.18}$$



(a)



(b)

Figure 5.4: Estimated productivity-employment relationship in the concrete industry.

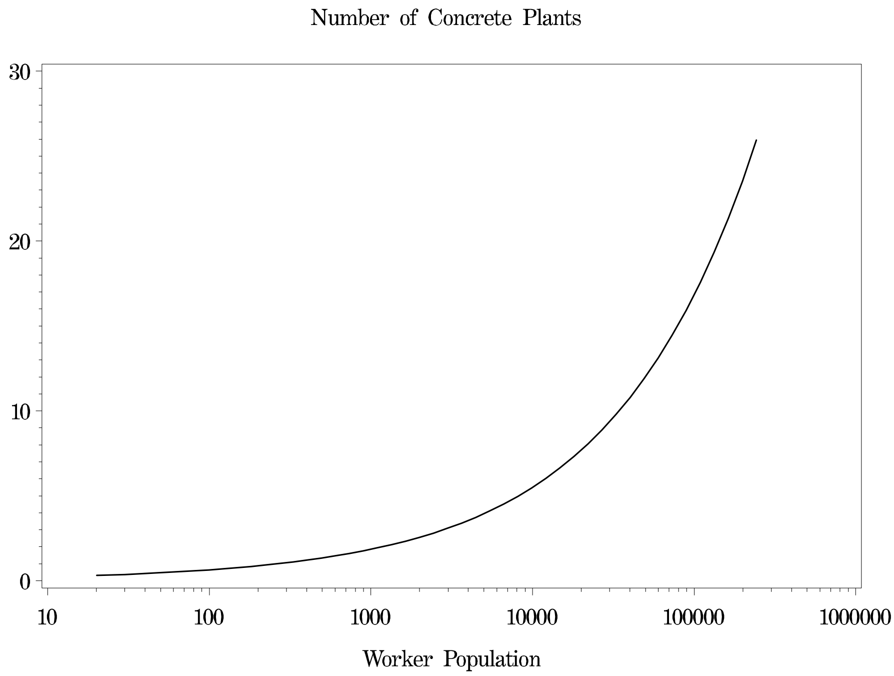


Figure 5.5: Estimated number of concrete plants as a function of construction employment.

Here m indexes each CBSA, and n_m is the number of plants operating in market m in log levels. L is the population of construction workers in the corresponding market. The value of the penalty parameter λ is set to 10 in my preferred specification, and the result is shown in Figure 5.5. The picture suggests that larger markets are host to a larger number of plants, as expected. As will be shown in the simulation results below, the theoretical model can replicate the same relationship very closely under both constant and decreasing returns to scale production.

5.4 Numerical Simulation

The first-order condition (5.7) simplifies to a linear equation when $\nu = 0.5$ or $\nu = 1$. In these two special cases, closed form solutions for output and input size are obtainable, and the solution to each case constitutes an extreme case of the model behavior[14].

At the same time, existing evidence suggests that most industries produce with close to constant returns to scale (Basu & Fernald 1997). Concrete is one such industry, with an estimated returns to scale of 0.996 (Syverson 2004). Therefore, I treat the case $\nu = 1$ with special attention, while deferring some analysis with $\nu = 0.5$ to Section 5.5.2 as a robustness check. Also, as in the previous section, the size of plants are expressed in total employment.

5.4.1 Estimation Methodology

Simulating the theoretical model entails estimating a set of parameters $\Lambda = \{\alpha, \eta, \gamma, \phi_M, w, f_E, f\}$ that minimizes the weighted squared error between the data provided moments and the simulated moments from the model. Section 5.3 provided two sets of data moments that can be of use in the estimation: (1) the relationship between employment and productivity, and (2) the relationship between number of plants and the market size.

Before describing the estimation method, it is useful to note that the profit function (5.6) can be written in the following form[15]

$$\pi_j = \left(\frac{f}{(1-\nu)q_*} - \frac{\gamma}{L} \frac{2\nu - 1}{1-\nu} q_* - \frac{\gamma}{L} q_j \right) q_j - w \frac{q_j^{1/\nu}}{\phi_j^{1/\nu}} - f. \tag{5.19}$$

From (5.9) and (5.19), q_* depends only on ϕ_* and the reduced set of parameters

[14] Finding the behavior with these two values is sufficient to picture the range of behaviors generated by ν changing from 0.5 to 1. Proposition 1 states that the solution is unique and does not bifurcate as ν changes. The solution having a single path as ν changes, combined with the continuity of (5.7) with respect to ν, guarantees that for two values ν_1 and ν_2 that are close enough, their corresponding bell-curve solutions resulting from the model for the same market size will be close too. In more technical terms

$$\forall \epsilon_2, ||E(\phi, L; \nu_2) - E(\phi, L; \nu_1)|| < \epsilon_2 \quad \Rightarrow \quad \exists \epsilon_1, |\nu_2 - \nu_1| < \epsilon_1,$$

where $E(\phi, L; \nu)$ is a bell-curve solution to the model using ν for a given L and the full range of ϕ.

[15] Please refer to the Appendix B for details of how to derive this relation.

$\Lambda_0 = \{\gamma, \phi_M, w, f_E, f\}$. It then follows from (5.8) that ϕ_* is implicitly a function of Λ_0. As a result, both output and input size can be written as functions of $q_*(\phi_*)$, ϕ_* and Λ_0, making it easier first to estimate the reduced set Λ_0 using a weighted nonlinear least squares method, and then use the estimated parameters to pin point α and η [16]. However, the dependency of ϕ_* on Λ_0 creates an identification problem: I need to know ϕ_* to estimate Λ_0, but to compute ϕ_* the parameter set Λ_0 must be known. At the same time, the unavailability of data on physical output productivity makes separate identification of some parameters impossible when ϕ_* is not known. These facts, together, make it impractical to use standard nonlinear least squares methods to estimate the parameters.

Instead, I use a recursive method of simulated moments estimator described in the following algorithm.

Algorithm 1

1. An initial Λ_0 is assigned,

2. Using Λ_0, ϕ_* is computed and simulated moments are produced as a function of revenue productivity θ (not ϕ).

3. Using a search method, a new parameter set Λ_0 is found that reduces the sum of squared errors between the empirical and simulated moments[17].

4. Steps 2 and 3 are repeated until the change in parameter set falls below an acceptable tolerance[18].

[16] α and η enter the performance measures in a certain form easily replaceable by (B.13) from Appendix B.

[17] I use a pattern search with trust region adjustments to perform this search. Due to presence of implicit and complicated functions, finding analytical gradients and Hessians proved to be non-trivial. Gradient methods using numerical gradient computations with BFGS adjustments also got stalled. Alternatively, pattern search is completely insensitive to such irregularities.

[18] I use 10^{-12} as tolerance bound.

The form of the weighted nonlinear least squares problem is

$$\min_{\Lambda_0} \quad (E^{data} - \hat{E}(\Lambda_0))W_0(E^{data} - \hat{E}(\Lambda_0))', \tag{5.20}$$

where E^{data} is the vector of estimated employment moments from Section 5.3.2 stacked for four market size classes. I use the relationship estimated for rTFP data in Figure 5.4(a). $\hat{E}$ is the vector of corresponding simulated moments generated by the model when using Λ_0. The moments are simulated using revenue productivities for each market class separately and plugging the average market size in that class as L into the model. In this way, the dimension of vectors E^{data} and $\hat{E}$ is the same.

W_0 is a weighting matrix that governs the importance of different moments in setting the parameters. The estimated productivity-employment relationship from the data are estimated to be smooth, with much of the noise already filtered out, therefore I will use uniform weighting. The only irregularity in the shape of the estimated moment happens when rTFP is larger than 100. This behavior seems to be a result of truncation error caused by limiting the power of estimated polynomial. This reasoning is affirmed by the fact that the scatter plot of Figure 5.3 does not show any mass of observations with large employment and rTFP higher than 100. For this reason I chose a diagonal W_0 in which all diagonal elements are 1 when rTFP is less than or equal to 100, and zero otherwise. This weighting effectively prevents the truncation error in the upper tail of the estimated moment to affect the parameter estimates. As a result, employment size is fitted using 57 points and for three market sizes, providing a total of 171 points for my parameter estimation.

In the second stage, I take the estimated Λ_0 from (5.20) and estimate the parameters α and η, using nonlinear least squares to minimize the following error function

$$\min_{\alpha,\eta} \quad (N^{data} - \hat{N}(\Lambda))W_1(N^{data} - \hat{N}(\Lambda)) + \lambda_1 \hat{N}(\Lambda) I[\hat{N}(\Lambda) < 0]\hat{N}(\Lambda)'. \tag{5.21}$$

Here, N^{data} is the vector of the number of plants operating at each market size. I use data from Figure 5.5(a), which provides me with 48 point estimates of N over L. $\hat{N}$ is the simulated number using the complete set of parameters Λ, where Λ_0 are the estimated values from stage 1 and are fixed. $\hat{N}$ is estimated for the market size classes from the data, so that N^{data} and $\hat{N}$ are forced to have the same dimension. W_1 is the weighting matrix, and, because of the smoothness of the estimated moment, I use identity matrix that weights all estimates uniformly. The extra term in (5.21) is a penalty term that forces the simulated vector $\hat{N}$ to have non-negative values. λ_1 is the penalty parameter and $I[\ ]$ is the diagonal matrix of indicator functions. By imposing a large penalty parameter λ_1, I make sure that the estimated values for α and η will not result in a negative simulated number of plants for any market. In my exercise, I set the value of λ_1 to 1000.

5.4.2 The Constant Returns Case

With a CRTS production function ($\nu = 1$) a closed form solution to (5.7) can be found as follows

$$q_j = \frac{L}{2\gamma}\left(\frac{\alpha\gamma + \eta N\bar{p}}{\gamma + \eta N} - \frac{w}{\phi_j}\right). \tag{5.22}$$

Using (5.22), the optimal profit can be computed and used in (5.9) to solve for the cutoff productivity, which yields

$$\phi_* = \frac{w}{\frac{\alpha\gamma+\eta N\bar{p}}{\gamma+\eta N} - \sqrt{2\frac{\gamma f}{L}}} \tag{5.23}$$

A feature of the above cutoff productivity is that it summarizes the effects of endogenous variables N and $\bar{p}$ on the plant behavior. As discussed before, all plant performance measures can be expressed as a function of ϕ_* and model parameters as follows:

$$\begin{aligned}
q_j &= \frac{Lw}{2\gamma}\left(\frac{1}{\phi_*} - \frac{1}{\phi_j} + K\right), && (5.24)\\
p_j &= \frac{w}{2}\left(\frac{1}{\phi_*} + \frac{1}{\phi_j} + K\right), && (5.25)\\
l_j &= \frac{Lw}{2\gamma\phi_j}\left(\frac{1}{\phi_*} - \frac{1}{\phi_j} + K\right), && (5.26)\\
\pi_j &= \frac{Lw^2}{4\gamma}\left(\frac{1}{\phi_*} - \frac{1}{\phi_j} + K\right)^2 - f, && (5.27)\\
N &= \frac{2\gamma}{\eta w}\left(\frac{\alpha - \frac{w}{\phi_*} - wK}{\frac{1}{\phi_*} - \overline{\left(\frac{1}{\phi_j}\right)} + \frac{1}{2}K}\right), && (5.28)\\
K &= \frac{2}{w}\sqrt{\frac{\gamma f}{L}}.
\end{aligned}$$

with l_j being the employment size (replacing x_j). Also, using the definition (5.5), the revenue productivity θ_j can be expressed as

$$\theta_j = \frac{w}{2}\left(1 + \phi_j\left(\frac{1}{\phi_*} + K\right)\right). \tag{5.29}$$

Obviously, θ is a function of the input price w and the market elasticity of demand embodied in ϕ_* and K, as well as the efficiency of production ϕ. The analysis of Section 5.1.3 together with the definition of K show that larger L or smaller γ decrease the coefficient multiplying ϕ_j in (5.29). That, in turn, causes revenue productivity to under-represent the efficiency of production, especially in large markets. The dispersion of revenue productivity is actually affected by the scaling effects of w, ϕ_* and K, as well as by ϕ_* cutting the distribution from below. Since the under-representation affects larger markets more seriously, the dispersion of revenue productivity should fall faster with employment than its physical productivity counterpart.

Also, the revenue cutoff productivity θ_* can be found by replacing ϕ_j with ϕ_* in (5.29) so that

$$\theta_* = w + \phi_* \sqrt{\frac{\gamma f}{L}}. \tag{5.30}$$

In (5.30), w is the cutoff revenue resulting from the input price. The second term is caused by presence of fixed costs and changes with market size. However, at this point, it is not clear if the relationship is positive or negative. Later analysis of this section reveals that ϕ_* increases with market size at a lower rate than $\sqrt{L}$, which causes θ_* to slowly decline with market size.

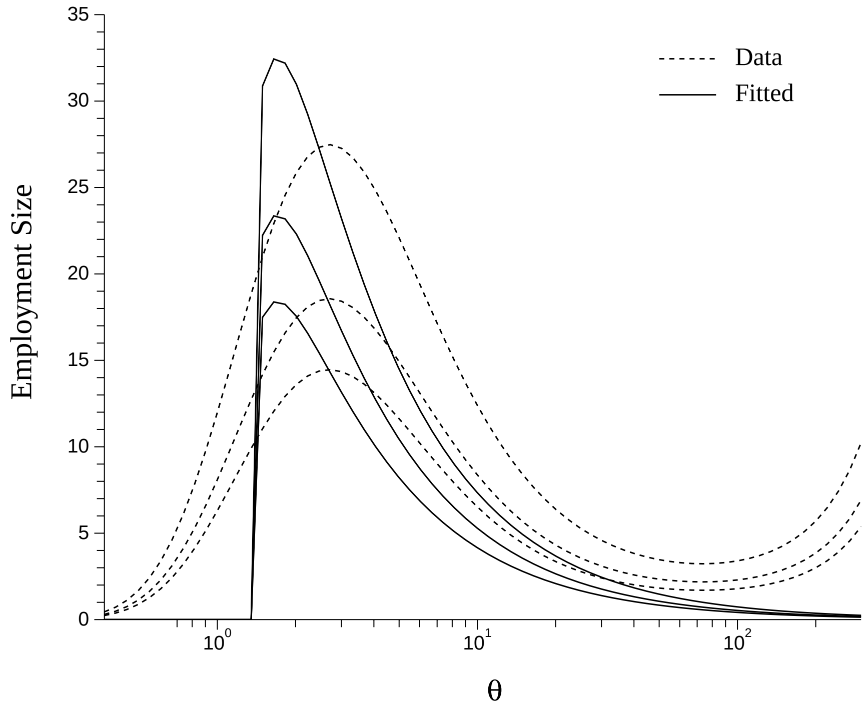
Upon entry, plants draw their random productivity from a variant of the Pareto distribution, whose cumulative distribution function is

$$G(\phi) = \frac{\log(1+\phi)}{\log(1+\phi_M)}. \tag{5.31}$$

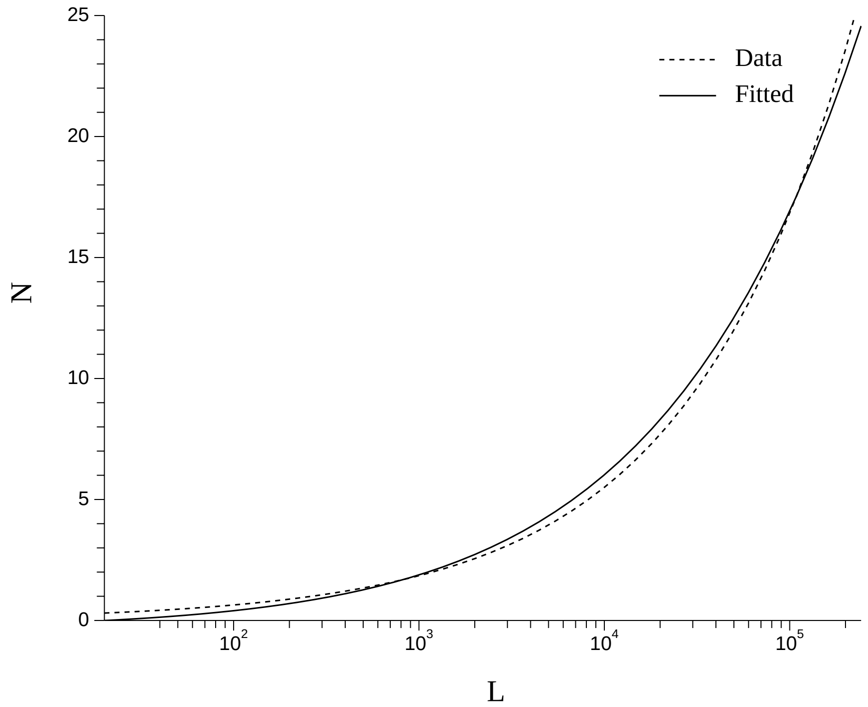
This distribution has a number of advantages for my analysis. As in the data, it implies a low probability of high productivity draws. Also, this functional form reduces the computational burden and improves the convergence of solutions. With this productivity distribution, the free entry equilibrium condition (5.8) can be solved for ϕ_* as a function of L, w, γ, f, and f_E, as elaborated in Section 5.4.1.

