COMPETITION, INNOVATION, AND REGULATION:
ACCOUNTING FOR PRODUCTIVITY DIFFERENCES

by

Pedro Bento

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Economics
University of Toronto

© Copyright 2013 by Pedro Bento
Abstract

Competition, Innovation, and Regulation:
Accounting for Productivity Differences

Pedro Bento
Doctor of Philosophy
Graduate Department of Economics
University of Toronto
2013

The relationships between competition, innovation, and regulation have long been studied in an attempt to understand and evaluate the effect of regulation on the wealth and growth of nations. Recent empirical work has emerged taking advantage of the still ongoing proliferation of ever more disaggregated data to shed more light on these relationships and at the same time uncover new puzzles in need of explanations. This thesis is an attempt to address the discrepancies between some of these newly discovered phenomena and current theory.

In Chapter 1 I introduce an insight of Friedrich Hayek — that competition allows a thousand flowers to bloom, and discovers the best among them — into a conventional model of Schumpeterian innovation. I show how the model can account for two seemingly contradictory empirical phenomena, a positive relationship between competition and industry-level productivity growth, and an inverted-U relationship between competition and firm-level innovation. In Chapter 2 I extend the model to investigate the effects of patent protection on competition and innovation, and to understand the interaction between patent policy and product-market regulation. I calibrate the model to show that patent protection in the U.S. is depressing competition, innovation, growth, and welfare. Using patent and citation data, I further provide empirical evidence supporting the implications of the model.

In Chapter 3 I investigate the impact of regulatory entry barriers to new firms on aggregate output and total factor productivity. Following recent work by Thomas J. Holmes and John J. Stevens, I extend a standard model of monopolistic competition to account for the existence of both niche markets and mass markets within industries. Calibrating the model using U.S. manufacturing data, I show this extension goes a long way towards explaining the large gap between empirical estimates of the impact of barriers to entry and the quantitative predictions of current models.
Acknowledgements

I am grateful to Diego Restuccia, Xiaodong Zhu, and Gueorgui Kambourov for their guidance and supervision over the past few years. I would also like to thank Victor Aguirregabiria, Michelle Alexopoulos, David Andolfatto, Margarita Duarte, Gilles Duranton, Alexander Karaivanov, Nicholas Li, Angelo Melino, Matt Mitchell, Peter Morrow, Carolyn Pitchik, and participants in various seminars for valuable comments and suggestions.
## Contents

1 Competition as a Discovery Procedure .................................................. 1
1.1 Introduction .......................................................... 2
1.2 The Model .......................................................... 4
  1.2.1 Environment .................................................. 4
  1.2.2 Market Structure ............................................. 6
  1.2.3 Competitive Equilibrium .................................... 7
  1.2.4 Results ....................................................... 9
1.3 Empirical Support .................................................... 13
1.4 Extension: Antitrust Policy .......................................... 15
  1.4.1 Collusion and Anti-Competitive Mergers ......................... 15
1.5 Conclusion ........................................................ 17

2 Patent Protection as a Tax on Competition and Innovation .................. 19
2.1 Introduction ........................................................ 20
2.2 The Model ........................................................ 21
  2.2.1 Environment .................................................. 22
  2.2.2 Innovation ................................................... 22
  2.2.3 Patents ......................................................... 23
  2.2.4 Competitive Equilibrium .................................... 23
  2.2.5 Results ....................................................... 27
2.3 Quantitative Analysis ................................................. 29
2.4 Patents, Citations, and Patent Protection .................................... 31
  2.4.1 Data ........................................................ 32
  2.4.2 Econometric Models ......................................... 33
  2.4.3 Results ....................................................... 33
2.5 Conclusion ........................................................ 35

3 Niche Firms, Mass Markets, and Income Across Countries .................. 36
3.1 Introduction ........................................................ 37
3.2 The Model ........................................................ 39
  3.2.1 Environment .................................................. 39
  3.2.2 Competitive Equilibrium .................................... 40
  3.2.3 Heterogeneous Firms ........................................ 42
3.3 Calibration ......................................................... 43
List of Tables

2.1 Calibrated Parameter Values ...................................................... 30
2.2 Quantitative Results ................................................................. 31
2.3 Descriptive Statistics ............................................................... 32
2.4 Results ...................................................................................... 34
2.5 Results ...................................................................................... 34

3.1 Descriptive Statistics: Mean ....................................................... 45
3.2 Cross-Country Experiment: TFP .................................................... 47
3.3 Cross-Country Experiment: Output per Worker ............................... 47

C.1 Estimation Results ................................................................. 64
List of Figures

1.1 Optimal Research per Firm and Free Entry ........................................... 10
1.2 Number of Firms and Research per Firm with variation in $z$ ........................ 11
1.3 Competition, Innovation, and Growth, with variation in $z$ .......................... 12

3.1 Model vs Data ..................................................................................... 49
3.2 Standard CES Model vs Data ............................................................... 49

A.1 $e^*$: Number of Firms at which Research per Firm Peaks ......................... 55
Chapter 1

Competition as a Discovery Procedure:
Schumpeter Meets Hayek in a Model of Innovation

Abstract:
I incorporate an insight of Friedrich Hayek — that competition allows a thousand flowers to bloom, and discovers the best among them — into a model of Schumpeterian innovation. Firms face uncertainty about the optimal direction of innovation, so more innovations implies a higher expected value of the ‘best’ innovation. The model accounts for two seemingly contradictory relationships reported in recent empirical studies — a positive relationship between competition and industry-level productivity growth, and an inverted-U relationship between competition and firm-level innovation. Notwithstanding the positive relationship between competition and growth, I find antitrust policy reduces industry-level growth.
1.1 Introduction

Schumpeter (1942) argues that the expectation of monopoly power is necessary to induce innovation. This implies that a higher level of competitive rivalry should translate into lower levels of innovation, and thus lower productivity growth. Early models of innovation include this Schumpeterian mechanism and share the conclusion that competition is harmful to growth.¹ Recent empirical studies, however, come to very different conclusions. Productivity growth at the industry level has been shown to be positively correlated with competition, while the average level of innovation per firm in an industry exhibits an inverted-U relationship with competition — that is, a positive relationship when competition is relatively low, and a negative relationship when competition is high.² These two stylized facts suggest the need for a model that both explains how an increase in competition can be accompanied by higher growth, and allows for industry-level outcomes to differ from those of the average firm. The present chapter accomplishes this by introducing uncertainty into a model of quality-improving innovations with endogenous firm entry. I assume that entrepreneurs are uncertain about the relative value of each possible direction of innovation, until an innovation has actually been introduced to the market. In such an economy, a greater number of firms implies a greater number of innovations tried, which in turn implies a higher expected value of the ‘best’ innovation. If the best innovation can capture a market (say, through Bertrand competition), it becomes possible that measured productivity growth for an industry can be increasing in the number of firms, even if the average level of research per firm is declining (due to the usual Schumpeterian mechanism).

Imposing uncertainty about the optimal direction of innovation (call it Hayekian uncertainty) captures an insight of F.A.Hayek (2002). In Hayek’s words, competition is (p.9):

> a procedure for discovering facts which, if the procedure did not exist, would remain unknown or at least would not be used.

When there is nothing to discover, competition results in static efficiency — identical firms compete for inputs and customers, driving price to marginal cost and unwittingly attaining allocative efficiency. When firms are uncertain about which new products or production processes will turn out to be the most valuable, however, competition acts to sort the best from the worst ideas tried. Daft (2004) reports 80 percent of new products fail in their first year, suggesting this Hayekian uncertainty plays an important role in the competitive process.

Why are entrepreneurs so uncertain about the value of innovations? When setting out to improve a product or production process, there is no one-dimensional measure of quality to progress along. One firm may improve the carrying capacity and handling of its car, only to lose market share to a competitor’s more stylish alternative. A fast-food chain may increase the speed of its drive-thru service, only to lose market share to a competitor who introduces fresher ingredients.

Key here is the distinction between uncertainty about how to accomplish a given goal (i.e., improve the handling of a car), and uncertainty about the relative values of competing goals (i.e., improving the handling of a car versus making a car more stylish). While experimenting with different ways of achieving

---

¹For example, see Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992). More recent models of R&D-based endogenous growth continue to share the implication that lower entry barriers (and thus more competition) should lower the growth rate. For example, see Dinopoulos and Thompson (1998), Connolly and Peretto (2007), and Etro (2007).

²For example, see Nickell (1996), Blundell et al. (1999), Dutz and Hayri (2000), Aghion et al. (2005) and Aghion et al. (2008).
Chapter 1. Competition as a Discovery Procedure

A goal may lead to a better outcome, such experimentation can be done within a firm when outcomes are easily measurable. On the other hand, if the relative values of competing innovations can only be determined after they are introduced to the market, firms face a strong disincentive to experiment with competing products. As a result, experimentation will arise primarily through the introduction of new products or production processes by multiple firms. In such a setting, competition works to ‘discover’ the best innovation.

Incorporating Hayekian uncertainty into a model of quality innovations creates a positive link between industry-level productivity growth and the endogenously-determined number of firms in an industry. By allowing the quality of its rivals to affect a firm’s pricing decision (through Bertrand competition), I create a link between a firm’s markup over marginal cost (a standard measure of competition) and the number of firms in its industry. By further allowing each firm to decide the magnitude of its quality improvement through research, I can examine the interplay between competition, industry-level productivity growth, and firm-level innovation in general equilibrium. Throughout this chapter the Schumpeterian mechanism, whereby lower rents discourage research, is always present. But when Schumpeter meets Hayek in this model of innovation, more competition is accompanied by higher industry-level productivity growth, even if firm-level research is reduced.

An advantage of modeling competition and innovation as endogenous outcomes is it allows for policy evaluation that is less susceptible to the Lucas Critique. When competition is endogenous, the evaluation of antitrust policy requires one to model a policy intervention as a change in the parameters of the model or in the constraints faced by firms, rather than as an assumption that more enforcement is equivalent to more competition. Whereas studies like Geroski (1990) and Dutz and Hayri (2000) interpret the positive empirical relationship between competition and growth as implying greater antitrust enforcement is good for growth, I find that restrictions on collusion between firms have a detrimental effect on industry-level productivity growth.

The model developed in this chapter is similar in many respects to those that have been developed in the innovation-driven endogenous growth literature. Closest to the present model is Grossman and Helpman (1991), who consider a continuum of industries, each made up of firms attempting to innovate by building on the technology of an incumbent and thus capture the market. Grossman and Helpman incorporate uncertainty in the timing of innovations, keeping the magnitude of innovations constant, while I keep the timing constant but make the magnitude of each innovation uncertain. While in Grossman and Helpman R&D conducted by a firm affects its probability of capturing the market, R&D in the present model has the additional function of increasing the expected magnitude of an innovation. In addition, when the magnitude of each innovation is uncertain, the number of innovating firms itself affects the expected value of the best innovation. Aghion et al. (2005) model an economy with a continuum of industries, each structured as a duopoly, where industries differ in the potential for collusion between competitors. The authors show their model can generate an inverted-U between competition and firm-level innovation, when variation in competition is driven by variation in the potential for collusion. This chapter instead focuses on variation in the economic and regulatory costs faced by innovators. In addition, the present model incorporates endogenous entry, while Etro (2007) suggests the inverted-U in Aghion et al. disappears once entry is allowed for. By allowing firms to compete for the same market as

---

3 The intuition here is straightforward. Consider a firm that is introducing one product and has some expectation of a resulting stream of profits. If this same firm was to introduce a second product in direct competition, it would reduce the expected profits from the first product. Such a firm would be better off introducing a differentiated second product, not in direct competition with the first.
in Aghion et al., but extending the model to allow for free entry, I can account for an inverted-U at the firm level while still generating a positive relationship between competition and productivity growth at the industry level. By incorporating free entry into the market, I also retain a connection between the conventional definition of ‘competition’ — more firms in the market — and the measure of competition used in recent empirical studies. To my knowledge, no paper that allows for the free entry of firms has found a mechanism through which lower entry barriers could result in higher growth either at the firm level or at the level of an industry. Peretto (1996) can generate an increase in growth and in the number of firms when fixed operating costs are reduced, but requires an assumption that a firm’s R&D directly improves the productivity of rival firms. The present chapter assumes only intertemporal spillovers, as new innovations build on the old. Boldrini and Levine (2002, 2005, 2008a) study the conditions under which innovation can occur under perfect competition, while the present chapter considers variation in the level of competition.

Several papers have modeled selection among heterogeneous firms (for example, see Melitz (2003) and Hopenhayn (1992)), where low-productivity firms cannot cover a fixed cost of operation. In this chapter as in Bernard et al. (2003), the best firm can capture its market, thus causing selection to arise without fixed operating costs.

In the next section I present the model and its implications for the relationships between competition, innovation, and productivity growth. Section 1.3 follows with a survey of the empirical evidence that supports the central mechanism in the model and its implications. In Section 1.4 I extend the model to evaluate the effects of antitrust policies on growth. Section 1.5 concludes.

1.2 The Model

1.2.1 Environment

Consider a closed economy where time is discrete and indexed by \( t \). There is a representative final-good (\( y \)) firm which uses as inputs a continuum (of measure 1) of products from a representative intermediate-good (\( x \)) industry. The final-good firm produces according to the following standard CES production function:

\[
y = \left( \int_0^1 X^\alpha(j) dj \right)^{\frac{1}{\alpha}},
\]

where \( j \) indexes intermediate products and \( \frac{1}{\alpha} \) is the constant elasticity of substitution (CES) between products. \( X(j) \) is defined in the following way;

\[
X(j) = \sum_{i=1}^{e(j)} A_i(j) x_i(j),
\]

where \( x_i(j) \) denotes the amount of product-\( j \) demanded from firm-\( i \), \( A_i(j) \) is a measure of the quality of \( x_i(j) \), and \( e(j) \) is the (endogenously-determined) number of firms that can potentially produce product-\( j \).

Any firm-\( i \) can produce according to;

\[
x_i(j) = L_i(j),
\]

---

4Peretto (1999) shows this result disappears when spillovers are removed from the model.
5Throughout the paper, I omit the time subscript unless clarity requires it.
where \( x_i(j) \) is the amount of product-\( j \) produced by firm-\( i \) of quality \( A_i(j) \), and \( L_i(j) \) is its labor input. Each period, any of a large number of firms can choose to invest in innovation, thereby acquiring a quality of \( A(j) > 0 \) for output produced in the subsequent period. I assume any firm-\( i \) that does not innovate in \( t - 1 \) receives a quality of \( A_i(j) = 0 \). The total investment required to innovate is made up of a fixed cost of introducing an innovation, \( z\psi(j) \), and the cost of research, \( m\psi(j)n_i(j) \), where \( z \) and \( m \) are exogenous, and \( n_i \) is the level of research chosen by firm-\( i \). The research cost is the cost of actually improving the quality of a product, while the fixed cost can be interpreted as the cost of introducing this product to the market. Introducing a new product to the market might entail holding an inventory for a time before demand is realized, or fulfilling regulatory requirements like efficacy trials for drugs. As such, a firm’s quality \( A_i(j) \) is not realized until after both costs are sunk. \( \psi(j) \) denotes total sales of product-\( j \). Assuming the cost of innovation scales up with output can be rationalized as reflecting the phenomenon that each incremental quality improvement is more costly than the last (this phenomenon is modeled in a similar fashion in Aghion and Howitt (2005)), but \( \psi(j) \) will also serve to make research and innovation decisions independent of past innovations and total population. Finally, firms finance investments by issuing equity to households.

At the beginning of each period, the quality \( A_i(j) \) associated with each firm-\( i \) that invested in innovation in the previous period becomes known to all agents. For each firm-\( i \) the introduction of an innovation in period \( t \) results in a quality of:

\[
A_{it}(j) = A_{t-1[i]}(j)h_{it}(j),
\]

where \( A_{t-1[i]}(j) \) is the highest ([1]) quality for product-\( j \) in period \( t - 1 \), and \( h_{it}(j) \) is the realized value of a random draw. I assume \( h_{it}(j) \) is distributed according to some twice continuously differentiable distribution \( F_i(h) \), bounded by 0 and \( n_i^0(j) \), \( \theta \in (0, 1) \), where \( n_i(j) \) is the level of research invested in by firm-\( i \) and \( f_i(n_i^0) \) is assumed to be finite. The randomness of the draw represents each firm’s uncertainty about the value of the direction in which it has chosen to innovate relative to other directions, while the level of research \( n \) determines the magnitude of the improvement in that direction. Note the lower bound of zero implies an innovator may misread its potential customers so badly that its new product is less valuable than the previous period’s best. I call this the \textit{New Coke Phenomenon}.\footnote{Changing the support of \( F_i(h) \) to \([1, 1+n_i^0(j)]\) complicates the analysis without changing any of the qualitative results. Allowing innovators with very low draws to use \( A_{t-1[i]} \) does the same.}

There is a representative household which supplies \( L \) units of labor to each intermediate product market.\footnote{Making labor mobile across product markets complicates the analysis without changing any of the qualitative results. I discuss this further in Section 1.2.4.} Households only value consumption, and have a constant discount rate \( \beta \in (0, 1) \). Income can be consumed or saved, and the only vehicle for savings is the purchase of equity in innovating intermediate firms, earning a gross rate of return of \( R \). With a continuum of intermediate industries and no aggregate uncertainty, each household faces a deterministic interest rate and economy-wide growth rate. The household’s problem is therefore to choose consumption \( \{c_t\}_{t=0}^{\infty} \) and savings \( \{s_t\}_{t=0}^{\infty} \) to solve:

\[
\max_{c_{t},s_{t}} \sum_{t=0}^{\infty} \beta^t \ln(c_t), \quad \text{s.t.} \quad c_t + s_t \leq L \int_0^1 w_t(j) \, dj + s_{t-1} R_t, \quad (1.1)
\]
where \( L \int_{0}^{1} w_{i}(j) dj \) is the total wage income from labor supplied to intermediate firms, and savings are assumed to be spread equally across all firms.