Using (5.26) and (5.28) to generate simulated relationships between size and productivity and between market size and the number of plants, the model parameters are estimated as shown in Table 5.1. The goodness of fit is tested by computing the corresponding standard deviation of the residual for each estimation stage separately, displayed in the same table. Figure 5.6 presents over-imposed plots of the fitted curves and the data moments to demonstrate the degree to which simulation fits the data.

With these estimated parameters, I simulate my model for several market sizes



(a)



(b)

Figure 5.6: Graphical demonstration of goodness of fit with CRTS production function.

α	η	γ	ϕ_M	w	f_E	f
2.724	7.030	1.178	704.69	1.133	21.137	5.649
Stage 1 estimation error					$\sigma_{error,1} = 4.615$	
Stage 2 estimation error					$\sigma_{error,2} = 0.424$	

Table 5.1: Parameter estimates with CRTS production function.

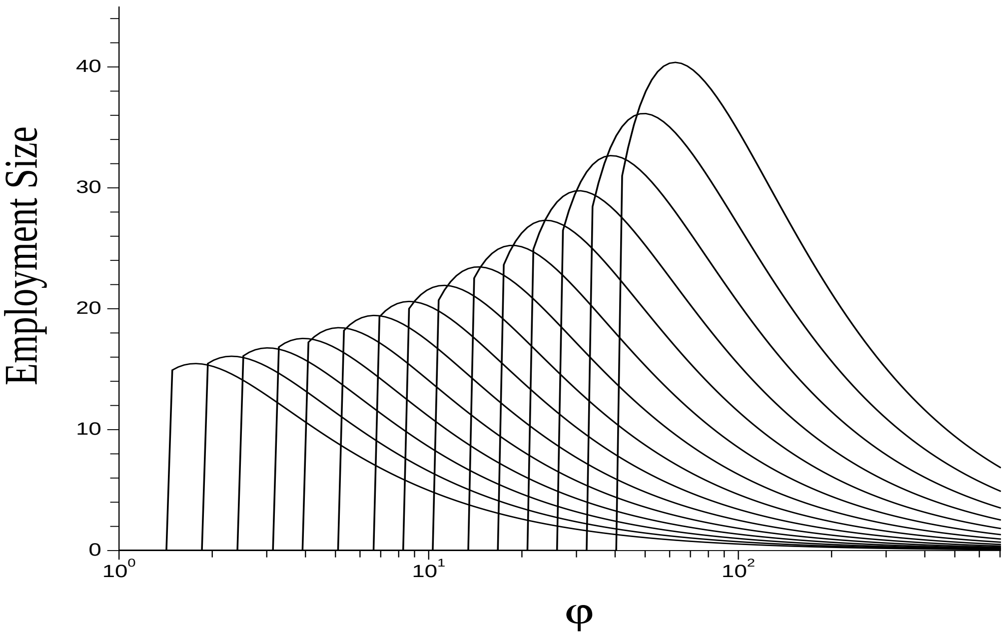
ranging from 100 to 327,397, which covers almost all the market sizes observed in the data[19]. Figure 5.7 illustrates plant behavior at different productivities and market sizes. The simulated values for a selection of market sizes are listed in Table 5.2. The bell-shaped relationship between employment and both revenue and physical output productivity and the effect of market size is demonstrated in plots (a) and (b). As expected, in larger markets, plants can get larger and are more productive on average. The cutoff productivity and number of plants per market are illustrated in Figure 5.8. As an auxiliary observation, it is interesting to note that both ϕ_* and N grow at a slower rate than L [20], although this is somehow due to the fact that the analysis of Section 5.1.3 proves that ϕ_* will eventually hit an upper bound and therefore cannot grow too fast. Below, it will be convenient to approximate the ϕ_* and N relationships by functional forms $a_1 L^{b_1}$ and $a_2 L^{b_2}$. Applying a simple regression model to the simulated data, I estimate

$$\phi_* = 0.229 L^{0.413}, \quad N = 0.092 L^{0.451}. \tag{5.32}$$

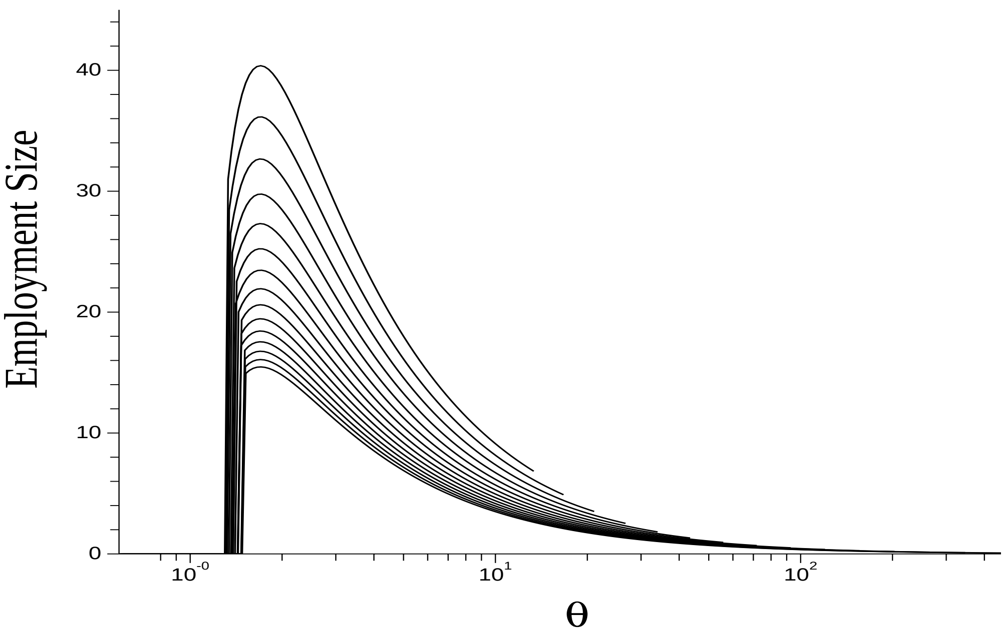
It is also interesting to note that as a consequence of the bell-shaped relationship between employment and productivity, the correlation between employment and productivity is not necessarily positive as in other economic models. Note that the linearity of (5.29) causes the correlation between employment size l and revenue productivity θ within a market to be identical to the correlation using physical pro-

[19]Using very small market sizes caused convergence problem when computing the cutoff productivity. Hence, I am starting the market sizes from 100, above the minimum market size observed in the data.

[20]Notice that the horizontal axis is in log space.

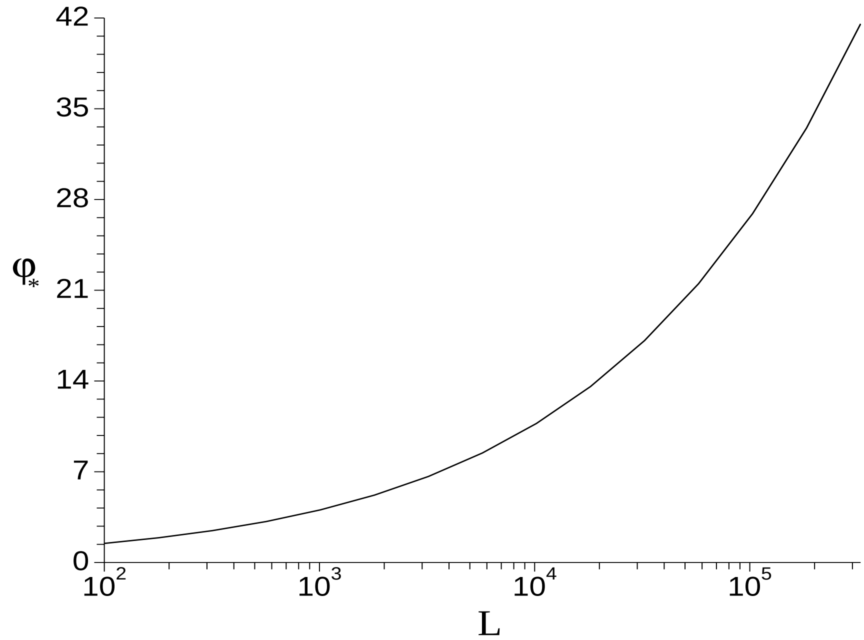


(a)

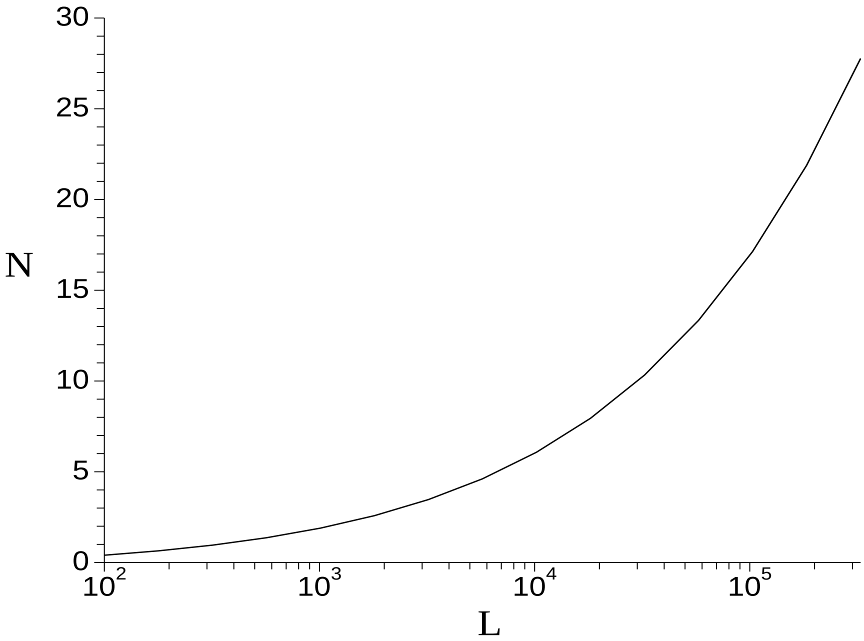


(b)

Figure 5.7: Productivity-employment relationship with CRTS production function.



(a)



(b)

Figure 5.8: Cutoff productivity and number of operating plants with CRTS production function.

L	ϕ_*	N	$corr(\theta, l)$
100	1.473	0.4	-0.469
1,009	4.060	1.9	-0.398
5,721	8.449	4.6	-0.290
57,781	21.502	13.3	-0.050
327,397	41.550	27.8	0.212

Table 5.2: Cutoff productivity, variety measure, and size-productivity correlation by market size.

Population of construction workers	#Obs	corr(rLP,TE)	corr(rTFP,TE)
Any	3970	-0.122	-0.031
$\leq$ 1st Qrtl.	1348	-0.233	-0.084
$\geq$ 3rd Qrtl.	883	-0.141	-0.048

Table 5.3: Correlation between productivity and employment from the data.

ductivity ϕ, therefore only the correlations with revenue productivity are reported. Table 5.2 shows several negative correlations between employment and productivity for the smaller markets, although the correlation increases toward positive values as markets get bigger. This is expected, since larger markets give more productive plants the chance to be larger and still be profitable, hence driving the productivity-employment relationship toward a more monotonic one. To correspond these results to those coming from the data, Table 5.3 lists correlations between rLP and Total Employment (TE) and between rTFP and TE. First row pools plants across concrete industry. To see the effect of market size on the correlations, the next two rows list correlations when selecting plants belonging to the lower and upper quartile of worker population, respectively, representing small and large markets. All correlations listed in the table are negative, and the correlations seem to increase from small to large markets.

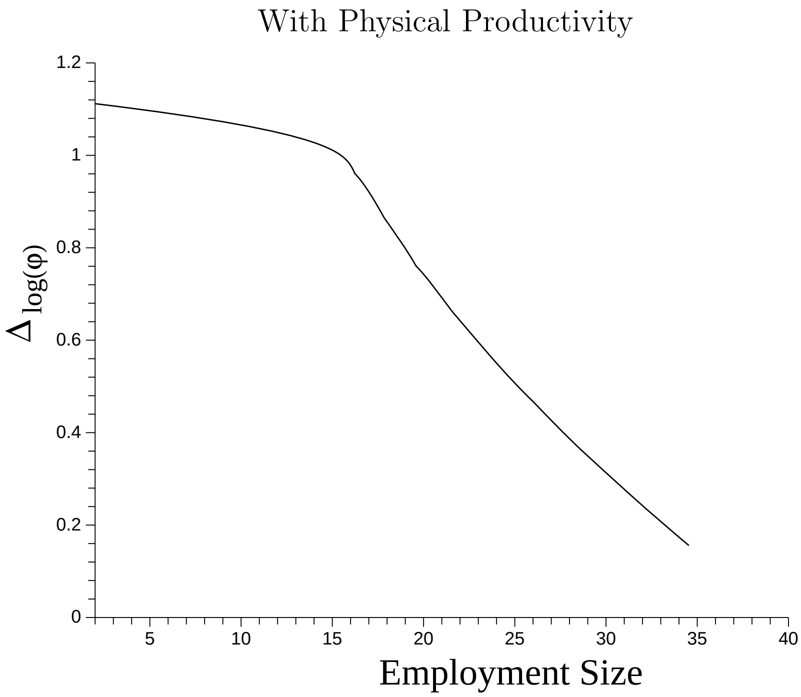
The simulated productivity dispersion curves are numerically obtained by performing a Monte Carlo simulation of productivity and market size draws. Market sizes are drawn from a uniform distribution in the range 100 to 327,397. Draws of ϕ are independently taken from the distribution (5.31). 100,000 random draws of productivity and market size are taken from the resulting joint distribution and productivity dispersions are computed as ranges of productivity for both ϕ and θ and by size classes spaced logarithmically. The curves are illustrated in Figure 5.9 along with the actual rTFP dispersion by employment size in concrete industry. To compare the slopes, the simulated curves are normalized so that they start from the same point as the actual curve. The simulated curve shows a very steep decline at the starting point, where productivities can range from cutoff all the way to the maximum productivity, causing a very large dispersion. After that, the actual and simulated curve almost follow the same slope, suggesting that market localization is able to account for most of the declining productivity dispersion by input size. This fact leaves a uniform additive variation, probably caused by technology or other supply-side frictions, to account for the gap in between the two curves.