### 1.2.2 Market Structure

The final-good sector is perfectly competitive, and so the representative final-good producer takes all prices as given and chooses \( x_{i}(j) \) from each firm-\( i \) in each market-\( j \) to solve:

\[
\max_{\{x_{i}(j)\}_{i,j}} \left( \int_{0}^{1} \left( \sum_{i=1}^{e(j)} A_{i}(j)x_{i}(j) \right)^{\alpha} dj \right) \frac{1}{\int_{0}^{1} \sum_{i=1}^{e(j)} P_{i}(j)x_{i}(j) dj}, \tag{1.2}
\]

where \( P_{i}(j) \) is the price of \( x_{i}(j) \) in terms of the final good, \( A_{i}(j) \) is the quality of \( x_{i}(j) \), and \( e(j) \) is the number of firms in the market for product-\( j \). The combination of perfect substitutability between firms in the same product market and a constant-returns-to-scale intermediate production function ensures that only one firm will produce each product in equilibrium. As a result, the final-good firm demands \( x_{i}(j) \) such that:

\[
P_{i}(j) = A_{i}^{\alpha}(j)x_{i}^{1-\alpha}(j)y^{1-\alpha}, \quad \text{if} \quad A_{i}(j) > \frac{A_{k}(j)}{P_{k}(j)}, \quad \forall k \neq i \in \{1, ..., e(j)\}, \tag{1.3}
\]

\[
x_{i}(j) = 0 \quad \text{otherwise}.
\]

Each firm-\( i \) in each market-\( j \) faces a given wage \( w(j) \) and chooses \( P_{i}(j) \) to solve:

\[
\max_{P_{i}(j)} P_{i}(j)x_{i}(j) - w(j)x_{i}(j), \tag{1.4}
\]

s.t. \( P_{i}(j) = A_{i}^{\alpha}(j)x_{i}^{1-\alpha}(j)y^{1-\alpha}, \quad A_{i}(j) > \frac{A_{k}(j)}{P_{k}(j)}, \quad \forall k \neq i \in \{1, ..., e(j)\}. \]

As a result only the best firm in market-\( j \) produces, charging the following price for \( x_{[1]}(j) \):

\[
P_{[1]}(j) = \begin{cases} w(j)/\alpha, & \text{if } \frac{A_{2}(j)}{A_{[1]}(j)} < \alpha \\ w(j)\frac{A_{[1]}(j)}{A_{2}(j)}, & \text{if } \frac{A_{2}(j)}{A_{[1]}(j)} > \alpha \end{cases} \tag{1.5}
\]

where \( A_{[\ell]}(j) \) is the \( \ell \)th-highest quality of product-\( j \).

Labor-market clearing \( (L = x_{[1]}(j)) \) combines with the demand for \( x_{[1]}(j) \) (1.3) and the price of \( x_{[1]}(j) \) (1.5) to determine the wage rate in market-\( j \):

\[
w(j) = \begin{cases} \alpha A_{[1]}^{\alpha}(j)L^{\alpha-1}y^{1-\alpha}, & \text{if } \frac{A_{2}(j)}{A_{[1]}(j)} < \alpha \\ \frac{A_{2}(j)}{A_{[1]}(j)} L^{\alpha-1}y^{1-\alpha}, & \text{if } \frac{A_{2}(j)}{A_{[1]}(j)} > \alpha \end{cases}
\]

- Throughout this chapter, I use ‘market-\( j \)’ and ‘the market for product-\( j \)’ interchangeably.
Any firm-$i$ innovating in $t - 1$ will therefore face the following expected discounted profits:\(^{11}\)

$$E_{t-1} \left( \frac{\pi_{it}}{R_t} \right) =$$

$$\frac{\psi_{t-1}(1 + g_t)^{1-\alpha}}{R_t} \cdot \Prob \left[ h_{it} = h_{i[1]} > \frac{h_{i[2]}}{\alpha} \right] \cdot E_{t-1} \left( \frac{(1 - \alpha)A^\alpha_{i[t]} L_{it}^{\alpha} y_{it}^{1-\alpha}}{R_t} \bigg| A_{it} = A_{i[1]} > \frac{A_{i[2]}}{\alpha} \right)$$

$$+ \frac{\psi_{t-1}(1 + g_t)^{1-\alpha}}{R_t} \cdot \Prob \left[ h_{it} = h_{i[1]} < \frac{h_{i[2]}}{\alpha} \right] \cdot E_{t-1} \left( \frac{\left( A_{i[t]} - A_{i[2]} \right) L_{it}^{\alpha} y_{it}^{1-\alpha}}{A_{i[t]}^{1-\alpha}} \bigg| A_{it} = A_{i[1]} < \frac{A_{i[2]}}{\alpha} \right)$$

$$- \psi_{t-1}(z + mn_{i,t-1})$$

where $\psi$ is the cost of introducing an innovation, $mn_{i,t}$ is the cost of firm-$i$’s research $n_i$, and $\psi_{t-1}(j) \equiv P_{t-1[1]}(j) L = A_{t-1[1]}(j) L y_{it-1}^{1-\alpha}$. Substituting for $A_{i[t]} = A_{i-1[1]} A_{i[t]}$, expected discounted profits are;

$$E_{t-1} \left( \frac{\pi_{it}}{R_t} \right) =$$

$$\frac{\psi_{t-1}(1 + g_t)^{1-\alpha}}{R_t} \cdot \Prob \left[ h_{it} = h_{i[1]} > \frac{h_{i[2]}}{\alpha} \right] \cdot E_{t-1} \left( \frac{(1 - \alpha)h_{i[1]}^{\alpha} y_{it}^{1-\alpha}}{R_t} \bigg| h_{it} = h_{i[1]} > \frac{h_{i[2]}}{\alpha} \right)$$

$$+ \frac{\psi_{t-1}(1 + g_t)^{1-\alpha}}{R_t} \cdot \Prob \left[ h_{it} = h_{i[1]} < \frac{h_{i[2]}}{\alpha} \right] \cdot E_{t-1} \left( \frac{h_{i[1]} - h_{i[2]}}{h_{i[1]}^{1-\alpha}} \bigg| h_{it} = h_{i[1]} < \frac{h_{i[2]}}{\alpha} \right)$$

$$- \psi_{t-1}(z + mn_{i,t-1})$$

where $h_{i[t]}$ denotes the $t$th-highest realized draw in period $t$, and $g_t$ is the growth rate of final-good output \( \left( \frac{y_{t-1}}{y_{t-1} - 1} \right) \).

Each firm-$i$ draws from the distribution function $F_t(h)$, which is bounded from above by $n_i$. I consider only symmetric equilibria, so for ease of notation I use $n_k$ to denote the research of every other firm except $i$. It follows that the joint density function of $h_{i[1]}$, $h_{i[2]}$, and $h_{it}$, conditional on $n_{i,t-1}$, $n_{i\neq i,t-1}$, and $e_{i-1}$, is;

$$f(h_{i[1]} = u_1, h_{i[2]} = u_2, h_{it} = h_{i[1]} \mid n_i, n_{i\neq i,t-1}, e) = (e - 1) f_k(u_1) f_k(u_2) F_k(u_2)^{e-2},$$

defined over the relevant intervals of $u_1$ and $u_2$, where $F_k(v)$ is the probability that $h_t$ is less than $v$ given an upper bound of $n_k$, and $f_k(v) = F_k(v)$.\(^{12}\)

### 1.2.3 Competitive Equilibrium

A competitive equilibrium is defined as a sequence of allocations \{\(c_t, s_t, y_t, x_{it}(j), L_{it}(j), e_t(j), n_{it}(j)\}\) and prices \{\(P_{it}(j), w_{it}(j), R_t\), \(\forall t \geq 0, \forall i \in \{1, ..., e_t(j)\}, \forall j \in [0, 1]\), such that: (i) households take prices as given and choose \(c_t\) and \(s_t\) to solve (1.1); (ii) in each period $t$, each innovating firm-$i$ in each market-$j$ takes wages as given and chooses $x_{it}(j)$ to solve (1.4); (iii) in each period $t$, the final-good firm takes prices as given and chooses $x_{it}(j)$, \(\forall i \in \{1, ..., e_t(j)\}, \forall j \in [0, 1]\), to solve (1.2); (iv) in each period $t$, each innovating firm-$i$ in each market-$j$ takes $e_t(j)$, $n_{i\neq i,t}(j)$, $g_{t+1}$, and $R_{t+1}$ as given, and chooses $n_{it}(j)$ to maximize (1.6); (v) free entry is satisfied, which implies \(E_t \left( \frac{\pi_{i,t+1}(j)}{R_{t+1}} \right) = 0 \) in each market-$j$ in each period $t$, where $E_t \left( \frac{\pi_{i,t+1}(j)}{R_{t+1}} \right)$ is given by (1.6); (vi) in each period $t$, all goods, labor,

\(^{11}\)The product indicator $j$ will henceforth be dropped unless required for the sake of clarity.

\(^{12}\)This joint distribution is derived in Appendix A.1.1.
and equity markets clear.

The cost functions for innovation faced by innovating firms ensures decisions about whether to innovate and what level of research to undertake are independent of all past innovations. Combined with the lack of any aggregate uncertainty, this implies the competitive equilibrium is stationary, in the sense that \( e_t(j) = e(j) \) and \( n_t(j) = n(j) \), \( \forall t \), in each market-\( j \). The number of firms \( e(j) \), the level of research per firm \( n(j) \), and the gross rate of return on equity \( R \) can be solved for by using the following three conditions:

The free-entry condition:
\[
E \left( \frac{\pi_t(j)}{R} \right) = 0,
\]

The research condition:
\[
\frac{\partial}{\partial n_t(j)} E \left( \frac{\pi_t(j)}{R} \right) = 0,
\]

The Euler condition:
\[
1 = \frac{\beta R}{1 + g},
\]

where the Euler condition is derived from the household’s problem (1.1), and \( g \) refers to the growth rate of the economy.

For the rest of this chapter, I assume firms draw from a uniform distribution - i.e., \( h_i \sim U(0, n_i^\theta) \), where \( n_i \) is the level of research chosen by firm-\( i \), and \( h_i \) is the realization of its random quality draw. This allows equilibrium to be solved for analytically.\(^{13}\) In Appendix A.2 I derive firm-\( i \)’s expected discounted profits and the corresponding equilibrium conditions for a uniform distribution.\(^{14}\) Given the gross interest rate \( R \) and the growth rate of the economy \( g \), the research and free-entry conditions become:

\[
\text{research: } n = \left[ \frac{(1 + g)^{1-\alpha}(1 - \alpha^\varepsilon)\theta(e + \alpha - 1)}{mRe(e + \alpha)} \right]^{\frac{1}{\alpha-1}}, \tag{1.8}
\]

and

\[
\text{free-entry: } z + mn = \frac{(1 + g)^{1-\alpha}(1 - \alpha^\varepsilon)n^\alpha\theta}{Re(e + \alpha)}, \tag{1.9}
\]

where the number of firms \( e \), research per firm \( n \), and all parameters should be understood to be specific to market-\( j \). \( z \) and \( m \) are the fixed cost of innovation and the marginal cost of research.

The growth rate of the economy \( g \) is equal to:
\[
g = \frac{y}{y-1} - 1 = \left( \frac{\int_0^1 A_{-1}[1](j)h_{1}[1](j)dj}{\int_0^1 A^\alpha_{-1}[1](j)dj} \right)^\frac{1}{\alpha} - 1,
\]

where \( A_{[1]}(j) \) is the highest quality in market-\( j \) in the current period, equal to \( A_{-1[1]}(j)h_{1}[1](j) \), and \( h_{1}[1](j) \) is the realization of the best quality draw in market-\( j \) in the current period. For a symmetric economy in which parameter values are common across markets, the growth rate can be shown to reduce

\(^{13}\)The model has been solved for using a two-parameter distribution called the Kumaraswamy distribution, similar to the Beta distribution but with a simple closed form density function. As long as the restriction \( f_i(n_i^\theta) < \infty \) is maintained, all qualitative results hold.

\(^{14}\)Equations 1.8 through 1.13 are also derived in Appendix A.2.
to:
\[ g = E(h_{[1]}^{\alpha}) - 1 = \left( \frac{en^{\alpha\theta}}{e + \alpha} \right)^{\frac{1}{\theta}} - 1. \] (1.10)

Given the economy’s growth rate \( g \), the gross interest rate is:
\[ R = \frac{1 + g}{\beta}. \] (1.11)

Using the free-entry and research conditions, I can now characterize the number of firms \( e \) in each product market and the level of research per firm \( n \) in a symmetric economy:
\[ \text{research: } n = \frac{\beta(1 - \alpha^e)\theta(e + \alpha - 1)}{me^2} \] (1.12)
and
\[ \text{free-entry: } z + mn = \frac{\beta(1 - \alpha^e)}{e^2}, \] (1.13)
where all variables and parameters are common across markets.

1.2.4 Results

In this section I discuss how equilibrium is affected by changes in the values of exogenous parameters and variables. Most recent empirical work (including each of the studies cited above) focuses on within-industry variation over time to identify the relationships they report between competition, innovation, and growth. I therefore keep the economy-wide growth rate and interest rate fixed for most of this section (unless otherwise noted), in order to examine the effects of changes in industry-specific parameter values on outcomes in a single industry. I focus on the effect of changes in \( z \), the cost of introducing an innovation, as the most plausible driver of the relationships identified in the data. In Appendix A.4 I prove the following results hold for all parameter values (for which equilibria exist): (i) the equilibrium number of firms is decreasing in the fixed cost of introducing an innovation \( z \); (ii) the level of competition (measured as in Aghion et al. (2005)) is decreasing in \( z \); (iii) both research per firm and the average level of innovation per firm exhibit an inverted-U relationship with \( z \); and (iv) industry-level productivity growth is decreasing in \( z \). To illustrate these relationships, I show how the number of firms and the level of research per firm (and functions thereof) change across stationary equilibria associated with different fixed costs of introducing an innovation, keeping constant all other parameters and exogenous variables. I use the following parameter values: \( \alpha = 0.816 \) (implying an elasticity of substitution of 5.4), \( \theta = 0.06 \), and \( m = 6 \cdot 10^{-5} \). For industry-specific changes, the economy-wide growth rate and interest rate are also held constant at 0.03% and 5%. The calibration strategy used to obtain these values is explained in Appendix A.5.

Figure 1.1 illustrates the relationships between the level of research per firm \( n \) and the number of firms \( e \) implied by the research (1.8) and free-entry (1.9) conditions, for each product market. The research curve shows the optimal level of research per firm as a function of the number of firms, while the free-entry curves show the level of research per firm necessary to ensure zero expected profits given the number of firms, for different fixed costs of introducing an innovation. For a given level of research, profits are decreasing in the number of firms. At the same time, given the number of firms, profits are 15

In Section 1.3, I discuss direct and indirect evidence that supports the relevance of variation in the fixed cost of innovations.
decreasing in research. As the number of firms increases, therefore, research must decrease to maintain zero expected profits, resulting in the negative relationship between \( n \) and \( e \) illustrated by the free-entry curves. The leftward shift of the free-entry curve associated with an increase in the fixed cost of introducing an innovation \( z \) shows that for a given level of research per firm \( n \), fewer firms can be accommodated under free entry.

![Image of Figure 1.1: Optimal Research per Firm and Free Entry](image)

Figure 1.1: Optimal Research per Firm and Free Entry

The research curve in Figure 1.1 illustrates an inverted-U relationship between the number of firms and each firm’s optimal level of research. From the research condition (1.8), each firm’s optimal level of research is clearly independent of \( z \), except through its effect on the number of firms \( e \). As \( e \) increases from one to two, research always increases. Above some threshold \( e^* \), however, \( n \) decreases with \( e \). To understand this relationship, it is useful to note that an increase in the number of firms \( e \); (i) decreases each firm’s probability of winning; (ii) increases the expected value of a firm’s innovation, given that it wins; and (iii) decreases the expected markup that the winner can charge. Now consider a variation of the model in which the winning firm can charge a monopoly price regardless of how close its competitors may be. In Appendix A.6, I show the optimal level of firm-level research in a monopoly-only model is both increasing in the number of firms, and always higher than in the full model (except, of course, when the market can only sustain one firm). This is the result of the assumption that the number of firms does not affect the winner’s markup. In the full model, an increase in the number of firms will thus increase the optimal level of research given monopoly pricing, but make monopoly pricing less likely. The net result is an inverted-U relationship between \( n \) and \( e \). Note that firm-level research in Figure 1.1 peaks at around three firms. In general, the location of this peak is weakly increasing in the elasticity of substitution.\(^{16}\)

Equations 1.8 and 1.9 can be combined to find the equilibrium number of firms \( e \) as an implicit

\(^{16}\)The relationship between \( e^* \) and the elasticity of substitution is discussed in Appendix A.4.2. Allowing for labor to move freely across industries results in \( n \) peaking with two firms, regardless of the elasticity of substitution.
Figure 1.2: Number of Firms and Research per Firm with variation in $z$

function of the fixed cost of introducing an innovation $z$;

$$z = \left[ \frac{(1+g)^{1-\alpha}(1-\alpha^e)}{Re(e+\alpha)} \right]^{\frac{1}{1-\alpha\theta}} \left[ \frac{\theta(e+\alpha-1)}{m} \right]^{\frac{\theta}{1-\alpha\theta}} \left[ 1 - \theta(e+\alpha-1) \right].$$  \hspace{1cm} (1.14)

Figure 1.2a plots the equilibrium values of the number of firms $e$ for different values of $z$, while Figure 1.2b does the same for the level of research per firm $n$.\(^{17}\) Not surprisingly, free entry ensures the number of firms $e$ adjusts downwards as the fixed cost of innovating increases. With $e$ a monotonic function of $z$, the qualitative relationship between $n$ and $z$ is just the reverse of that between $n$ and $e$. The result is the inverted-U relationship in Figure 1.2b. Although not illustrated here, the research intensity of an industry (the fraction of total output spent on R&D) can be expressed simply as $e \cdot n \cdot m$. While Figure 1.2b illustrates an inverted-U between firm-level research and $z$, industry-level research intensity decreases monotonically from a high of 6% to a low of 1%, as $z$ increases over the illustrated range.

To analyze the relationships between productivity growth, innovation, and competition, I first find the equilibrium values of each variable for different values of $z$, the cost of introducing an innovation. I then plot these variables against each other, where each point represents the set of equilibrium values associated with a particular $z$. Since there is only one representative industry and population is fixed at one, the rate of total factor productivity growth in an industry is simply equation 1.10;

$$g = E(h_{[1]}^\alpha)^{\frac{3}{2}} - 1 = \left( \frac{en^{\alpha\theta}}{e+\alpha} \right)^{\frac{3}{2}} - 1,$$

where $e$ and $n$ are now specific to the industry in question. The average level of innovation per firm is the expected value of each draw;

$$Average \ Innovation = E_{i=1} \left( \frac{A_{ii}(j)}{A_{i-1}[j]} \right) = E(h_i) = \frac{n^\theta}{2}.$$  \hspace{1cm} (1.15)

\(^{17}\)The range of values for $z$ has been chosen to generate the range of values for measured competition reported in Aghion et al. (2005). This variation in competition is illustrated in Figure 1.3.
Following Aghion et al. (2005), I use one minus the average Lerner Index in an industry as the measure of competition in that industry. The Lerner Index is equal to the ratio of price minus marginal cost over price for the best firm, and zero for all other firms.\(^{18}\) The value of this measure of competition for an industry is thus:\(^{19}\)

\[
\text{Competition} = 1 - \frac{1}{e} E \left( \frac{\text{Price} - \text{Marginal Cost}}{\text{Price}} \right) = \frac{e^2 + \alpha e - 1}{e^2}.
\] (1.16)

Figure 1.3a plots the level of innovation per firm against competition. Each point represents the values of both variables associated with a particular value of \(z\), the fixed cost of introducing an innovation.\(^{20}\) Competition is monotonically decreasing in \(z\), so average innovation per firm has an inverted-U relationship with competition just as did research per firm in Figure 1.2b.

\[\text{(a) Firm-Level Innovation} \quad \text{(b) Industry-Level Productivity Growth}\]

Figure 1.3: Competition, Innovation, and Growth, with variation in \(z\)

Figure 1.3b plots industry-level productivity growth against competition as the cost of introducing an innovation varies. Even while firm-level innovation follows an inverted-U, industry-level productivity growth always increases with competition. The decrease in the expected markup of the winning firm due to more competition tends to discourage research, all else equal, since a drop in the fraction of value captured by the winner means firms are less interested in winning the market \(ex\ antee\) — this is the Schumpeterian effect. But competition also tends to increase the expected value of the best innovation — this is the Hayekian effect. At low levels of competition, an increase in competition is associated with more firm-level research, more innovations, and higher growth. At relatively high levels of competition, where the Schumpeterian effect causes firm-level research and average innovation to decline with competition, the Hayekian effect more than compensates, increasing the expected value of the best innovation, and thus increasing industry-level productivity growth. The range of productivity growth rates illustrated in Figure 1.3b is comparable to the range across 6-digit manufacturing industries in the

\(^{18}\) This means measured competition is equal to \(1 - E(Lerner)/e\), where \(Lerner\) is the Lerner Index for the winning firm and \(e\) is the number of innovating firms. A better industry measure might weight each firm’s Index by its market share, so that competition would be measured here as \(1 - E(Lerner)\). The choice does not affect any qualitative results as both measures are increasing in the number of firms, and so I use the same measure as Aghion et al. (2005) for the sake of easier comparison.