5.5 Robustness Tests

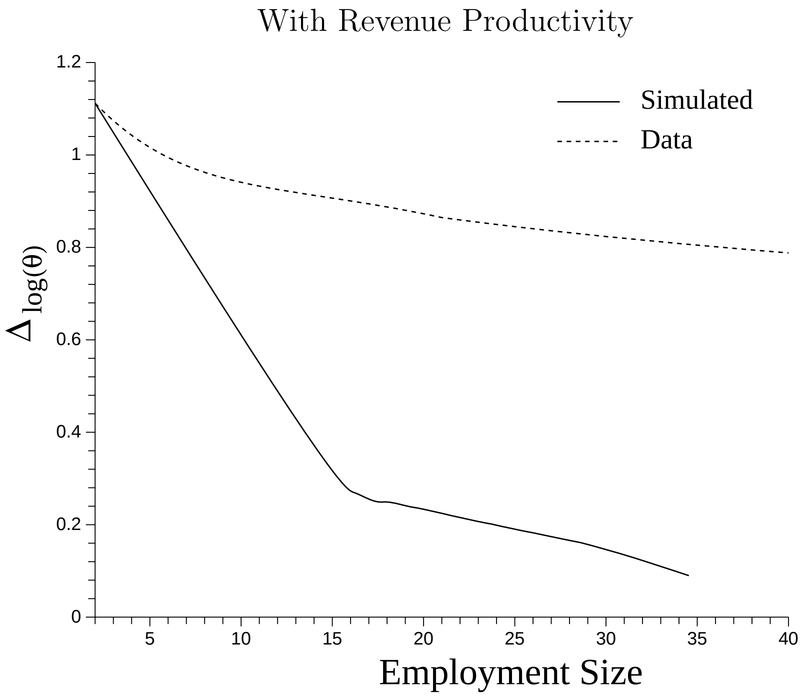
So far the model of Section 5.1 has proven successful in mimicking the behavior of productivity dispersion by employment for the concrete industry very closely. In this section, several different tests are run to check the robustness of model's results to various changes in its setting. More specifically, the effect of wage variations, decreasing returns to scale, use of output as measure of plant size, choice of market size definition, and finally, the degree of market localization will be studied and discussed.

5.5.1 The Effect of Wage Variations

A fixed wage within a market is justified in the absence of worker skill heterogeneity when homogeneous workers are mobile within a market. In equilibrium, wages will level across producers to make workers indifferent between staying with



(a)



(b)

Figure 5.9: Plots of productivity dispersion by employment size with CRTS production function.

their current employers or changing jobs. However, ruling out worker mobility across markets creates different labor supply and demand curves in markets of different sizes. The most likely outcome is wages that vary by market size. The sensitivity of the model behavior is tested by letting wages vary with the logarithm of market size as shown below.

$$w = w_0(1 + \Delta_w \log(L)), \tag{5.33}$$

where w_0 is an offset wage and Δ_w is a non-negative variation factor. The effect of this wage variation on the performance of plants can cause changes in cutoff productivities and the number of operating plants across markets as the cost of operation now varies from small to large markets. That will be the main effect causing differences in how productivity dispersion falls with employment.

In the data, the smallest market is Yazoo, Mississippi with a population of 48 construction workers, and the largest one is New York, New York with a population of 327,397 construction workers. The US Bureau of Labor Statistics (BLS) reports the 2006 mean annual wage of production workers in Yazoo and New York areas to be \$27,880 and \$31,430, respectively[21]. This amounts to a roughly 13% wage difference. I will simulate the model with a 5%, 15%, and 25% maximum wage difference to cover a range of possible variations. Note that from (5.33) the total percentage difference in wages across markets (Δw) relates to Δ_w in the following way:

$$\Delta_w = \frac{\Delta w}{\log(L_{max}) - (1 + \Delta w)\log(L_{min})}. \tag{5.34}$$

I am using the number of construction workers from Yazoo and New York as L_{min} and L_{max}, respectively. The offset wage w_0 is chosen so that the average simulated wage across markets weighted by number of plants in each market is equal to the previously estimated w. It is important to note that the number of operating plants

[21]BLS reports wags for metropolitan areas only. The closest metropolitan area to Yazoo is Jackson, whose average annual wage is used here.

Variation of Wages	Estimated Δ_w	Estimated w_0	Min. Wage	Max. Wage
5%	0.006	1.059	1.083	1.137
15%	0.018	0.929	0.995	1.144
25%	0.032	0.819	0.920	1.150

Table 5.4: Estimated wage variation parameters.

itself is determined endogenously by wage and market size. Therefore, I take a two step recursive approach to get an estimate for w_0. For simplification, I assume that the distribution of market sizes is uniform which helps me formulate the step 1 estimate of w_0 as

$$w_0 \int_{L_{min}}^{L_{max}} (1 + \Delta_w \log(L))N(L)dL = w \int_{L_{min}}^{L_{max}} N(L)dL. \tag{5.35}$$

Using $N = a_2 L^{b_2}$ as an approximation to $N(L)$, (5.35) can be solved analytically to yield

$$w_0 = \frac{w\big(L_{max}N(L_{max}) - L_{min}N(L_{min})\big)}{L_{max}N(L_{max})\left(1 + \Delta_w\left(\log(L_{max}) - \frac{1}{b_2+1}\right)\right) - L_{min}N(L_{min})\left(1 + \Delta_w\left(\log(L_{min}) - \frac{1}{b_2+1}\right)\right)}. \tag{5.36}$$

The initial values of a_2 and b_2 are picked from (5.32) and w_0 is computed from (5.36) for a given Δ_w. In step 2, the estimated w_0 is used to find the relationship between N and L when the wage varies according to (5.34), and new estimates of a_2 and b_2 are computed that are plugged back into (5.36). By repeating these two steps recursively, the method converges very fast and provides a stable estimate of w_0. For the three levels of variation used in the test runs, the estimated values for Δ_w and w_0 are listed in Table 5.4.

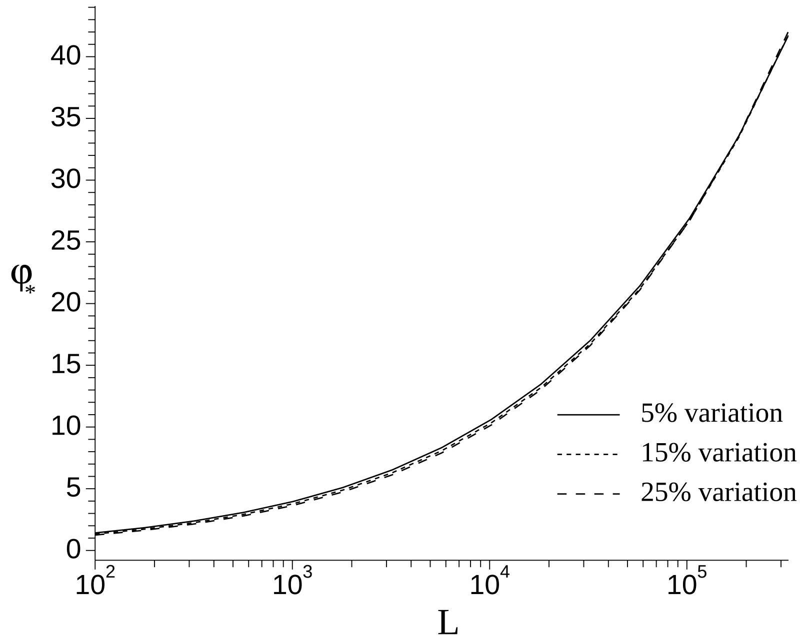
Having a full description of the wage equation, I can examine the effect of different levels of wage variations on the results. Figures 5.10 and 5.11 illustrate the results for a CRTS production function. The model shows very strong robustness to even large differences in wages. Since operating in larger markets is now costlier than operating in smaller markets, the cutoff productivity is expected to rise more sharply with L. This is observed in simulation results, except that the difference is not remarkably large. The same thing happens for the relationship between the number of plants and market size, where wage variation does not seem to play a significant part. The differences are more magnified when the dispersion of ϕ is plotted against employment. However, the effect on the dispersion of θ is not significant, so that, except at the starting point in the curve, the simulated curve still has a slope close to the empirical one. Thus, the predictive power of the model does not change when wage variation is added.

Note that, similar to the wage, the theoretical model assumes that fixed cost of operation f is also constant within and across markets. The effect of varying fixed cost across markets will act in the same way as varying wage. Higher fixed costs in larger markets raise the cutoff productivity and drive some plants out of the market. Therefore, I conjecture that the impact of varying fixed costs is similar to that of varying wages, and for that reason, I will not proceed with separate simulation of varying fixed costs.

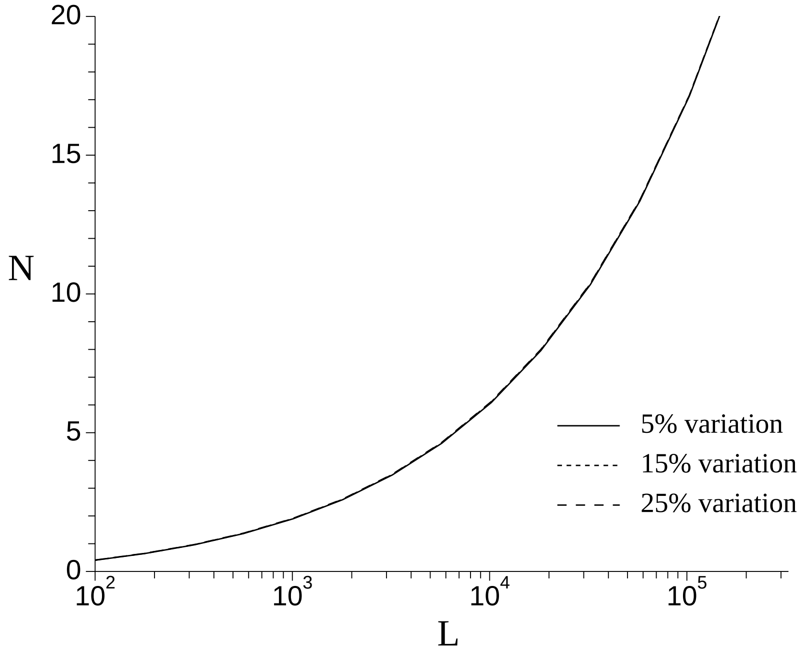
5.5.2 The Decreasing Returns Case

The first-order condition (5.7) can also be solved analytically when $\nu = 0.5$. Industries are not believed to operate with such low returns to scale production. But, with the continuity of solutions as ν changes, the behavior of the industry with $\nu = 0.5$ ensures that the model implications are still in place when return to scale are slightly below 1. With $\nu = 0.5$ the solution to (5.7) is

$$q_j = \frac{\alpha\gamma + \eta N \bar{p}}{\gamma + \eta N} \frac{L\phi_j^2}{2(\gamma\phi_j^2 + Lw)}. \tag{5.37}$$

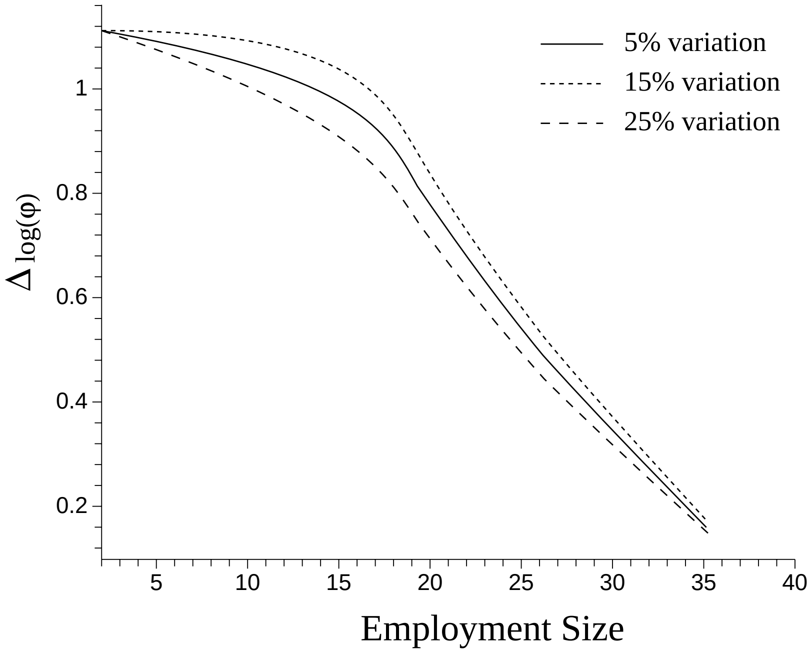


(a)

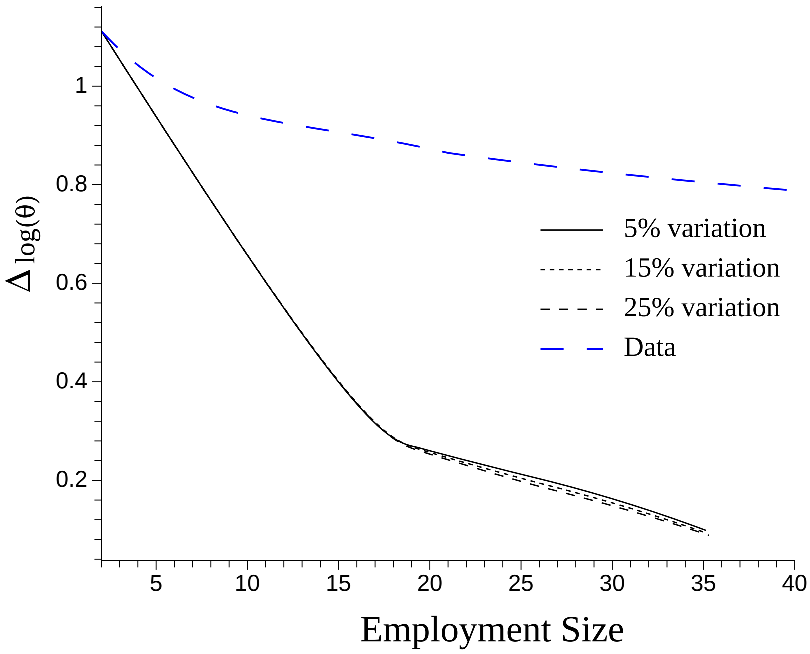


(b)

Figure 5.10: Cutoff productivity and the number of plants per market when wage variations are present and with CRTS production function.



(a)



(b)

Figure 5.11: Plots of productivity dispersion by employment size when wage variations are present and with CRTS production function.

Plugging (5.37) into the profit function (5.6) and some algebraic simplification of the profit function yields

$$\pi_j = \frac{1}{2}\frac{\alpha\gamma + \eta N\bar{p}}{\gamma + \eta N} q_j - f. \tag{5.38}$$

The cutoff productivity is found by setting the above profit function to zero. Define

$$A^2 = \frac{4f(\gamma\phi_*^2 + Lw)}{L\phi_*^2}, \quad K = \frac{Lw}{\gamma\phi_j^2 + Lw}.$$

Again, all the plant performance measures can be expressed as a function of the cutoff productivity embodied in A :

$$\begin{aligned}
q_j &= \frac{AL\phi_j^2}{2(\gamma\phi_j^2 + Lw)}, && (5.39)\\
p_j &= \frac{A}{2}\left(1 + \frac{Lw}{\gamma\phi_j^2 + Lw}\right), && (5.40)\\
l_j &= \frac{A^2L^2\phi_j^2}{4(\gamma\phi_j^2 + Lw)^2}, && (5.41)\\
\pi_j &= \frac{fLw(\phi_j^2 - \phi_*^2)}{\phi_*^2(\gamma\phi_j^2 + Lw)}, && (5.42)\\
N &= \frac{2\gamma(\alpha - A)}{\eta A(1 - \bar{K})}, && (5.43)
\end{aligned}$$

where l is the employment size. Using the definition (5.5), the revenue productivity θ_j can be expressed as

$$\theta_j = \frac{\gamma}{L}\phi_j^2 + 2w. \tag{5.44}$$

Looking at (5.44), again it is clear that a higher elasticity of demand (lower ν) results in an under-representation of productivities in revenue terms. Therefore, using revenue productivities will result in steeper productivity dispersion curves.

Noting that the concrete industry shows returns to scale very close to 1, estimating model parameters using $\nu = 0.5$ has more rhetorical than practical value. Using (5.41) and (5.43) to generate the simulated relationships between employment

α	η	γ	ϕ_M	w	f_E	f
1.609	2.631	0.483	1672.64	0.629	15.191	6.617
Stage 1 estimation error					$\sigma_{error,1} = 4.148$	
Stage 2 estimation error					$\sigma_{error,2} = 0.721$	

Table 5.5: Parameter estimates with decreasing returns to scale production function.