\(^{19}\) Equation 1.16 is derived in Appendix A.3, as is the alternative measure of competition discussed above.

\(^{20}\) The range of values used for \(z\) is the same as that in Figure 1.2.
Chapter 1. Competition as a Discovery Procedure

U.S. Average productivity growth from 1959 to 2005 ranged from -2% for Other Apparel Manufacturing to +13% for Electronic Computer Manufacturing, with a full one third of industries experiencing a decline in productivity over the period.\footnote{Productivity growth here refers to five-factor productivity growth, as reported in NBER-CES Manufacturing Industry Database.}

The experiment illustrated in Figures 1.2 and 1.3 (changing equilibrium outcomes by changing \( z \)) can be repeated for each of the exogenous parameters in the model. Given that the number of firms \( e \) is a decreasing function of \( z \) (as illustrated in Figure 1.2a), (1.14) implies \( e \) is decreasing in the marginal cost of research \( m \), as is research per firm \( n \). Both \( e \) and \( n \) have an inverted-U relationship with the elasticity of substitution. The parameter \( \theta \) governs the elasticity of each firm’s upper bound with respect to research, as each firm-\( i \) draws from a distribution with support \([0, n^0_i]\). Research per firm \( n \) is increasing in \( \theta \), while the number of firms \( e \) is decreasing. An increase in the interest rate \( R \) or decrease in the economy’s growth rate \( g \) (both exogenous with respect to a single industry) lower discounted profits and therefore result in fewer firms and less research per firm in equilibrium.

The effects of economy-wide changes in exogenous parameters and variables (i.e., when the interest rate and growth rate are endogenous) are the same as those for industry-level changes. The only difference of note is that research (and innovation) per firm is always highest with two firms per industry, regardless of the elasticity of substitution. The intuition here is that when the cost of introducing an innovation decreases, the economy-wide growth rate increases and brings the interest rate \( R \) up with it. The net effect is to decrease expected profits for firms, which puts additional downward pressure on the incentive to do research.

It is worthwhile to digress here to discuss the difference between the mechanism that generates an inverted-U in the present model, and that in Aghion et al. (2005). In Aghion et al., the positive link between competition and innovation (over the initial section of their inverted-U) is driven by neck-and-neck firms innovating more to escape lower rents. With free entry, however, lower rents would presumably lower the incentive to innovate for lagging firms, thus dampening or even wiping out the positive effect of lower rents on the average level of innovation. Indeed, Etro (2007) finds the “escape competition” effect disappears completely when he incorporates free entry into a number of variants of the model in Aghion et al. In the present paper, free entry itself causes aggregate innovation to increase with competition, due to the Hayekian effect (by increasing the expected value of the best innovation for a given level of research per firm), while also offsetting the disincentive to innovate at the firm level at low levels of competition.

1.3 Empirical Support

In the preceding section, I focus on the equilibrium effects of a change in the cost of introducing an innovation \( z \). This focus is appropriate if changes in \( z \) are actually driving the relationships reported in empirical studies. In fact, each of the studies cited throughout this paper control for both year and industry fixed effects, and so are presumably reporting relationships driven by within-industry variation in an exogenous parameter over time. The regulatory portion of the fixed cost of introducing an innovation to the market is at least a plausible candidate for the source of this variation. Evidence from Aghion et al. (2005) more directly supports a focus on regulatory costs. In an attempt to control for endogeneity in their test of the effect of competition on firm-level innovation, Aghion et al. employ...
a series of stuttered waves of deregulation in different industries as a source of exogenous changes to competition. In the end, their IV results are almost identical to their OLS results (controlling for time and fixed effects) with a reduced-form $R^2$ of 0.8, suggesting that much of the correlation between competition and firm-level innovation can be explained by the effect of lower barriers to entry on the number of firms and thus the optimal level of innovation. Variation in the regulatory cost of entry maps well to the fixed cost $z$ of introducing an innovation to the market in the model. Another plausible source of the variation in competition in Aghion et al. is variation in import tariffs, which the authors also use as an instrument. This channel is neglected in the present paper, but could be examined in an extension to the model.

The primary implication of the model developed in this chapter, that competition and productivity growth are positively related, is well supported in empirical studies of industry-level growth like Nickell (1996), Blundell et al. (1999), and Aghion et al. (2008). To my knowledge, no empirical studies have provided evidence to the contrary. In addition, studies like Graham et al. (1983) and Nicoletti and Scarpetta (2003) have suggested both higher marketing costs and more burdensome regulations are associated with lower growth, which is consistent with the model. Aghion et al. (2005) test the relationship between competition and the average level of innovation per firm and report an inverted-U relationship, consistent with the present model.

In recent publications, economists have interpreted evidence of an inverted-U relationship between firm-level innovation and competition as evidence of a similar relationship between industry-level productivity growth and competition. Aghion et al. (2008) present evidence to the contrary. Their results are consistent with those of the model presented here, where variation in the cost of innovating induces positively correlated differences in competition and productivity growth at the industry level, but an inverted-U relationship between competition and firm-level innovation.

Boudreau et al. (2011) provide evidence of the plausibility of the central mechanism of this chapter - that a larger number of innovators will tend to increase the value of the best innovation, notwithstanding the disincentive effect on effort. The authors look at data from 9,661 software algorithm contests in which competitors try to solve a well-defined problem. All proposed solutions are observed and their quality measured. Multiple contests are held for the same problem with exogenous variation in the number of competitors. In contests where the best approach to solving a problem is uncertain, the authors find the quality of the best solution is increasing in the number of competitors, even when the average quality is decreasing.

The central mechanism of this chapter generates one additional implication. In Appendix A.4.5 I show that the average distance of firms within an industry from the frontier (i.e., the best firm in an industry) is increasing with competition. This is supported by Aghion et al. (2005), who report industries become less ‘neck-and-neck’ as measured competition increases.

The implications of the model with respect to changes in parameters other than fixed costs are broadly consistent with both existing models and the empirical literature on the determinants of R&D, market structure, and productivity growth. Vives (2008) and Etro (2007) provide detailed analyses of the various determinants of R&D and growth for a number of different theoretical models. Cohen and Levin (1989) and Gilbert (2006) provide excellent surveys of empirical studies looking at the effects of differences in appropriability, technological opportunity, and market size.

\footnote{See Bianco (2007), for example. Indeed, this view seems to have become the conventional wisdom.}
1.4 Extension: Antitrust Policy

In this section, I use the model to consider the effects of antitrust policies on innovation and productivity growth. An advantage of using a model in which competition is endogenous is the ability to specify exactly how various policies affect exogenous constraints or parameters in the model and evaluate the resulting outcomes, rather than merely assume antitrust policy exogenously increases competition. The baseline model developed in Section 1.2 implies that deregulation should increase competition and productivity growth. Where antitrust policy works to lower legal barriers to innovation and market entry, the model predicts a higher aggregate level of innovation will result. In contrast it is straightforward to show that a legislated ceiling on markups (enforced, for example, under ‘Abuse of Dominant Position’ provisions of antitrust legislation) will discourage both entry and innovation, resulting in less growth. In the following subsection, I extend the model to evaluate the effects of legal restrictions on collusion or (equivalently) ‘anti-competitive’ mergers. Whereas empirical studies of competition and productivity growth generally conclude with calls for greater antitrust enforcement (on the presumption that greater enforcement implies more competition, which is found to be positively correlated with growth)\(^{23}\), I find that restrictions on collusion (as with restrictions on markups) lead to lower productivity growth in equilibrium. The finding that restrictions on markups and collusive behavior are detrimental to innovation and growth is consistent with conventional models of innovation. The contribution of this section is to show that even in a model where measured competition is positively related to growth, restrictions on firm behavior nevertheless result in lower growth.

1.4.1 Collusion and Anti-Competitive Mergers

This section extends the baseline model by allowing the winning firm in any period to either purchase or collude with any other firm that might constrain the winning firm’s pricing decision. Williamson (1968), Mathewson and Winter (1987), and others have considered situations in which price-fixing and mergers thought to be anti-competitive might actually improve allocative efficiency. The possible benefits of joint research ventures with respect to technology diffusion have also been analyzed in studies like Katz (1986). The contribution of this section is to evaluate the effects of restrictions on collusion and mergers for which no efficiency defence exists - that is, behavior undertaken by firms that increases profits at the expense of consumers \textit{ex post}, and provides no benefit to allocative efficiency. I simplify the model by assuming an exogenous common upper bound \(\lambda > 1\) for the distribution firms draw from, as allowing for each firm to choose its own level of research adds little additional insight.