L	ϕ_*	N	$corr(\phi, l)$	$corr(\theta, l)$
100	5.464	0.6	-0.378	-0.317
1,009	14.321	2.8	-0.094	-0.209
5,721	27.450	6.4	0.360	0.108
57,781	53.022	14.2	0.944	0.810
327,397	67.128	19.1	0.980	0.984

Table 5.6: Cutoff productivity, variety measure, and size-productivity correlation by market size and with decreasing returns to scale production function.

size and productivity and between market size and the number of operating plants, the model parameters are estimated in this case as listed in Table 5.5. A look at the estimated parameters shows that fixed costs of operation are lower and maximum productivity is higher than estimated with CRTS production. With decreasing returns to scale production, giving larger plants a production disadvantage, the estimated parameters have moved in the right direction. A list of simulated moments is also listed in Table 5.6. Notice that because of the nonlinear relationship between ϕ and θ, the correlations of those productivity measures with size are not identical as in the CRTS case. The behavior of plants is illustrated in Figures 5.12 and 5.13.

To generate productivity dispersion, a Monte Carlo simulation is performed by drawing 100,000 random samples from the same distribution in Section 5.4.2, and productivity dispersion is computed for each employment class. The dispersions of ϕ and θ by employment are shown in Figure 5.14. The findings are consistent with what was observed in Section 5.4.2. Specifically, the slope by which productivity dispersion falls with employment is again much the same as the actual slope in the data, except at the starting point. The analysis of this section was conducted to show that the model's performance and qualitative implications do not change under

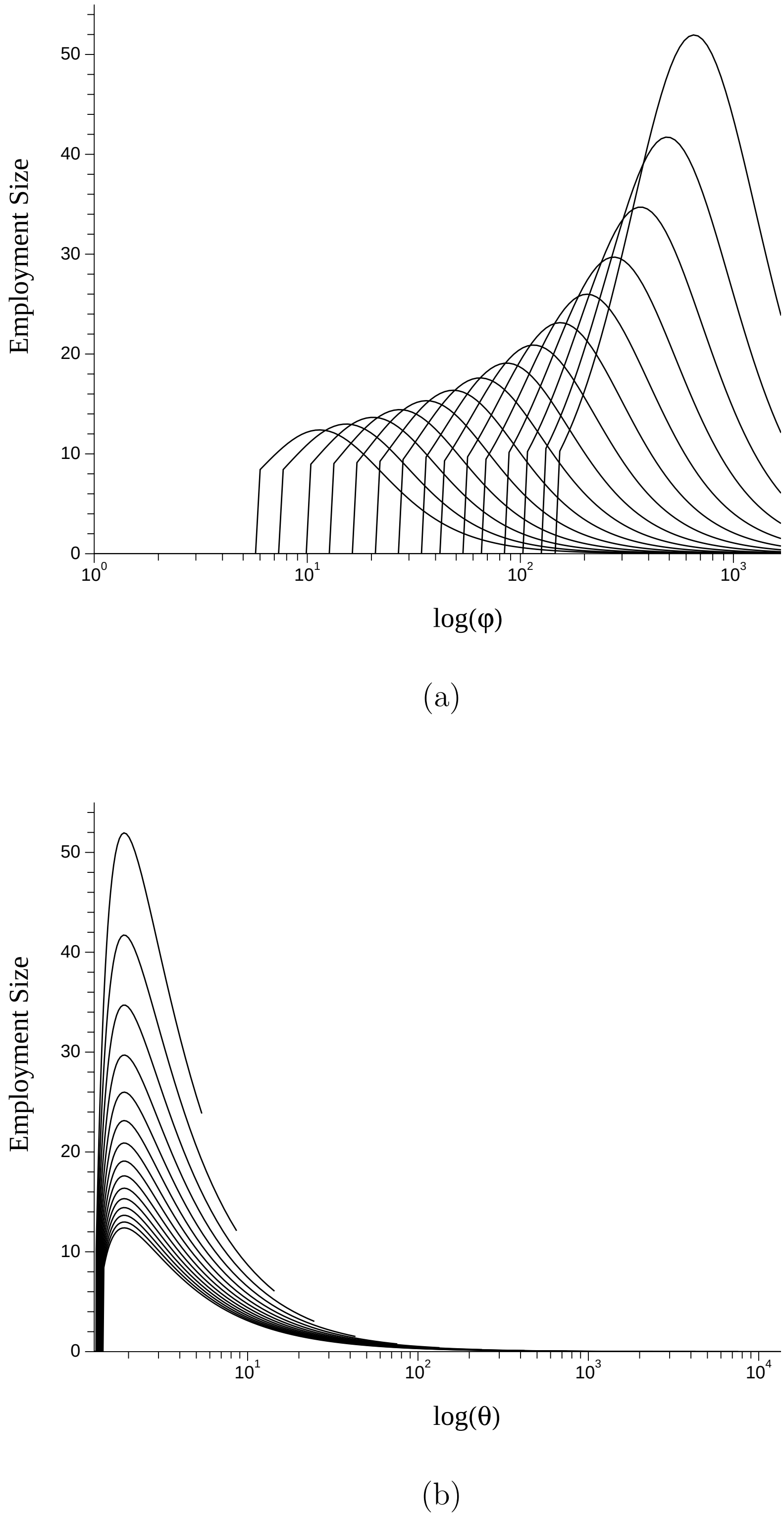
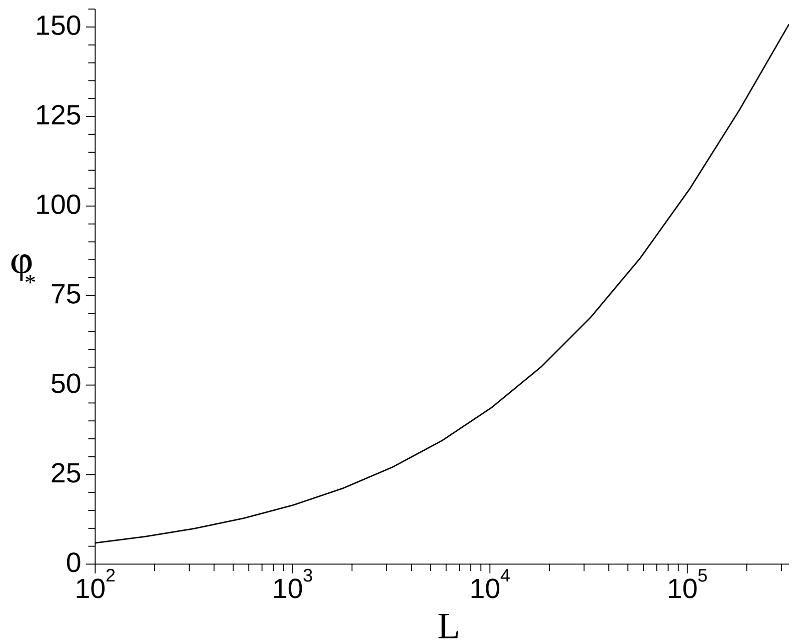
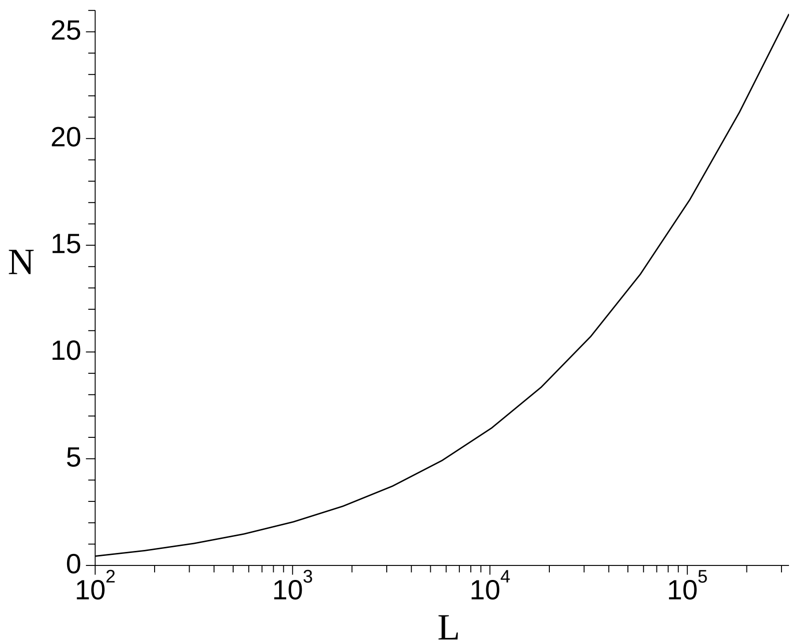


Figure 5.12: Productivity-employment relationship with decreasing returns to scale production function.



(a)



(b)

Figure 5.13: Cutoff productivity and the number of operating plants with decreasing returns to scale production function.

decreasing returns to scale production.

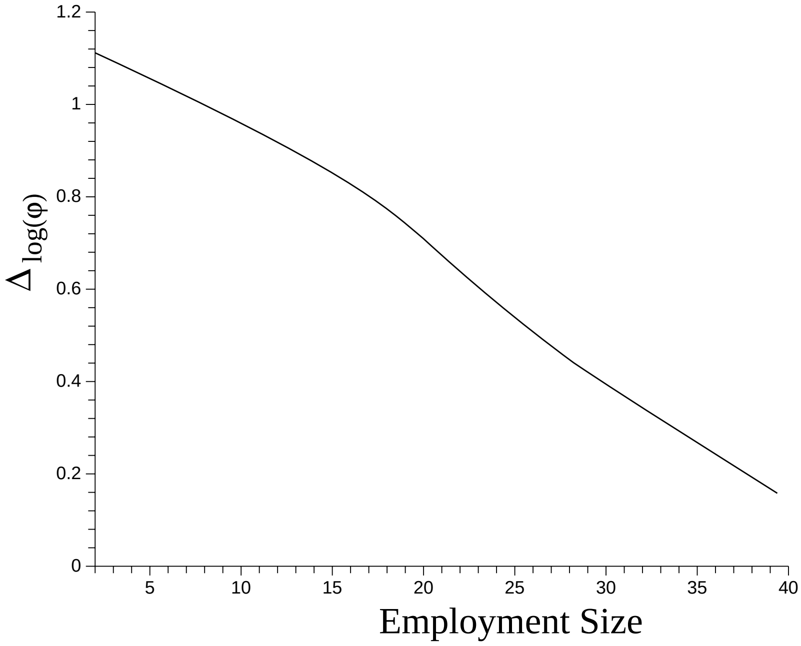
5.5.3 Measuring Market Size

So far, the discussion of market size relied on the population of construction workers, which strongly relates to the scale of construction activity in a CBSA. Later, I want to be able to compare the behavior of the concrete industry with that of other 4-digit industries, and for that, I will need a more universal measure of demand size, namely the resident population of CBSA. Hence, it is useful to know whether the relationship between employment and productivity, which affects the shape of productivity dispersion and employment relationship, shows sensitivity to such change in the choice of market size.

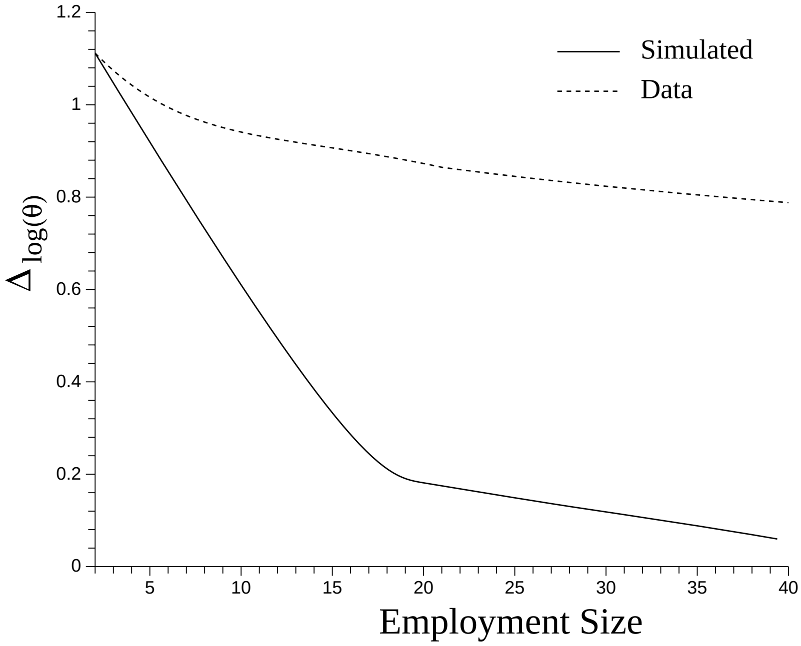
Moreover, because of high transportation costs, physical distances are important in determining the market reach for concrete. Population densities, either with resident or worker population, take account of the physical extent of an urban area and demand concentration. However, these measures do not provide an estimate of the actual demand size in the absence of information on the shipment radius at individual level, as a result, I had to rely on worker population to calibrate my model. I will show the insensitivity of this shift by estimating the productivity-employment relationship using population densities.

The Census Bureau's City and County Databook provides information on county population and land areas. When aggregated to CBSA level, three other measures of market size can be defined for a CBSA: resident population, resident population density, and construction worker population density. Summary statistics for each market definition is listed in Table 5.7.

For each of the above mentioned market definitions, I estimate the relationship between productivity and employment for the corresponding 10 and 90 percentile market sizes. The choice of those market sizes for demonstration provides a sense of



(a)



(b)

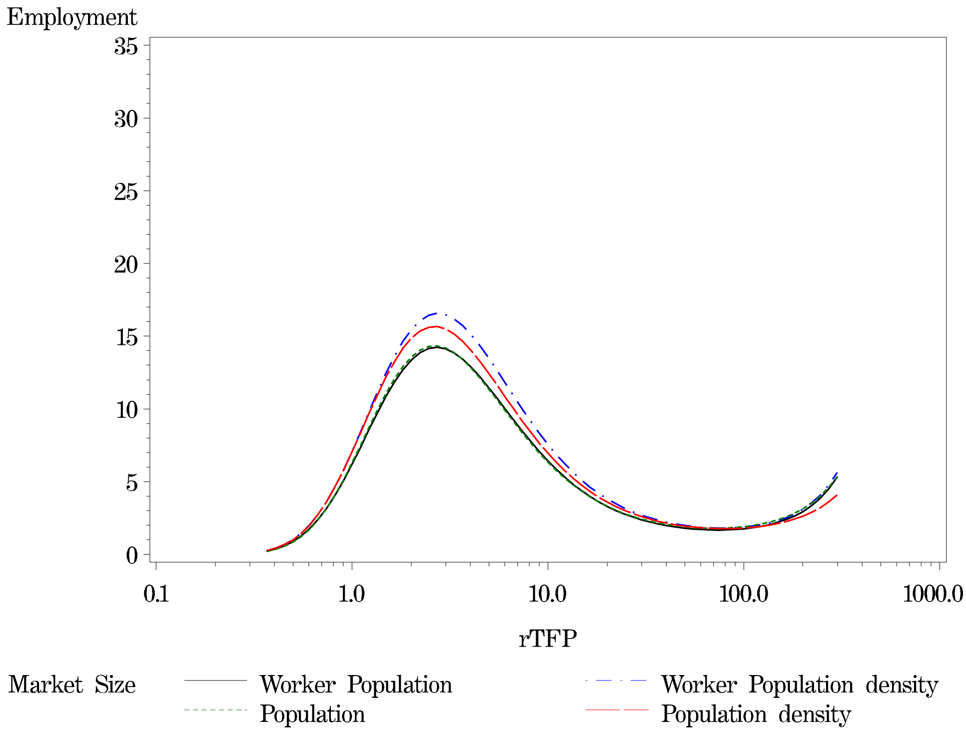
Figure 5.14: Productivity dispersion by employment with decreasing returns to scale production function.