I model the cost of purchasing or colluding with all firms that pose a threat as an exogenous fraction \(\tau^C\) of the additional profit firm-[1] (the winning firm) stands to gain by charging a monopoly price, rather than a limit price. In addition, I assume firm-[1] must incur a cost to monitor its co-conspirators (or manage its larger size after merging), which I model in a similar fashion as a fraction \(\tau^M\).\(^{24}\) Antitrust policy here can be interpreted as a marginal increase in either of these costs, or simply as an assumption that antitrust enforcement makes either of these costs high enough to make collusion always a bad idea for firms.

\(^{23}\)Examples of this particular policy proposal can be found in the concluding sections of Geroski (1990) and Dutz and Hayri (2000).

\(^{24}\)It seems intuitive that each of these costs is increasing with the payoffs to collusion. The larger the payoff to the best firm from colluding, the higher the payment that its rivals could presumably extract. And the further a collusive price is from the competitive limit price, the greater is the incentive for rivals to cheat, which would presumably increase the cost of monitoring.
The total cost of achieving a monopoly price when \( \frac{h_{2\tau}}{h_{1\tau}} > \alpha \) is:

\[
(\tau^C + \tau^M) \frac{\psi(h_{2\tau} - \alpha h_{1\tau})}{h_{1\tau}^{1-\alpha}},
\]

where \( \psi \equiv A_{-1\tau}^\alpha L_1^\alpha \psi_{1-\alpha} \), \( A_{-1\tau} \) is the first-best quality of the previous period, and \( h_{\ell\tau} \) is the \( \ell \)-th-best draw of the current period. The payoff for each firm-[\( r \)] for which \( \frac{h_{2\tau}}{h_{1\tau}} > \alpha \) is equal to the fraction \( \tau^C \) of the additional profit the winning firm receives by colluding with firm-[\( r \)]. The payoff for an eligible firm-[\( r \)] is therefore:

\[
\tau^C \frac{\psi(h_{[r]} - \max\{h_{[r+1], \alpha h_{[1]}}\})}{h_{1\tau}^{1-\alpha}}.
\]

When \( \frac{h_{2\tau}}{h_{1\tau}} > \alpha \), the operating profits of a winning firm choosing a limit price are \( \frac{\psi(h_{1\tau} - h_{2\tau})}{h_{1\tau}^{1-\alpha}} \), while those of a winning firm choosing to collude are \((1 - \alpha)\frac{h_{1\tau}^\alpha}{h_{1\tau}^{1-\alpha}} - (\tau^C + \tau^M) \frac{\psi(h_{2\tau} - \alpha h_{1\tau})}{h_{1\tau}^{1-\alpha}}\). Firm-[1] will therefore choose to collude if \( \frac{h_{21}}{h_{11}} > \alpha \) and;

\[
(1 - \alpha)\frac{h_{11}^\alpha}{h_{11}^{1-\alpha}} - \frac{h_{21}}{h_{11}^{1-\alpha}} > (\tau^C + \tau^M) \frac{h_{21}^\alpha - \alpha h_{11}^\alpha}{h_{11}^{1-\alpha}},
\]

or if \( \tau^C + \tau^M \leq 1 \). Since the decision to collude or not depends only on \( \tau^C \) and \( \tau^M \) (whenever \( \frac{h_{21}}{h_{11}} > \alpha \)), I focus on the interesting case where \( \tau^C + \tau^M \leq 1 \).

I now derive the expected discounted profits facing each innovating firm. Because the payoffs to eligible losing firms are straight transfers and firms have no ability to affect their expected payoff (or payment) \textit{ex ante}, I need not account for payoffs or payments. The expected discounted profits of an innovating firm-i are thus:

\[
E \left( \frac{\pi_i}{R} \right) = \frac{(1 - \alpha)\psi(1 + g)^{1-\alpha}}{R} \cdot \text{Prob} \left[ h_i = h_{11} \right] \cdot E \left[ h_{11} \right] \cdot \text{Prob} \left[ h_{11} = h_{11} \right] - \frac{\psi(1 + g)^{1-\alpha}}{R} \cdot \text{Prob} \left[ h_i = h_{11} < \frac{h_{21}}{\alpha} \right] \cdot E \left[ h_{21} - \alpha h_{11} \left/ h_{11}^{1-\alpha} \right. \right] \cdot h_{11} = h_{11} < \frac{h_{21}}{\alpha} - \psi(z + mn_i),
\]

where the first term is the expected monopoly profits from production, and the second term is the expected cost of monitoring the winning firm’s co-conspirators. Using the density function derived in Appendix A.1.2 and setting the expected value of entry equal to zero, the collusive equilibrium can now be characterized by the following \textit{free-entry} condition:

\[
\textit{free-entry: } z = \frac{(1 + g)^{1-\alpha} \lambda^\alpha}{R} (e(1 - \alpha) - \tau^M [e(1 - \alpha) + \alpha^e - 1]), \quad (1.17)
\]

Note that if monitoring costs eat up the entire benefit of collusion \( \tau^M = 1 \), then this condition reduces to that of a competitive equilibrium. Further, the number of innovators is independent of \( \tau^C \), conditional on \( \tau^C + \tau^M \leq 1 \).

The right-hand side of equation 1.17 is simply the expected discounted operating profits of an innova-
tor divided by a constant ($\psi$), and so is decreasing in the number of firms $e$. A higher cost of monitoring $\tau^M$ reduces expected profits (given $e$), and so the number of firms must be decreasing in $\tau^M$ to satisfy the free-entry condition. The growth rate for an industry is $g = \left( \frac{e^{\lambda_0}}{e^{\lambda_0} + \alpha} \right)^{\frac{1}{\alpha}} - 1$ (analogous to equation 10), which is increasing in $e$. It follows that $g$ must be decreasing in $\tau^M$. Given that the growth rate for a collusive industry is equal to that in a non-collusive industry when $\tau^M$, it must be the case that growth is higher if collusion is possible, as long as $\tau^M < 1$.

Testing these results against real outcomes is difficult, due to the ubiquitousness of antitrust laws throughout the last century. The story of Standard Oil in the late nineteenth and early twentieth centuries, however, does provide a case study that seems consistent with the implications of the present model. Over a forty-year period, the company used a combination of process and product innovations, acquisitions, and price-fixing agreements to dominate the market, maintaining a market share of close to 90% until both politics and more able rivals began to drag them back down to earth. Just as the model predicts, a large number of new entrants appeared each year, many seemingly with the intention of becoming just competitive enough to get bought-out by Standard Oil. While some refineries were used to add to capacity, many were just bought and then shut down in an effort to reduce competition. Nevertheless, the period saw an enormous number of new process innovations, as well as new ways to turn the ‘waste’ from the production of kerosene into products like gasoline, paving tar, and petroleum jelly. Over the course of thirty years, just as Standard Oil first acquired and then maintained its monopoly, the deflation-adjusted price of kerosene dropped by over 65%, even while petroleum output increased by more than a factor of three.

Finally it is important to point out that the model is ill-suited for evaluating the welfare implications of collusion, as it does not allow for any static inefficiency resulting from a monopoly price (since labor is immobile). That collusion may be welfare-reducing even while increasing growth is most obvious in the limiting case where monitoring costs erode the entire benefit of charging a monopoly price. When $\tau^M$ is equal to one, growth will be no higher than in a competitive industry, but resources will nonetheless be wasted enforcing the collusive agreement. The policy implication of this model is that regulators should either allow price-fixing agreements to be enforced, thereby lowering monitoring costs, or else raise monitoring costs to a prohibitive level. The appropriate choice depends on the trade-off between dynamic and static efficiency (as well as on enforcement costs and ideology).

1.5 Conclusion

Hayek (2002) argued that the competitive process could be thought of as a procedure for discovering and making use of knowledge that would otherwise not emerge. When firms are uncertain about the optimal direction of innovation, the best innovation to emerge will tend to be of higher value when more innovations are tried. Throughout this chapter the Schumpeterian mechanism, whereby lower \textit{ex post} rents discourage lower \textit{ex ante} research, is always present. But when Schumpeter meets Hayek in this model of innovation, more competition is accompanied by higher industry-level productivity growth, even if research per firm is lower. Combining Hayekian uncertainty with endogenous entry, markups, and firm-level research results in a model able to generate a positive relationship between competition and

\footnote{McGee (1958) provides a detailed analysis of the competitive behavior of Standard Oil and its rivals. Boudreaux and Folsom (1999) provide a short summary of the innovations and price reductions that accompanied J.D. Rockefeller’s entertaining quest to dominate the market.}
industry-level productivity growth, and an inverted-U relationship between competition and firm-level innovation.

By treating competition, growth, and innovation as endogenous, I develop a framework for evaluating competition policy that is less susceptible to the Lucas Critique. Two such evaluations of antitrust policies (restrictions on markups and price-fixing) come to conclusions very much at odds with policy prescriptions based on simple interpretations of empirical studies. The results of recent empirical studies notwithstanding, I conclude that one of Schumpeter’s central messages should be taken to heart — regardless of any static benefits of antitrust enforcement and regulation, these policies come with the cost of less growth. Or alternatively, there ain’t no such thing as a free lunch.
Chapter 2

Patent Protection as a Tax on Competition and Innovation

Abstract:
I introduce patents into a general equilibrium model of innovation, where innovators choose between creating a new product market and entering an existing market. Patent holders demand royalties from sequential innovators, but are constrained by the ability of innovators to work around patents. I show patent protection acts as a net tax on innovation and reduces productivity growth, consistent with recent evidence. Calibrated to match moments from U.S. data, the model predicts that eliminating patent protection in the U.S. would generate a 23% increase in steady-state productivity growth as well as an increase in welfare equivalent to that from a 16% increase in annual consumption. I test several implications of the model using patent data from the U.S. and an index of patent protection across countries. Consistent with the model, the data suggests an increase in the strength of patent protection reduces the average quality of innovations, the number of innovations, and the overall level of innovation.
2.1 Introduction

The last decade has seen renewed interest in two strands of the innovation literature, and in more recent years the two have started to interact. Aghion et al. (2005) renewed interest in the relationship between competition and innovation, resulting in a wealth of both empirical and theoretical studies exploring the nuances of this relationship and its policy implications. A series of papers (as well as a book) by Boldrin and Levine (2002, 2005, 2008a, 2008b) questioning the need for and efficacy of patent rights have at the same time inspired a growing number of empirical and theoretical studies investigating the relationship between patent protection and innovation.