	Population of Construction Workers	Population Density of Construction Workers	Population	Population Density
Mean	43,173.4	9.14	2,541,081.9	534.4
Std.Dev.	58,022.6	9.79	3,793,412.7	632.2
Min.	48	0.04	12,457	3.6
Median	16,600	5.99	924,786	301.5
Max.	327,397	48.76	18,747,320	2792.2

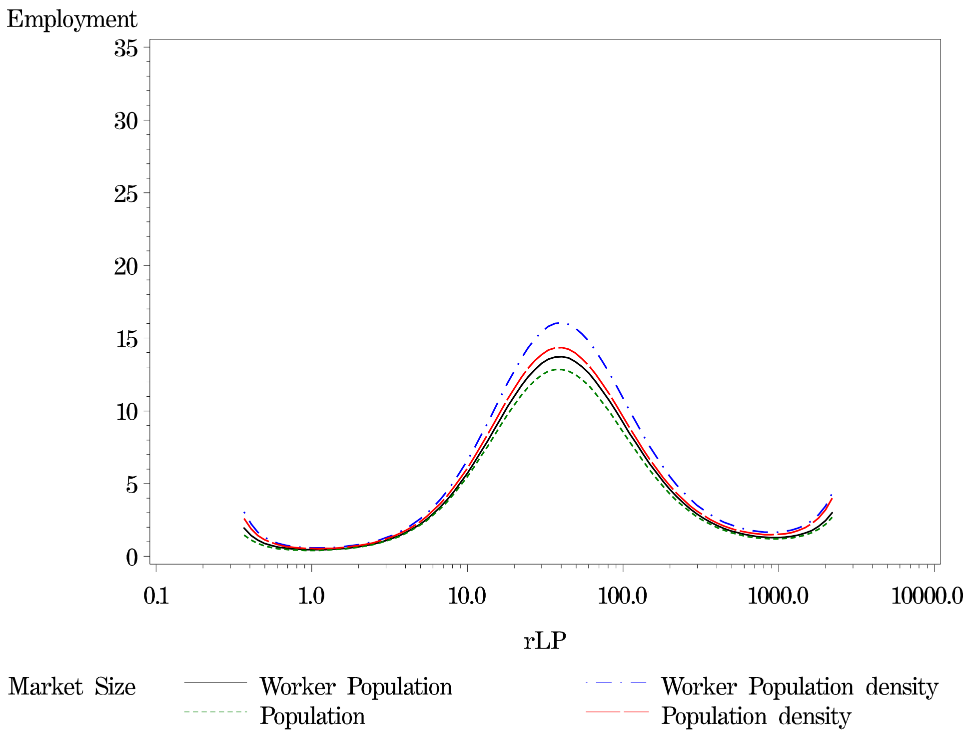
Table 5.7: Summary statistics for different definitions of market size.

full range of sizes and productivity dispersions achievable by each market definition, making it easy to picture the effect of each market definition on the slope by which productivity dispersion falls with employment. Again, to reduce computational burden, I classify market sizes by rounding the log of populations to the nearest 0.5 and rounding the log of population densities to the nearest 0.1.

Estimation results for 10th percentile market sizes are shown in Figure 5.15, and results for 90th percentile market sizes are shown in Figure 5.16. The fact that remains unchanged is that the bell-shaped relationship between productivity dispersion and employment and the effect of market size are the same no matter which definition of market is used. Interestingly enough, using population of construction workers or residents does not seem to really matter as the estimated plots almost overlap. The same thing can be said about population densities. Overall, the estimated productivity-size relationship with either of the market definitions are similar in both their shape and range of values. Since the range of productivities at each employment level is a direct outcome of the estimated bell-shape, the dispersions of productivity by employment size will be almost the same at large markets and small markets, irregardless of which of the four market definitions is used. Therefore, the average slope by which productivity dispersion declines, i.e. the slope of the line that connects dispersions in small and large plants, should not vary significantly when changing market definition among the four presented here.

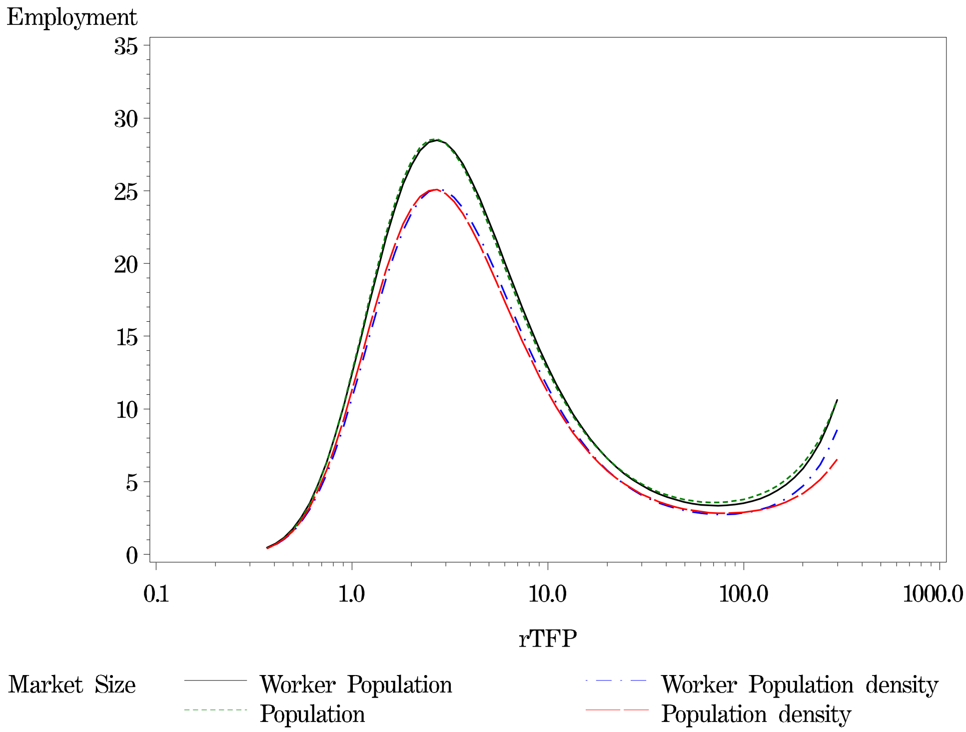


(a)

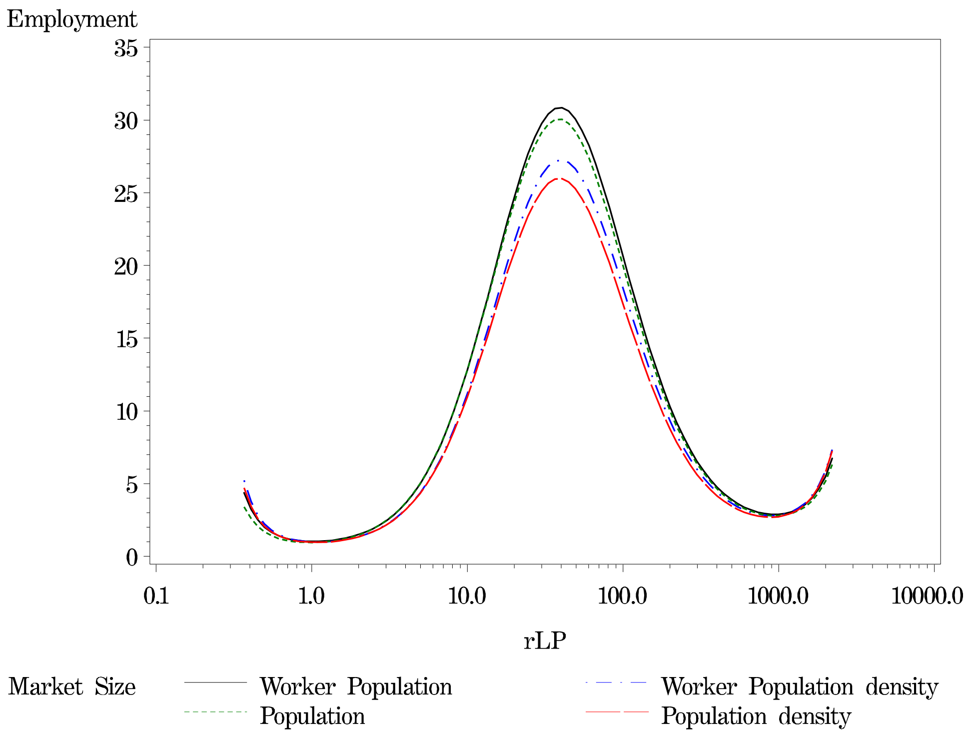


(b)

Figure 5.15: Comparing different market definitions at 10 percentile market size.



(a)



(b)

Figure 5.16: Comparing different market definitions at 90 percentile market size.

L	ϕ_*	N	$corr(\theta, pq)$
100	1.473	0.4	0.213
1,009	4.060	1.9	0.371
5,721	8.449	4.6	0.477
57,781	21.502	13.3	0.616
327,397	41.550	27.8	0.718

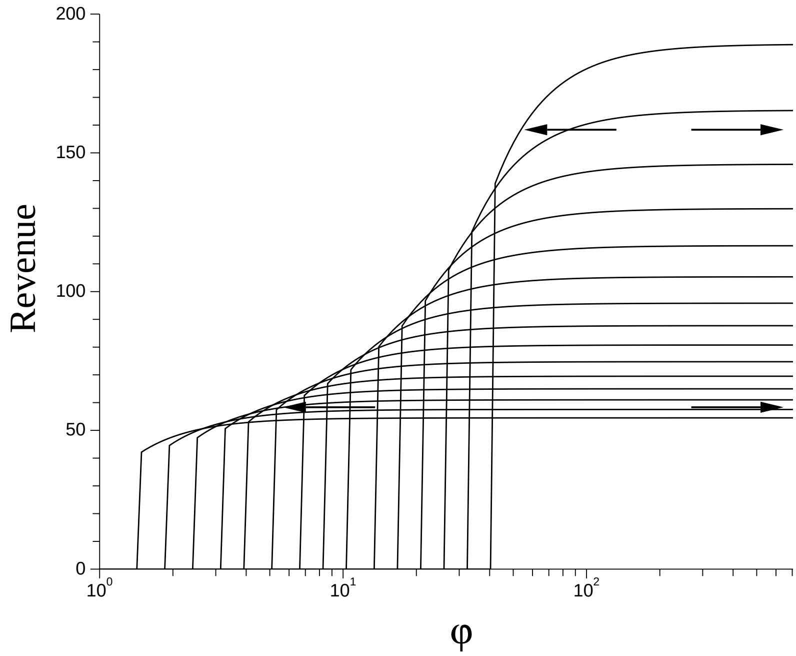
Table 5.8: Cutoff productivity, variety measure, and output-productivity correlation by market size.

5.5.4 Input versus Output Size

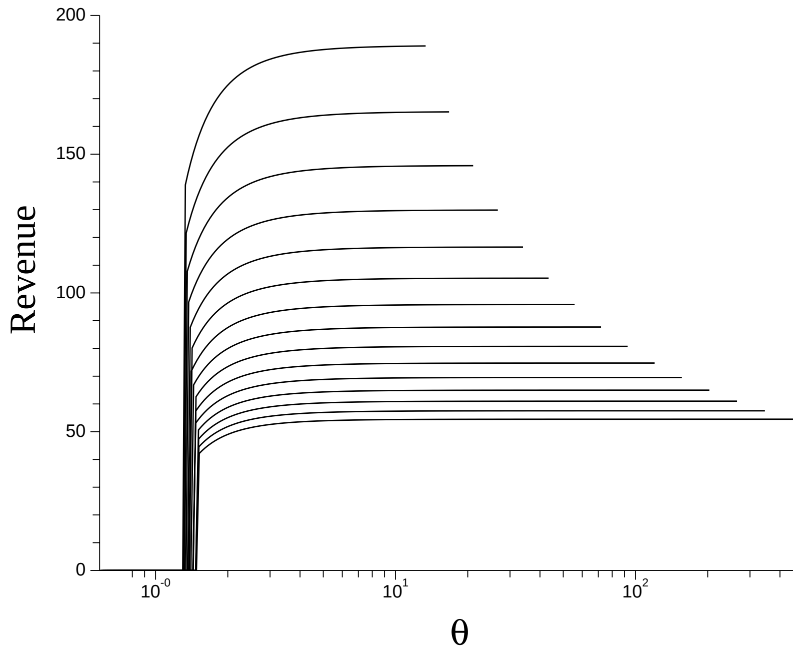
In the literature, both input and output sizes have been used to compare plants' scales of operation. So far, all my discussions have been based on input size and specifically the size of employment. In this section I will look at both empirical and theoretical implications when using output instead. For clarification, output is defined as deflated shipment value in the data and equivalently as the simulated revenue $p_j q_j$ in the model. Using (5.24) and (5.25), the resulting relationship between this revenue and productivity is monotonic and positive as illustrated in Figure 5.17.

The theoretical and empirical behavior of output with productivity and market size can be compared using correlations and also by estimating a model similar to (5.15). Tables 5.8 and 5.9 report correlations between revenue productivity and output from the simulation and from the data, respectively. As discussed earlier, using physical or revenue productivity produce identical theoretical correlations with CRTS production function, therefore, only the results with revenue productivities are reported. Similar to theory, the data correlations also show a positive relationship between output and productivity for small and large markets, though the correlations are not as strong as those from the theory.

I examine the details about the relationship between output and productivity by estimating a semi-parametric model similar to (5.15) in which employment size is replaced with real output from the data. Figure 5.18 illustrates the estimated relationships. The previously predicted productivity-employment relationships are



(a)



(b)

Figure 5.17: The theoretical relationship between revenue and productivity. Arrows show the range of productivities present at different revenue levels.

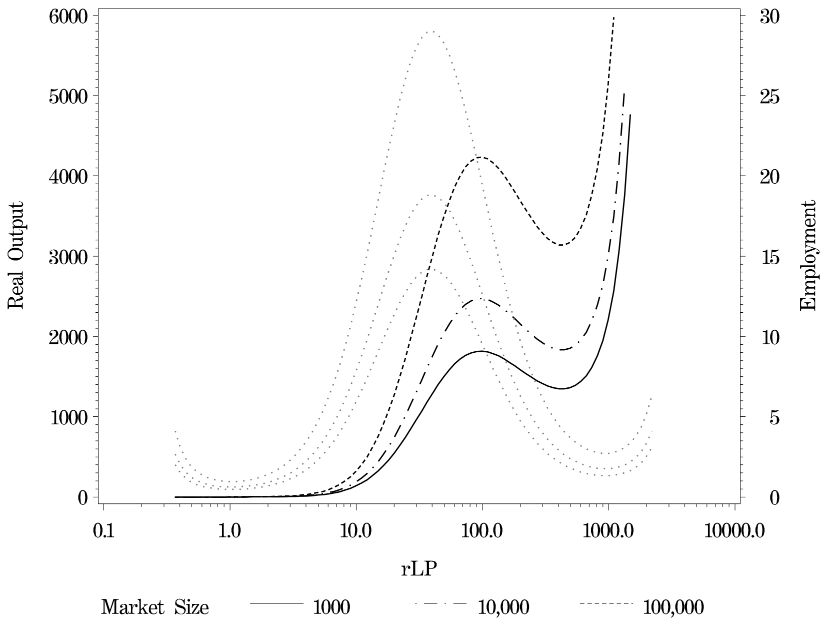
Population of construction workers	#Obs	corr(rLP,Q)	corr(rTFP,Q)
Any	3970	0.196	0.143
$\leq$ 1st Qrtl.	1348	0.136	0.120
$\geq$ 3rd Qrtl.	883	0.203	0.110

Table 5.9: Correlation between productivity and output (Q) from the data.

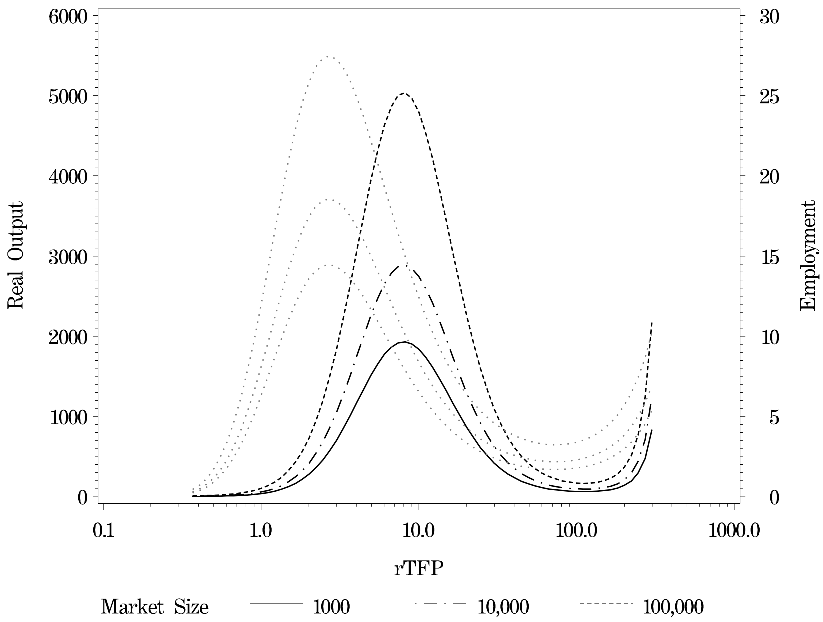
also plotted in the same picture (gray dotted lines) for a better understanding of input size and output size correspondences as seen in the data. The figure shows that the predicted relationship between output and rLP is rather monotonic, though not in a strict sense, and the estimated relationship justifies the positive correlation seen in the data and predicted by the theory.

On the other hand, the predicted relationship between output and rTFP is bell-shaped. However, one must be careful in interpreting this result as a contradiction to the theoretical prediction. With the productivity-employment curves present in the same plot, it is easy to see that output increases for a considerable range of rTFP for which the employment both rises and then falls. In fact, the rTFP at which the peak of output curve happens is actually about five times higher than where the employment curve peaks. It is useful to remark that some very highly productive concrete plants in the data are so because their listed real capital or energy and material consumption is very small or close to zero, but their employment information is more precise. With this fact in mind, it is possible that those plants could be playing a role in causing a declining upper tail in the predicted productivity-output curves, while not affecting the relationship with rLP.

The monotonicity of the relationship between productivity and output in theory implies zero productivity dispersion at a given revenue level at a given market size. However, Figure 5.17 shows that when a continuum of market sizes are present, it is possible to produce productivity dispersions that are nonzero and declining with revenue level.



(a)



(b)

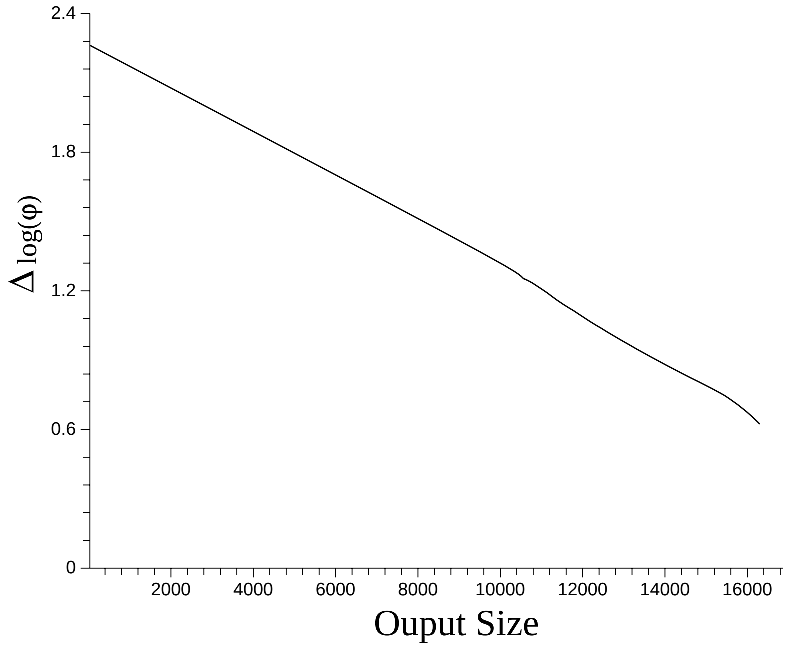
Figure 5.18: The predicted relationship between output and revenue productivity from the data. The predicted productivity-employment relationship is also plotted with gray dotted line.

Figure 5.19 shows the dispersion of productivity by output size, a counterpart of Figure 5.9 but using output size to classify concrete plants. The starting point of the simulated dispersions are normalized to be equal to the first point in the data. In addition, the simulated revenue is not up-to-scale with data values since prices in the model are normalized to the price of a numeraire consumption good. For that reason I re-scale all the simulated revenues so that their maximum coincides with the maximum in the data.

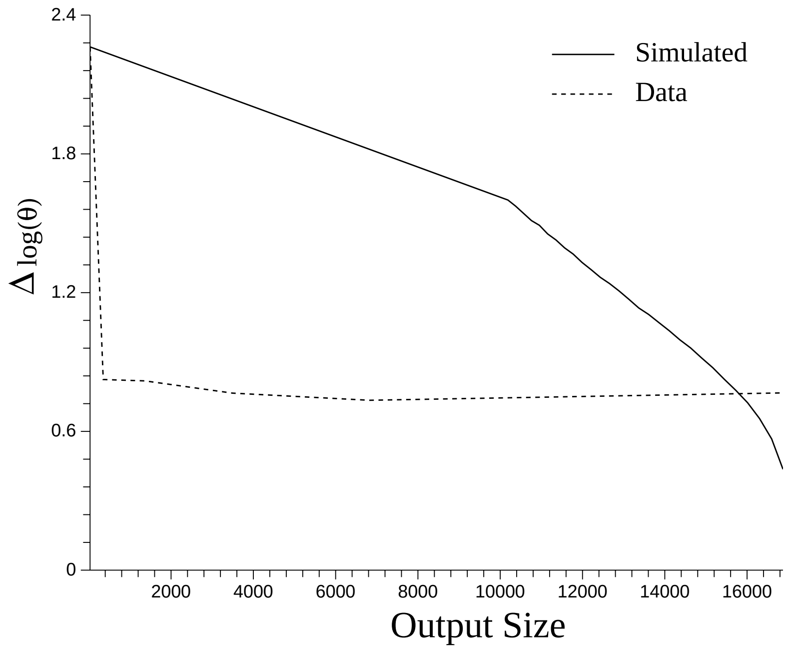
The picture shows that productivity dispersion is practically uniform across different output sizes. In contrast, the model again predicts a falling productivity dispersion with employment size in the same way as it did with input size. Notice that the model was not meant, nor calibrated, to mimic the behavior of productivity dispersion with output size. So far, the results of this section remain the main point of deviation between the model and the data implications.

5.5.5 The Effect of Market Structure

When dealing with the concrete industry, two characteristics of the industry had direct effect on the shape of the bell that was estimated as the relationship between productivity and employment. The spatial diversity of concrete output ensured that the range of productivities that are present is large, and an average shipping distance of 82 miles (from 1997 CFS report) caused many productive plants to be small. Changing any of these assumptions will change the shape of the bell in some way or the other. I investigate the effect of market structure on the formation of the bell-shaped relationship by looking at two other 4-digit industries: Manufactured Ice (SIC 2097) and Roasted Coffee (SIC 2095). Both of these industries have rather homogeneous outputs, putting them on par with concrete in reducing the effect of diversity in driving revenue productivities. In addition, each of these industries represents a situation where one of the model assumptions is relaxed.



(a)



(b)

Figure 5.19: Plots of productivity dispersion by revenue levels.

Manufactured ice is a very localized-market industry with an average shipment radius of 35 miles[22]. However, due to low cost of transportation, the output is not as spatially differentiated as in concrete. Under this situation, the range of productivities that survive and stay in the market must be narrower. At the same time, due to the localization of their markets, the ice plants are expected to have a low average employment size. In case of roasted coffee, output is shipped 183 miles away on average, qualifying as an industry with broader market. Here, average sizes are expected to be larger in response to a farther reaching market.

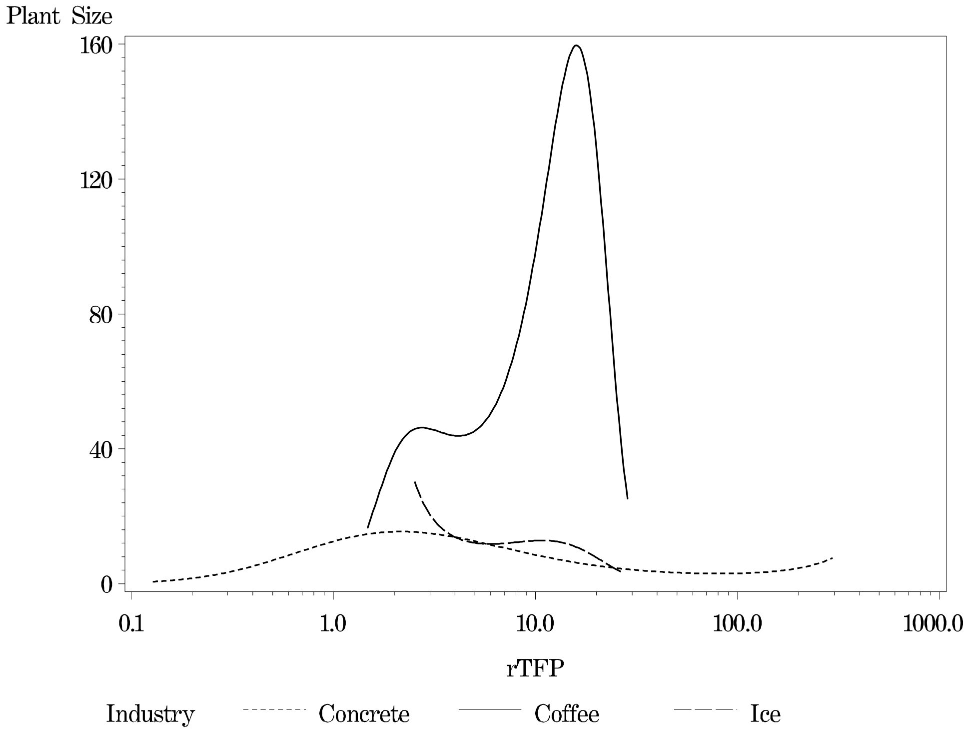
Summary statistics for each of these industries is shown in Table 5.10. The listed correlations between productivity and employment for each industry provide an early look into the behavior of each industry. Most notably, the correlations are positive for the coffee industry, where employment size is expected to be large for more productive plants.

The estimation results for each industry, along with concrete as control group, are shown in Figure 5.20. The results for rTFP and rLP are somewhat different, especially for the coffee industry. The ice industry shows a bell-shaped relationship between its productivity and employment similar to that of concrete, but with a much narrower range of operating productivities, most likely as a result of product substitutability. In the coffee industry, the average size is obviously higher and the range of productivities is also more limited in favor of more productive plants. The estimate with rTFP shows a bell-curve whose tip has moved in the upper-right direction as a result of more expanded demand market. With rLP, the picture is rather different, but still the average employment size and average productivity have both increased.

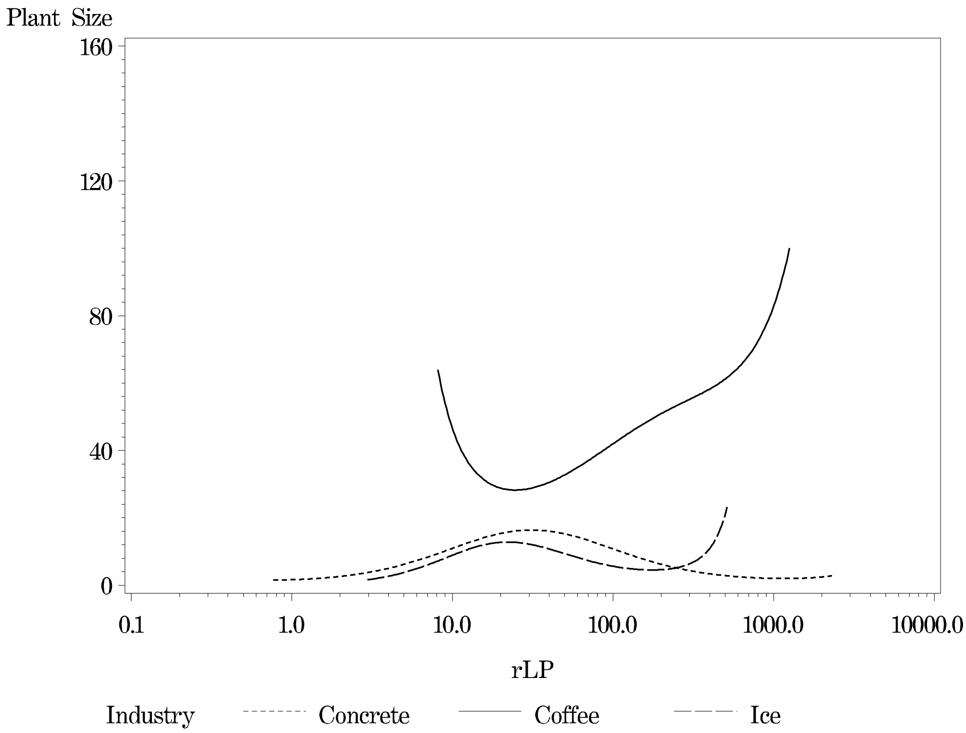
[22]The shipment distances are according to 1997 commodity flow survey.

Statistics	Concrete (SIC 3273)	Ice (SIC 2097)	Coffee (SIC 2095)
#Obs	3970	128	233
Mean Employment	20.8	15.8	102.1
Std.Dev. Employment	25.3	11.5	128.8
Max. Employment	513	105	999
Mean rTFP	1.574	2.112	1.095
Std.Dev. rTFP	0.285	0.437	0.265
Min. rTFP	-2.084	0.839	0.421
Max. rTFP	5.842	3.265	3.557
Mean rLP	4.033	3.291	4.898
Std.Dev. rLP	0.675	0.641	0.946
Min. rLP	-0.326	1.300	2.109
Max. rLP	7.710	6.473	7.167
All Markets			
corr(rLP,TE)	-0.122	-0.094	0.179
corr(rTFP,TE)	-0.031	-0.182	0.123
Population≤1st Qrtl.			
corr(rLP,TE)	-0.250	-0.402	0.103
corr(rTFP,TE)	-0.099	-0.350	-0.063
Population≥ 3rd Qrtl.			
corr(rLP,TE)	-0.177	0.005	0.295
corr(rTFP,TE)	-0.060	-0.225	0.103

Table 5.10: Summary statistics on concrete, ice and coffee.



(a)



(b)

Figure 5.20: Comparing productivity-employment relationships among concrete, ice, and coffee.

5.6 Beyond Ready-Mix Concrete

In light of Proposition 4, industries with completely globalized markets, i.e. no limit on demand and no trade or transportation costs, should have their employment size increase monotonically with productivity. In practice, these conditions are hardly satisfied for any industry. In addition, plants belonging to an industry are still heterogeneous in their degrees of market reach, so that a large average shipment distance does not necessarily exclude possibility of localized trade among some plants. However, industries with higher average shipment distances are expected to show higher average productivities and larger average sizes for their plants.

In this section, I primarily look at different classes of industries with respect to their average shipment distance and investigate the qualitative relationship between average shipment distance and the average productivity and size of plants. These results are meant to be complementary to those of the last section. While the results of the last section are more exact in the sense that product diversity was not a major issue there, the results of this section will extend the concept to more industries and show the universality of the implications.

In my first experiment, I focus on localized-market industries by pooling plants from those ones whose products are shipped on average no more than 100 miles away. The list includes 18 industries with a total of 18,529 plants (Appendix A). This experiment will further test the fact that the bell-shaped relationship between productivity and employment is not a peculiarity of the concrete or ice industry, but is common among industries whose markets are primarily localized.