Boldrin and Levine (2008b) argue that the large and still growing number of studies finding a positive relationship between competition and innovation is prima facie evidence against the presumed need for patent rights to encourage innovation. Meanwhile Aghion, Howitt, and Prantl (2013b) present evidence that the pro-innovation effects of product-market deregulation are increasing in the strength of patent rights, and argue Boldrin and Levine may be overlooking some potential benefits of patent protection. Given the inability of current workhorse models of R&D-based growth to generate a positive link between competition and innovation, they are of limited value in evaluating Boldrin and Levine’s claim. More generally, the models currently used to discuss patent policy are of limited use in thinking about ‘competition’, as traditionally defined (the number of actual or potential competitors). In extensions of Grossman and Helpman (1991) the number of firms is indeterminate, while in IO-style models the number of firms is fixed at two.\footnote{For extensions of Grossman and Helpman (1991) see, for example, Acemoglu and Akcigit (2012), Riis and Shi (2012), and Parra (2013). Examples of IO-style models include Bessen and Maskin (2009) and Aghion, Howitt, and Prantl (2013a).}

In this chapter I bring together ideas from both the competition and innovation literature and the patent rights literature to build a model which captures many of the key facts about competition, innovation, and the effects of patent rights documented in empirical studies. In the model, product-market deregulation results in more competition, innovation, and productivity growth (Bottasso and Sembenelli (2001), Nicoletti and Scarpetta (2003), Griffith, Harrison, and Simpson (2010)); more patent protection leads to less innovation and growth (Lerner (2009), Moser and Voena (2012), Moser (2013)); and product-market deregulation has a larger impact when patent rights are strong (Aghion, Howitt, and Prantl (2013b)). The model is ambiguous with respect to the effect of patent protection on welfare, as the lower growth caused by stronger protection is accompanied by a higher consumption share of output and an increase in the variety of products available. To estimate the welfare costs (or benefits) of patent protection in practice, I calibrate the model using U.S. manufacturing data. The calibrated model predicts that the elimination of patent rights in the U.S. would generate a 23% increase in productivity growth and a 21% drop in the number of product varieties, resulting in a net \textit{increase} in steady-state welfare equivalent to that from a 16% increase in annual consumption.

An additional implication of the model is that stronger patent protection should lead to a decrease in the average quality of innovations. Using data on U.S. patent applications originating in foreign countries, I provide new evidence that the average number of citations per patent (a proxy for the average quality of innovations) is indeed driven lower when patent protection is made stronger. In addition I provide corroborating evidence of the previously documented negative effects of patent protection on the number of patents (Lerner (2009)) and the number of citation-weighted patents, a proxy for the overall level of innovation (Moser and Voena (2012), Moser (2013)).

The model economy developed in this chapter is similar to Chu, Cozzi, and Galli (2012) in that the
number of differentiated products (varieties) within an industry is endogenously determined by a free entry condition. Competition within each differentiated product market is modeled as in Chapter 1. Each period, firms compete within a product market by introducing improved versions of an existing product. These sequential innovators conduct research to increase the magnitude of their respective improvements, but are ex ante uncertain about the optimal direction of a quality improvement. Once the cost of research and the cost of introducing a product into a market are sunk, firms discover the value of their own improvement and those of their competitors. By allowing for the ‘best’ innovation to capture a market I create a link between the number of innovations and the level of quality growth within a product market. This mechanism captures the insight of Hayek (2002) that competition allows a thousand ideas to bloom (to mix metaphors), then works to discover the best among them.

A key feature of the model is that patent protection acts as a tax on sequential innovation and competition. That this must be the case follows directly from the fact that the higher costs incurred by sequential innovators improving an existing product today (in the form of royalty payments to patent holders) are larger than the expected discounted revenue from royalty payments in the future, a point made by Chu, Cozzi, and Galli (2012). As these higher costs are effectively transfers from sequential innovators to creators of new differentiated products (who I assume do not infringe on previous patents), stronger patent protection involves a reallocation of resources away from growth to maintain a greater variety of products. As a result, patent protection is welfare-improving only if growth is otherwise ‘too high’.

This chapter contributes to the theoretical literatures on competition, innovation, and patent policy, as well as the empirical literature examining the efficacy of patents in encouraging innovation and growth. Aghion et al. (2005) develop a model where some measure of competition can be positively correlated with innovation and growth, but do not allow for the free entry of firms. Boldrin and Levine (2008a, 2009) develop models in which innovation occurs (and can even be optimal) under perfect competition, but are unable to speak to the relationship between innovation and differing levels of competition.

Scotchmer (1991) analyzes optimal patent policy when innovation is sequential, as do a large number of subsequent papers like Bessen and Maskin (2009), Acemoglu and Akcigit (2012), Riis and Shi (2012), and Parra (2013). In each of these papers the number of innovators is either fixed at two or indeterminate, while the model developed here features an endogenous number of sequential innovators.

In the next section I describe the model, characterize the competitive equilibrium, and discuss its key implications. In Section 2.3 I calibrate the model and show that eliminating patent protection in the U.S. would increase welfare. In Section 2.4 I provide empirical evidence supporting the implications of the model with respect to patent rights and innovation. The final section concludes.

2.2 The Model

Consider an economy in which a final good is produced using a variety of inputs from a representative intermediate industry. Intermediate firms (hereafter referred to as ‘firms’) produce these inputs one-for-one using labor. As such, output per worker in this economy is equivalent to aggregate total factor productivity. The final good can be used for consumption or innovation, and will also act as the numéraire. There are a large number of potential innovating firms, any of which can choose to introduce either a new product (thereby creating a new product market), or an improved version of an existing product. Each product market faces an exogenous probability of destruction each period, so new product
markets continue to be created in steady state. I study the stationary competitive equilibrium of the economy along a balanced growth path, in which firms take the economy-wide wage, growth rate, and interest rate as given, and free entry ensures zero expected profits for all entrants. I begin by describing the environment.

2.2.1 Environment

There is a representative consumer who supplies one unit of labor to intermediate firms. The consumer only values consumption \( C \) and has a constant discount factor \( \beta \in (0, 1) \). Preferences over the stream of consumption in each period are described by the following utility function:\(^2\)

\[
\sum_{t=0}^{\infty} \beta^t u(C_t).
\]

The market for final output is perfectly competitive, with a representative firm using inputs from a representative intermediate industry to produce output according to the following production function:\(^3\)

\[
y = \left( \int_0^J A^{1-\alpha}(j)x^\alpha(j)dj \right)^{\frac{1}{\alpha}},
\]

where \( j \) indexes a continuum of product markets of measure \( J \), \( x(j) \) is the quantity of input \( j \) used, \( A(j) \) is the quality of input \( j \) (or alternatively, the productivity of input \( j \) in final good production), and \( \frac{1}{1-\alpha} \) is the constant elasticity of substitution between differentiated products. Under the above assumptions, \( y \) is equivalent to both output per worker and aggregate total factor productivity.

I assume different products within a product market are perfectly substitutable. I further assume some epsilon fixed operating cost and that the firm with the 1st-best quality \( A = A[1] \) can commit to charging a limit price if any rivals attempt to produce, thus ensuring only the best firm in a market produces in equilibrium.\(^4\)

Finally, I assume product markets are destroyed with probability \( \delta \) each period after production takes place but before any costs associated with innovation are incurred.

2.2.2 Innovation

In each period, after both intermediate inputs and the final good have been produced, firms can choose to undertake the investment necessary to create and introduce a new or improved product. I first describe the innovative process for firms introducing an improved product into an existing product market.

Existing Product Market

Denote the quality of an incumbent’s product in period \( t \) as \( A_{t[1]}(j) \), where [1] indicates that the incumbent has the 1st-best quality in period \( t \) in market \( j \). Any firm-\( i \) wishing to incur the cost of

---

\(^2\)I focus my attention on the stationary equilibrium of the economy, so there is no need here to specify the form of \( u(\cdot) \).

\(^3\)Throughout the paper, I omit the time subscript unless clarity requires it.

\(^4\)In Chapter 1 I consider a market with no operating cost, where the best firm charges a price equal to the marginal cost of its closest rival (adjusted for quality) in equilibrium. Incorporating these variable markups in the present model would complicate the analysis without contributing any insights to those already discussed in Chapter 1.
introducing an improved version of product $j$ in the subsequent period receives a quality equal to $A_{t+1} = A_t[h_i]$, where $h_i$ is a random variable bounded by 1 and $1 + n_i$, and $n_i$ is chosen by the firm. I assume the cost of $n_i$ is $\Psi_A(j) \cdot n_i^{\theta}$, where $\theta > 1$ is the elasticity of research expenditure with respect to the upper bound chosen by the firm. Before realizing the value of its draw (as well as the draws of other firms) the innovating firm must also incur a fixed cost of introducing its product equal to $\Psi_A(j) \cdot z$. One can think of the level of research as determining the magnitude of a quality improvement, and the randomness of the draw reflecting a firm’s ex ante uncertainty about which characteristics of the product to improve upon.\footnote{Throughout the paper, I will often refer to a firm’s choice of $n$ as that firm’s ‘level of research’.
} Only once all of these costs are sunk do firms learn their own and each other’s quality. By multiplying these costs by $\Psi_A(j) \equiv y_t \cdot A_t[j]$, I ensure innovating firms will face the same decision each period in steady state.