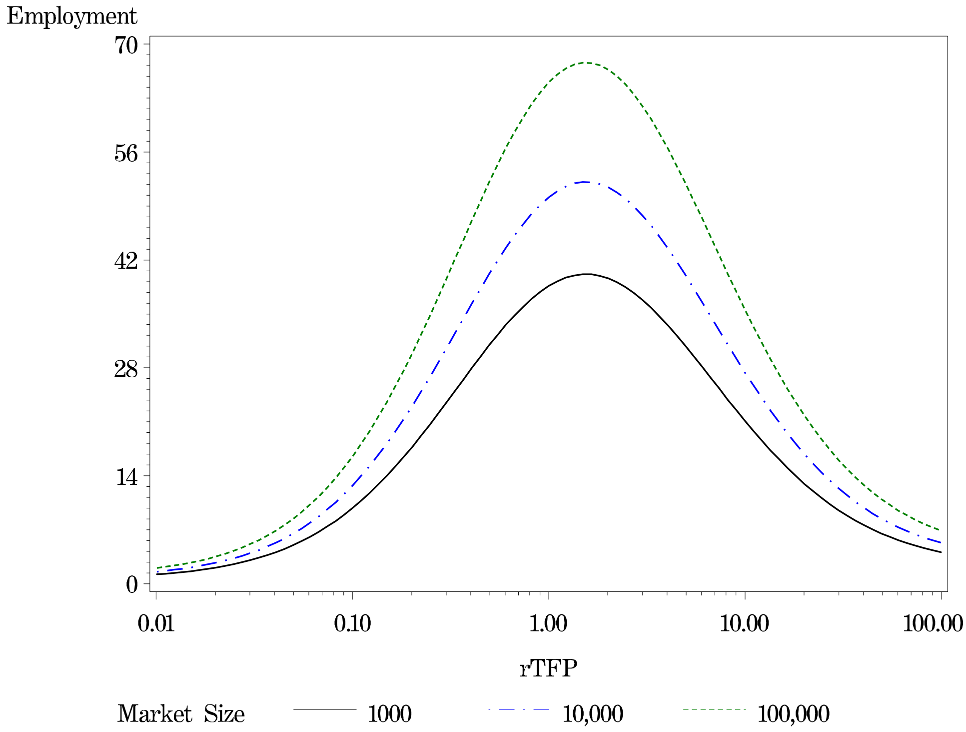
For my purpose, I rerun model (5.15) and further include a non-parametric term capturing an industry effect in addition to the already present market size and time effects, so that the model has the form

$$\log(l_{ijt}) = \sum_{p=0}^{P} \alpha_p \log(\theta_{ijt})^p + h(i, L_j, t) + \epsilon_{ijt}, \tag{5.45}$$

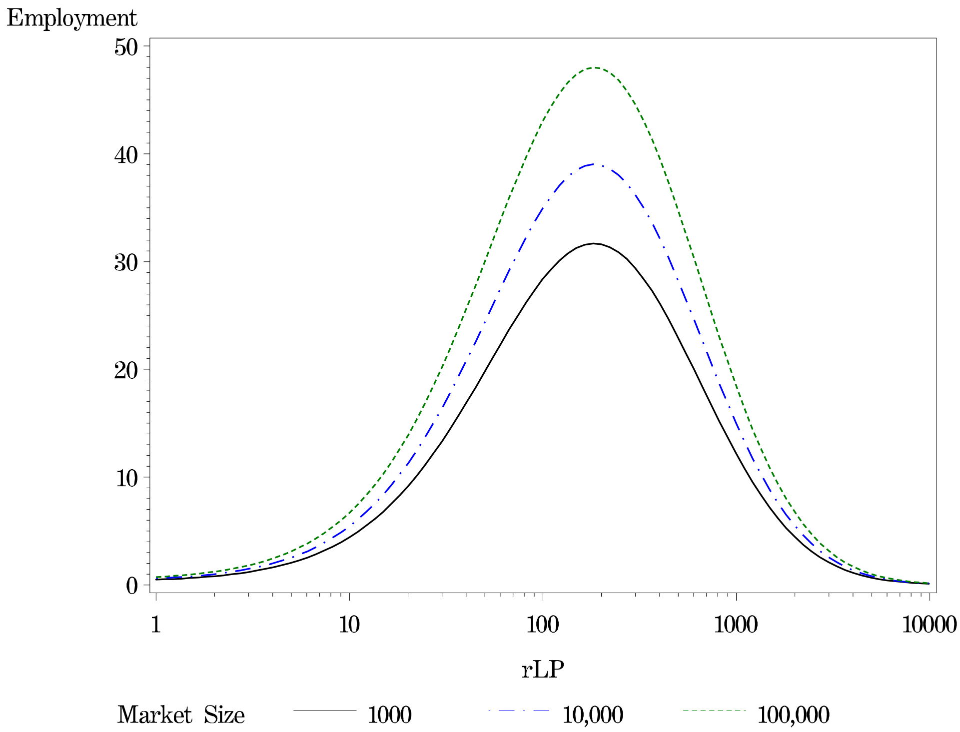
where, i indexes industries and other definitions follow as before. The other difference is that now market size is defined as a CBSA's resident population, because the same definition must be applicable to different industries. Results from Section 5.5.3 assure that, in the case of concrete, using resident population instead of worker population causes very little distortion in the final estimates.

The relationship between productivity and employment is estimated using the same penalized least-squared method discussed in Section 5.3.2, and the predicted results are shown for three market sizes 100000, 1000000, and 10000000 (resident population). The pictures are drawn using the concrete industry fixed-effect, to facilitate comparison with the previously available results from section 5.3.2, but the main goal is to demonstrate that market localization generates a bell-shaped relationship between productivity and employment. The estimation results are shown in Figure 5.21. The sequence of market sizes used here are again chosen to form a geometric series, so that comparing the estimated relations can offer clues about the role of market size in affecting the conduct of industry in localized markets. As can be seen from the figure, the qualitative form of the bell-curve is invariably present at any market size, while the average size of plants grows with market size non-linearly, much in the same way as in the concrete industry.

Industries with more globalized markets, on the other hand, should display higher average productivities and larger average sizes. Section 5.5.5 briefly touches on this issue by comparing the coffee industry to the concrete industry. In my second experiment in this section, I will study the effect of market structure in a broader sense by classifying industries according to their market reach. I still define



(a)

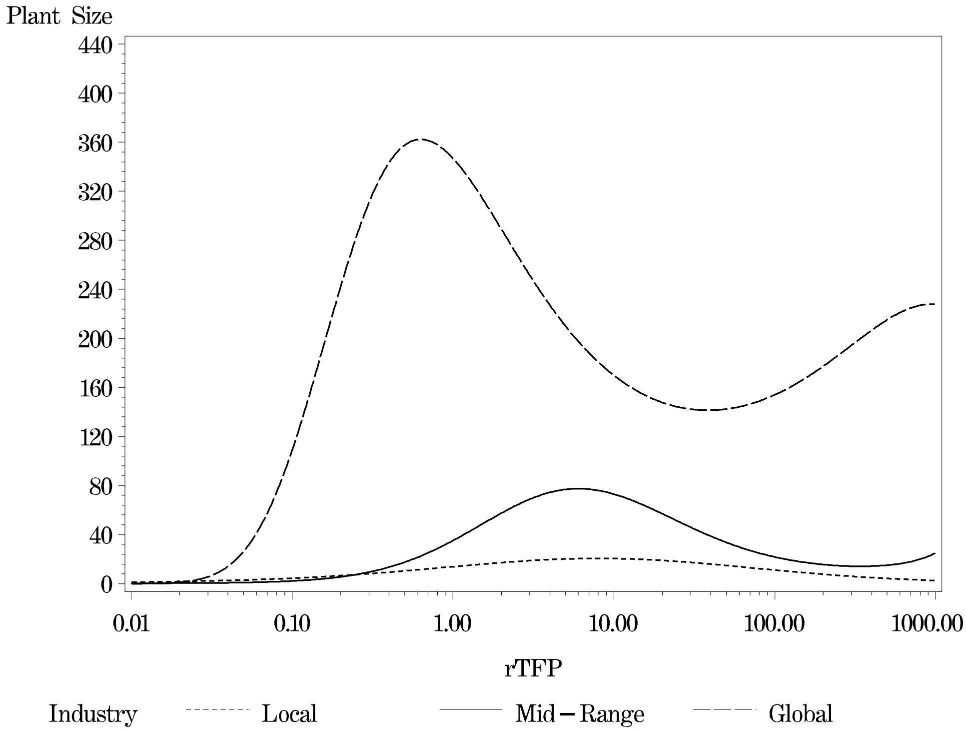


(b)

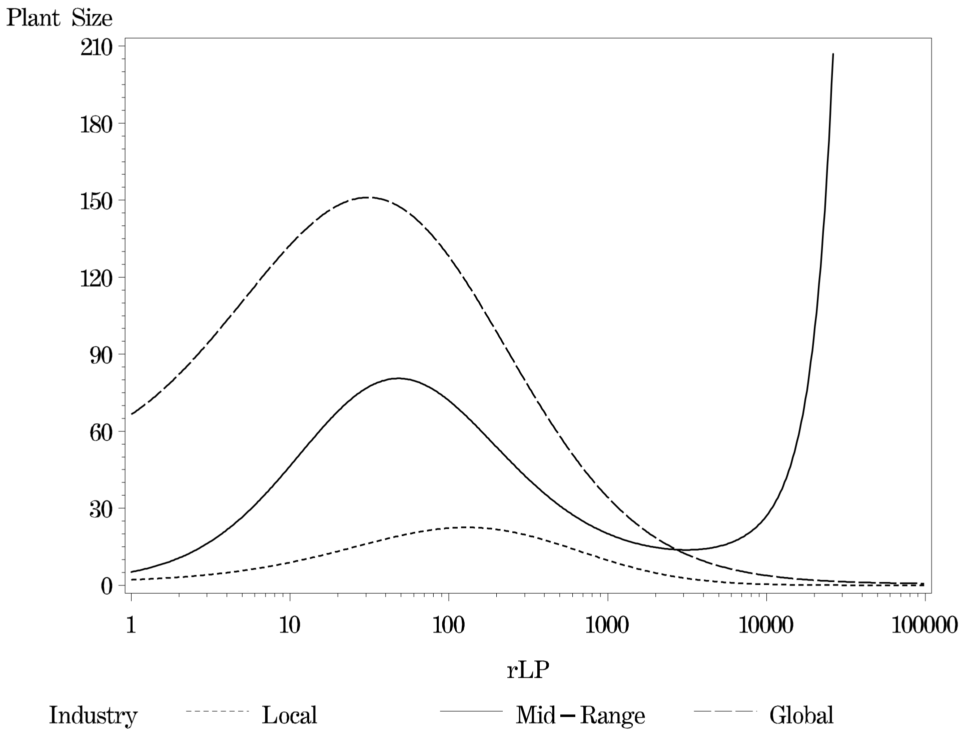
Figure 5.21: Bell-shaped relationship between productivity and employment when pooling across all localized-market plants.

localized-market industries as those industries with an average shipment distance of under 100 miles. Industries with average shipment distance of 500 to 550 miles will be named "mid-range" industries in market reach. Finally, I define industries that sent their products on average more than 950 miles away as globalized-market. The list of industries that fall into each group is listed in Appendix A. The mid-range industries consists of 17 4-digit codes and a total of 92,933 plants. The globalized industries, in turn, consists of 26 4-digit codes and a total of 59,506 plants. The gaps in between shipment distances of the defined classes should help to differentiate the behavior of each group more distinctively.

I use (5.45) to estimate the relationship between productivity and employment within each class by pooling all plants that belong to the corresponding industries. The estimation results are shown in Figure 5.22. In the figures, plants in globalized-market industries have by far the largest average size, while localized-market plants are the smallest on average, with medium-range plants located in the middle. Having said that, all industry classes seem to demonstrate some kind of a bell-shaped relationship between productivity and employment. What differentiates among these classes of industries is mostly the average and the peak size of plants, where more globalization means that more productive plants are much larger on average.



(a)



(b)

Figure 5.22: Estimated relationship between productivity and employment among different classes of market structure.

Chapter 6

Conclusion

Sector-specific factors play an important role in shaping the distribution of productivity among operating plants in that sector. Technology differences and market structure both play roles in decreasing or increasing productivity dispersion within industries. This study is motivated by the fact that productivity dispersion changes non-uniformly with the employment level. I find that the behavior of dispersion is due to long-run behavior of established plants rather than transitory dynamics and selection on productivity. Particularly, industries whose products are primarily traded locally show a significantly negative relationship between productivity dispersion and employment. This is explained by the fact that, in localized markets, the behavior of plants is not only influenced by their productivity, but also by their demand size, causing plants with the same productivity to behave differently in different markets. This effect was shown by constructing a differentiated-product model in which markets are assumed to be localized. The main result of the model is the emergence of a bell-shaped relationship between productivity and input size whenever market structure puts limits on demand size. This relationship served as the engine to produce a behavior of productivity dispersion that was consistent with empirical observations on the Ready-Mix Concrete industry. Particularly, the simulated slope by which productivity dispersion falls with employment almost coincides with the empirical one, showing the dominant role of market-localization in explaining such behavior.

The results of this dissertation can also be used in a broader sense to test the effect of demand structure on the overall conduct of an industry. As both the theoretical model and cross-industry observations show, more global markets raise competitiveness of markets, causing less productive plants to exit while more productive plants can now grow large without constraint. This reduces productivity dispersion at lower sizes and raise the average plant size. Therefore, the slope by which productivity dispersion changes with employment approaches positive values, a fact that is supported by data implications. Still, a more rigorous study of productivity dispersion in industries with more global markets would be useful to characterize in more detail the response of the slope to features of market.

Appendix A

Industry List

- Descriptions of four-digit industries is obtained from the US Department of Labor's SIC manual accessible at `http://www.osha.gov/pls/imis/sic_manual.html`.

- Description of commodities and their data are reported from 1997 Commodity Flow Survey.

Industry Description	SIC	Shipment Distance (miles)
Creamery Butter	2021	74
Natural and Processed Cheese	2022	74
Dry, Condense, and Evaporated Dairy Products	2023	74
Ice Cream and Frozen Desserts	2024	74
Fluid Milk	2026	74
Bread and other Bakery Products	2051	96
Malt beverages	2082	31
Malt	2083	31
Bottled and Canned Soft Drinks and Carbonated Waters	2086	35
Manufactured ICE	2097	35
Logging	2411	85
Asphalt Paving Mixtures and Blocks	2951	70
Asphalt Felts and Coatings	2952	70
Hydraulic Cement	3241	82
Concrete Blocks and Bricks	3271	82
Concrete Products, Except Blocks and Bricks	3272	82
Ready-Mix Concrete	3273	82
Lime	3274	32

Table A.1: List of localized-market industries with shipment distances is less than 100 miles.

Industry		Shipment Distance
Description	SIC	(miles)
Man-made Fiber and Silk	2221	509
Tire Cord and Fabrics	2296	509
Wooden Boxes and Shook	2441	520
Wood Pallets and Skids	2448	520
Wood Preserving	2491	520
Reconstituted Wood Products	2493	520
Cellulosic Fibers	2823	509
Other Organic Fibers	2824	509
Cosmetics and Toilet Products	2844	522
Adhesives and Sealants	2891	522
Unsupported Plastic Profile Shapes	3082	509
Laminated Plastic Profile Shapes	3083	509
Plastic Bottles	3085	509
Plastic Foam Products	3086	509
Custom Compound of Plastic Resins	3087	509
Screw Machine Products	3451	524
Bolts and Nuts	3452	524

Table A.2: List of medium-range industries with shipment distances between 500 and 550 miles.

Industry		Shipment Distance
Description	SIC	(miles)
Women Hosiery	2251	956
Hosiery	2252	956
Knit Outerwear Mills	2253	956
Knit underwear Mills	2254	956
Weft Knit Fabric Mills	2257	956
Lace and Warp Knit Fabric Mills	2258	956
Men's Suits and Coats	2311	956
Men's Shirts	2321	956
Men's Underwear and Nightwear	2322	956
Men's Neckwear	2323	956
Men's Trousers and Slacks	2325	956
Men's Work Clothing	2326	956
Women's Blouses and Shirts	2331	956
Women's Dresses	2335	956
Women's Suits, Skirts and Coats	2337	956
Women's Underwear and Nightwear	2341	956
Brassieres, Girdles, and Allied Garments	2342	956
Hats, Caps, and Milliney	2353	956
Children's Dresses, Blouses, and Shirts	2361	956

Table A.3: List of globalized-market industries with shipment distances of at least 950 miles.

(Cont.)

Industry		Shipment Distance
Description	SIC	(miles)
Dress and Work Gloves	2381	956
Robes and Dressing Gowns	2384	956
Waterproof Outerwear	2385	956
Textile Bags	2393	956
Household Audio and Video	3651	1087
Prerecorded Tapes and Disks	3652	1087
Magnetic and Optical Recording Media	3695	1079

Table A.4: List of globalized-market industries with shipment distances of at least 950 miles.

Appendix B

Technical Appendix

Proof of Proposition 1

For $\nu = 1$ the proof is trivial. Let $0 < \nu < 1$. Testing (5.7) for two extreme values $q_j = 0$ and $q_j \to \infty$ and with the continuity of $\partial \pi_j / \partial q_j$, at least one crossing point is found in the range $q_j > 0$. Moreover, the second derivative of the profit function is

$$\frac{\partial^2 \pi_j}{\partial q_j^2} = -\frac{2\gamma}{L} - \frac{(1-\nu)w}{\nu^2 \phi_j^{1/\nu}} q_j^{1/\nu - 2}, \tag{B.1}$$

which is always negative for any $q_j > 0$. Since two maxima cannot appear next to each other without any local minimum in between them, then, there is only one positive solution to (5.7).

To show that the solution can never be negative, let $q_j < 0$ be the solution to (5.7). We notice that a negative solution can always be written in complex form as $q_j = q e^{\imath \pi}$, where $q > 0$ and q is real. Replacing this in (5.7) results in a left hand side with nonzero imaginary part for any $\nu < 1$. Having a real right hand side, this contradicts the fact that q_j is a solution. $\diamondsuit$

Proof of Proposition 2

Since prices must be non-negative, it follows from (5.3) that output size is bounded above within a certain market. Now if we let $\phi_j \to \infty$ in (5.7) and having

$\bar{p} \leq \alpha$ (and therefore $(\alpha\gamma + \eta N\bar{p})/(\gamma + \eta N) \leq \alpha$), then q_j will converge to $L\alpha/2\gamma$. $\diamondsuit$

Proof of Lemma 1

From (5.7) the optimal price for each plant can be written as

$$p_j = \frac{\gamma}{L} q_j + \frac{w}{\nu}\frac{x}{q}. \tag{B.2}$$

Therefore, using (B.2) and knowing that $q_j = \phi_j x_j^{\nu}$, the revenue productivity can be written as

$$\theta_j = \frac{p_j q_j}{x_j} = \frac{\gamma}{L}\phi_j^{1/\nu} q_j^{2-1/\nu} + \frac{w}{\nu}. \tag{B.3}$$

Taking derivatives with respect to ϕ results in

$$\frac{d\theta_j}{d\phi_j} = \frac{\gamma}{L\nu}\phi_j^{\frac{1}{\nu}-1} q_j^{2-\frac{1}{\nu}} + \frac{\gamma(2\nu-1)}{L\nu}\phi_j^{\frac{1}{\nu}} q_j^{1-\frac{1}{\nu}}\frac{dq_j}{d\phi_j} > 0, \tag{B.4}$$

and the above result follows because of (5.10). $\diamondsuit$

Proof of Proposition 4

First, I show that the endogenous term $\frac{\alpha\gamma+\eta N\bar{p}}{\gamma+\eta N}$ can never converge to zero. If so, then the only possible case is when $N \to \infty$ and $\bar{p} \to 0$. But it means that $p_j \to 0, \forall j$. In turn, (5.3) implies that $q_j \to 0, \forall j$. But this means that all plants will exit the market, driving N to zero. This contradicts the original assumption that $N \to \infty$. Hence, $0 < \frac{\alpha\gamma+\eta N\bar{p}}{\gamma+\eta N} \leq \alpha < \infty$.

To complete the proof, two cases must be treated separately.

Case 1, $\nu = 1$: Then using (5.7) and knowing that $q_j = \phi_j x_j$, I can write

$$x_j = \frac{L}{2\gamma}\left(\frac{\alpha\gamma + \eta N\bar{p}}{\gamma + \eta N}\phi_j - w\right). \tag{B.5}$$

Then since the term $\frac{\alpha\gamma+\eta N\bar{p}}{\gamma+\eta N}$ is always positive, size of plants with productivities above a certain cutoff productivity will go to infinity.

Case 2, $\nu < 1$ **:** Using (5.7) and with boundedness of the right-hand side, it is clear that q_j cannot grow faster than L. Then as $L \to \infty$, (5.7) converges to

$$a_0 + \frac{w}{\nu \phi_j^{1/\nu}} q_j^{1/\nu - 1} = A_0, \tag{B.6}$$

where a_0 is a positive constant if L and q_j grow at the same rate, and zero if q_j grows slower than L. A_0 is the limit value of $\frac{\alpha\gamma+\eta N\bar{p}}{\gamma+\eta N}$ and non-negative. Note that (B.6) requires that $q_j \geq 0$ and $a_0 \leq A_0$. Using $q_j = \phi_j x_j^{\nu}$ results in

$$x_j = \left(\frac{(A_0 - a_0)\nu}{w} \phi_j \right)^{\frac{1}{1-\nu}}. \tag{B.7}$$

The case where $a_0 = A_0$ (q_j and L grow at the same rate) can be immediately rejected here as it implies that $q_j \to 0$, and that contradicts the fact that $a_0 > 0$. Therefore, x_j will be exponentially increasing in ϕ_j.

From Lemma 1, it also follows that he limit relationship between input size and revenue productivity is a monotonic one, and that completes the proof. ♢

Algebraic Steps to (5.14)**:** Let's define

$$A = \frac{\alpha\gamma + \eta N\bar{p}}{\gamma + \eta N}. \tag{B.8}$$

At this point A is an endogenous variable that will facilitate further algebra. The cutoff conditions will be

$$\frac{2\gamma}{L} q_* + \frac{w}{\nu \phi_*^{1/\nu}} q_*^{(1-\nu)/\nu} = A \tag{B.9}$$

$$\left(A - \frac{\gamma}{L} q_* \right) q_* - \frac{w}{\phi_*^{1/\nu}} q_*^{1/\nu} = f \tag{B.10}$$

Eliminating A between (B.9) and (B.10), and substituting $q_* = \phi_* x_*^\nu$, gives

$$\frac{\gamma}{L} q_*^2 + \frac{(1-\nu)w}{\nu} x_* = f. \tag{B.11}$$

Equation (B.11) can be rewritten in the following way

$$q_* \left(\frac{2\nu - 1}{1-\nu} \frac{\gamma}{L} q_* + \frac{2\gamma}{L} q_* + \frac{w}{\nu \phi_*^{1/\nu}} q_*^{(1-\nu)/\nu} \right) = \frac{f}{1-\nu}. \tag{B.12}$$

Looking at (B.12), it is easy to recognize and replace the term from (B.9). Thus, with some simple algebra, (B.12) yields

$$A = \frac{f}{(1-\nu) q_*} - \frac{2\nu - 1}{1-\nu} \frac{\gamma}{L} q_*. \tag{B.13}$$

By replacing A in the profit function, a plant's profit at an optimum can be expressed as below which is a function of q_* only:

$$\pi_j = \left(\frac{f}{(1-\nu) q_*} - \frac{\gamma}{L} \frac{2\nu - 1}{1-\nu} q_* - \frac{\gamma}{L} q_j \right) q_j - w \frac{q_j^{1/\nu}}{\phi_j^{1/\nu}} - f. \tag{B.14}$$

To find $\partial \phi_* / \partial L$, I need to find $\partial q_* / \partial L$ first. To find the derivatives, note that a change in market size affects q_* both directly and indirectly, through ϕ_*. With this in mind, taking partial derivatives of (B.11) with respect to L results in

$$\frac{\partial q_*}{\partial L} = \frac{\frac{\gamma}{L^2} q_*^2 + \frac{(1-\nu)w}{\nu^2 \phi_*} x_* \frac{\partial \phi_*}{\partial L}}{\frac{2\nu - 1}{\nu} \frac{\gamma}{L} q_* + \frac{f}{\nu q_*}}. \tag{B.15}$$

Next, insert (B.14) into (5.8) and take partial derivatives with respect to L. Replacing $\partial q_* / \partial L$ from (B.15) leads to (5.14). ♢

Proof of Proposition 5: Both the nominator and the denominator in the (5.14) will be unambiguously positive in the light of (5.10) and if $\nu \geq 0.5$. Therefore, it immediately follows that $\partial\phi_*/\partial L > 0$. Having a fixed maximum productivity, the higher the cutoff productivity goes, the smaller the productivity dispersion will become. $\diamondsuit$

Bibliography

Abel, A. B. & Eberly, J. C. (1996), 'Optimal investment with costly reversibility', *Review of Economic Studies* **63**, 581–593.

Abowd, J. M., Kramarz, F. & Margolis, D. N. (1999), 'High wage workers and high wage firms', *Econometrica* **67**(2), 251–333.

Acs, Z. J. & Audretsch, D. B. (1991), *Innovation and Technological Change*, The University of Michigan Press, Ann Arbor, chapter 4,6.

Asplund, M. & Sandin, R. (1999), 'The number of firms and production capacity in relatrion to market size', *The Journal of Industrial Economics* **47**(1), 69–85.

Baily, M., E.J.Bartelsman & J.Haltiwanger (1994), Downsizing and productivity growth: Myth or reality?, Technical Report 4741, NBER Working Paper series.

Baily, M. N., Hulten, C., Campbell, D., Bresnahan, T. & Caves, R. E. (1992), 'Productivity dynamics in manufacturing plants', *Brookings Papers on Economic Activity: Microeconomics* **1992**, 187–267.

Bartelsman, E. J. & Dhrymes, P. J. (1998), 'Productivity dynamics: U.s. manufacturing plants 1972-1986', *Journal of Productivity Analysis* **9**(1), 5–34.

Bartelsman, E. J. & Doms, M. (2000), 'Understanding productivity: Lessons from longitudinal microdata', *Journal of Economic Literature* **38**(3), 569–594.

Basu, S. & Fernald, J. G. (1997), 'Returns to scale in u.s. production: Estimates and implications', *Journal of Political Economy* **105**(2), 249–283.

Bernard, A. B., Redding, S. J. & Schott, P. K. (2006), Multi-product firms and product switching, Technical Report 12293, NBER Working Paper series.

Berry, S. & Waldfogel, J. (2003), Productivity quality and market size, Technical Report 9675, NBER Working Paper series.

Bertola, G. & Garibaldi, P. (2001), 'Wages and the size of firms in dynamic matching models', *Review of Economic Dynamics* **4**(2), 335–368.

Bontemps, C., Robin, J.-M. & Van Den Berg, G. J. (2000), 'Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation', *International Economic Review* **41**(2), 305–358.

Brown, C. & Medoff, J. (1989), 'The employer size-wage effect', *Journal of Political Economy* **97**(5), 1027–1059.

Campbell, J. R. & Hopenhayn, H. A. (2002), Market size matters, Technical Report 9113, NBER Working Paper series.

Chew, W., Clark, K. B. & Bresnahan, T. (1990), *Measure for Manufacturing Excellence*, Harvard Business School Press, Boston, Massachusetts, chapter 5. edited by Robert S. Kaplan.

Chiang, H. (2005), Essays on Employer-Employment Relationships and Firm Performance, Ph.d. dissertation, University of Maryland.

Churchill, N. C. & Lewis, V. L. (1983), 'The five stages of small business growth', *Harvard Business Review* pp. 30–50.

Davis, S. J. & Haltiwanger, J. (1991), 'Wage dispersion between and within u.s. manufacturing plants, 1963-1986', *Brookings Papers on Economic Activity, Microeconomics* **1991**, 115–200.

Davis, S. J., Haltiwanger, J. & Schuh, S. (1996), *Job Creation and Destruction*, Massachusetts Institute of Technology.

Dixit, A. K. & Stiglitz, J. E. (1977), 'Monopolistic competition and optimum product diversity', *American Economic Review* **67**(3), 297–308.

Dunne, T., Foster, L., Haltiwanger, J. & Troske, K. (2000), Wage and productivity dispersion in u.s. manufacturing: The role of computer investment, Technical Report 7465, NBER Working Paper series.

Ellison, G. & Glaeser, E. L. (1997), 'Geographic concentration in u.s. manufacturing industries:a dartboard approach', *J. of Political Economy* **105**(5), 889–927.

Ericson, R. & Pakes, A. (1995), 'Markov-perfect industry dynamics: A framework for epirical work', *The Review of Economic Studies* **62**(1), 53–82.

Evans, D. S. (1987), 'The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries', *The Journal of Industrial Economics* **35**(4), 567–581.

Feenstra, R. C. (1996), U.s. imports, 1972-1994: Data and concordances, Technical Report 5515, NBER Working Paper Series.

Feenstra, R. C. (1997), U.s. exports, 1972-1994: With state exports and other u.s. data, Technical Report 5990, NBER Working Paper Series.

Fisher, F. M. (1993), *Aggregation*, 1st edn, MIT Press, Cambridge.

Foster, L., Haltiwanger, J. & Krizan, C. (1998), Aggregate productivity growth: Lessons from microeconomic evidence, Technical Report 6803, NBER Working Paper Series.

Foster, L., Haltiwanger, J. & Syverson, C. (2008), 'Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?', *American Economic Review*.

Gollop, F. M. & Monahan, J. L. (1991), 'A generalized index of diversification: Trends in u.s. manufacturing', *The Review of Economics and Statistics* **73**(2), 318–330.

Greiner, L. E. (1998), 'Evolution and revolution as organizations grow', *Harvard Business Review* pp. 55–67.

Hall, B. H. (1987), 'The relationship between firm size and firm growth in the us manufacturing sector', *The Journal of Industrial Economics* **35**(4), 583–606.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, Springer Verlag.

Haltiwanger, J. (1997), 'Measuring and analyzing aggregate fluctuations: The importance of building from microeconomics evidence', *Fedreal Reserve Bank of St. Louis Review, May/June,* pp. 55–78.

Haltiwanger, J. C., Lane, J. I. & Spletzer, J. (2000), Wages, productivity, and the dynamic interaction of businesses and workers, Technical Report 7994, NBER Working Paper series.

Hamermesh, D. S. (1995), 'Labour demand and the source of adjustment costs', *The Economic Journal* **105**(430), 620–634.

Hopenhayn, H. (1992), 'Entry, exit, and firm dynamics in long run equilibrium', *Econometrica* **60**(5), 1127–1149.

Jarmin, R. & Miranda, J. (2002), The longitudinal business database, Technical Report CES-WP-2002-17, CES Discussion papers.

J.Bartelsman, E. & Gray, W. (1996), The nber manufacturing productivity database, Technical Report 205, NBER Working Paper series.

Jovanovic, B. (1982), 'Selection and the evolution of industry', *Econometrica* **50**(3), 649–670.

Katayama, H., Lu, S. & Tybout, J. R. (2003), Why plant-level productivity studies are often misleading, and an alternative approach to inference, Technical Report 9617, NBER Working Paper series.

Kemeny, J. G. & Snell, J. L. (1983), *Finite Markov Chains*, 3rd edn, Springer Verlag.

McGuckin, R. H. & George A. Pascoe, J. (1988), The longitudinal research database (LRD): Status and research possibilities, Technical Report CES-88-2, CES Discussion papers.

Melitz, M. & Ottaviano, G. I. (2005), Market size, trade, and productivity, Mimeo, Harvard University.

Melitz, M. J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica* **71**(6), 1695–1725.

Moussa, M. A. A. & Cheema, M. Y. (1992), 'Non-parametric regression in curve fitting', *The Statistician* **41**(2), 209–225.

Olley, G. S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), 1263–1297.

Sutton, J. (1991), *Sunk Costs and Market Structure*, Massachusetts Institute of Technology.

Syverson, C. (2003), Product substitutability and productivity dispersion, Technical Report 10049, NBER Working Paper series.

Syverson, C. (2004), 'Market structure and productivity: A concrete example', *Journal of Political Economy* **112**(6), 1181–1222.

Ziou, D. & Tabbone, S. (1998), 'Edge detection techniques - an overview', *International Journal of Pattern Recognition and Image Analysis* **8**(4), 537–559.