**New Product Market**

Any firm choosing to introduce a new product (thus creating a new product market) in period $t + 1$ must incur a fixed cost in period $t$ equal to $y_t \cdot z_0$, where the 0-subscript reflects the fact that a new market is being created. I assume such a firm receives a quality equal to the industry average, $E(A_{t+1}[1])$.\footnote{The model can easily be extended to allow for a research decision by new-product innovators, but such an extension complicates the model without contributing any interesting implications.} For an equilibrium with both new-product and sequential innovation to exist, $z_0$ must be sufficiently large relative to $z$.

**2.2.3 Patents**

Upon introducing a product into a market, a firm can costlessly obtain a patent. I assume firms introducing new products never infringe on existing patents, while sequential innovators always do. I further assume patent holders commit ex ante to charging a royalty fee to subsequent infringers equal to some fixed fraction $\rho$ of both operating profits and any future royalty revenue the infringer may receive. I assume an innovator who pays these royalty fees to a patent holder takes ownership of the previous patent, so that a patent holder producing in period $t$ only receives royalty payments from an innovator producing in period $t + 1$ (who then passes along payments from the next innovator, etc...). At the same time I assume a patent holder’s choice of $\rho$ is constrained by the ability of sequential innovators to work around a patent (and thus not infringe) by incurring larger innovation costs, equal to $\hat{\rho} \cdot \Psi_A(j) \left( z + n_i^{\theta} \right)$, $\hat{\rho} > 1$. I interpret $\hat{\rho}$ as a measure of the strength of patent protection. But given the royalty fee $\rho$ will be a strictly increasing function of $\hat{\rho}$, from this point on I simply use $\rho$ as my measure of patent protection.

**2.2.4 Competitive Equilibrium**

I focus on the stationary competitive equilibrium of the model. In such an equilibrium, the interest rate $r$ and growth rate $g$ are constant, as is the measure of product markets $J$, and the wage $w$ grows at the same rate as total output. In addition, the assumptions made about the environment above will ensure both a constant number of sequential innovators $e$ per market in each period and a constant level of research per (sequential) innovation $n$. I begin by describing the decision problems of each agent, and then define and solve for the stationary equilibrium.
Chapter 2. Patent Protection as a Tax on Competition and Innovation

Consumer

In each period the consumer chooses both consumption and savings, and the only vehicle for savings is the purchase of equity in innovating intermediate firms, earning a rate of return of \( r \).\(^7\) The consumer’s problem is therefore to choose consumption \( C \) and savings \( S \) in each period \( t’ \), given \( w \), \( r \), and \( g \), to maximize:

\[
\sum_{t=t’}^{\infty} \beta^t u(C_t), \quad \text{s.t.} \quad C_t + S_t \leq w_t(1 + g)^{t’ - t} + S_{t-1}(1 + r).
\]

The first order conditions for this problem imply the following interest rate:

\[ r = \frac{1 + g}{\beta} - 1. \]

Final-Good Producer

In each period, the final-good producer takes the prices of all intermediate inputs as given, and demands inputs from each intermediate firm to maximize profits:

\[ y - \int_0^J P(j) x(j) dj, \]

where \( y \equiv \left( \int_0^J A^{1-\alpha}(j)x^\alpha(j) dj \right)^{\frac{1}{\alpha}} \) and \( P(j) \) is the price of input \( j \). The first order conditions for the final-good firm’s problem imply the following inverted demand function for each intermediate input:\(^8\)

\[ P(j) = y^{1-\alpha} A^{1-\alpha}(j) x^{\alpha-1}(j). \]

Intermediate Firms

In each period, once the quality of each innovator is realized, all but the highest-quality firm in each product market will choose not to produce (given the assumptions above). The remaining firms face the downward-sloping demand curves implied by the final-good firm’s problem above, and demand labor \( x(j) \) given the wage \( w \), to maximize operating profits:

\[ \pi(j) = P(j) x(j) - wx(j). \]

First order conditions imply an optimal price equal to:

\[ P(j) = \frac{w}{\alpha}, \]

and optimal output equal to:

\[ x(j) = y A(j) \left( \frac{\alpha}{w} \right)^{\frac{1}{1-\alpha}}. \]

Together, these imply a firm with quality \( A(j) \) earns operating profits equal to:

\[ \pi(j) = y A(j)(1 - \alpha) \left( \frac{\alpha}{w} \right)^{\frac{\alpha}{1-\alpha}}. \]

---

\(^7\)I take it as given that the consumer will diversify across all innovating firms, as they will all share the same expected value before innovating.

\(^8\)I have taken it for granted here that only one intermediate firm will produce in each product market.
A sequential innovator that realizes the highest quality $A$ in market $j$ will earn a fraction $1 - \rho$ of the above operating profits (with the remainder transferred to the previous patent holder), as well as a fraction $\rho(1 - \rho)$ of the operating profits and royalty revenue of the following period’s best firm (again, with $\rho^2$ being transferred to the previous patent holder). The value of introducing an improved product into an existing market $j$ for some firm-$i$ in period $t$ can now be expressed as:

$$ V_{i,t-1}(j) = -\Psi_{t-1}(j) \left( z + \frac{n^0_{i,t-1}(j)}{\theta} \right) + \text{Prob}(A_{i,t}(j) = A_{t-1}(j)) \frac{(1 - \rho)(1 - \alpha)\alpha^{t-\alpha}}{1 + r} \sum_{t'=t}^{\infty} E(A_{t'[1]}(j)|A_{i,t}(j) = A_{t-1}(j)) \frac{(1 - \delta)^{t'-t} \rho^{t'-t} y_{t'}}{(1 + r)^{t'-t} w_{t'-1}^{t'-\alpha}}. $$

By taking for granted that the wage grows at the same rate as final output, and using $A_{i,t}(j) = A_{t-1}(j)h_{i,t}(j)$ and $1 + r = (1 + g)/\beta$, the above can be expressed more compactly as:

$$ V_{i,t-1}(j) = -\Psi_{t-1}(j) \left( z + \frac{n^0_{i,t-1}(j)}{\theta} \right) + \text{Prob}(h_{i,t}(j) = h_{t}(j)) E(h_{t}[j]) \frac{A_{t-1}(j)(1 - \rho)\beta(1 - \alpha)y_{t-1}}{(1 + g)^{t-\alpha} - \rho\beta(1 - \delta)E(h_{t}[j])} \left( \frac{\alpha}{w_{t-1}} \right)^{t-\alpha}, $$

where $z$ is the fixed cost of introducing an improved product, $n_i$ is firm-$i$’s level of research, $h_i$ is the realization of firm-$i$’s random draw, $h_{t}[j]$ is the best draw, $A_{t-1}[j]$ is the best quality in period $t - 1$, $E(h_{t}[j])$ is the expected value of the best draw in future periods (constant in equilibrium), $\delta$ is the probability of product-market destruction, and $\Psi_{t-1} = y_{t-1} A_{t-1}[j] E(h_{t}[j])$. Given a firm-$i$’s decision to introduce an improved product in market $j$, it will choose its level of research $n_i$ to maximize expected discounted profits, resulting in the following optimal research condition:

$$ \frac{\partial}{\partial n_{i,t-1}(j)} V_{i,t-1}(j) = 0, $$

where $\text{Prob}(h_i = h_{t}[j] E(h_{t}[j])$ depends on firm-$i$’s level of research $n_i$, every other firm’s level of research $n_{-i}$, and the number of products being introduced $e$.

An innovator choosing to introduce a new product, thus creating a new product market, pays no royalty and faces no uncertainty, but must incur a fixed cost of $z_0 \cdot y$. The value of introducing a new product is therefore:

$$ V_{0,t-1} = -y_{t-1} \cdot z_0 + \frac{(1 - \alpha)\alpha^{t-\alpha}}{1 + r} \sum_{t'=t}^{\infty} E(A_{t'[1]}(j) \frac{(1 - \delta)^{t'-t} \rho^{t'-t} y_{t'}}{(1 + r)^{t'-t} w_{t'-1}^{t'-\alpha}}, $$

or

$$ V_{0,t-1} = -y_{t-1} \cdot z_0 + E(h_{t}[j]) \frac{A_{t-1}(j)\beta(1 - \alpha)y_{t-1}}{(1 + g)^{t-\alpha} - \rho\beta(1 - \delta)E(h_{t}[j])} \left( \frac{\alpha}{w_{t-1}} \right)^{t-\alpha}. $$

Finally, note that firms will never choose to introduce more than one improved product into an existing market in the same period. Introducing a second product would reduce the expected value of introducing the first, so that a firm will always prefer to introduce a second product into a differentiated market. Further, if new-product creators were allowed a research decision (assumed away here), they
would similarly prefer to introduce any second product into a differentiated product. This all implies that the number of firms in an industry is indeterminate, although the number of firms in each product market is pinned down in equilibrium.9

Stationary Competitive Equilibrium

Here I take for granted that all sequential innovators face the same problem and all new-product creators also face the same problem. A stationary competitive equilibrium is a constant number of sequential innovators per market $e$, level of research per innovation $n$, measure of product markets $J$, and markup over marginal cost $P/w$, as well as a constant growth rate $g$ of intermediate prices $P$, final-good output $y$, and the wage rate $w$, such that:

(i) Consumer Optimization: $r = 1 + \frac{1 + g}{\beta} - 1$

(ii) Final-Good Firm Optimization: $P(j) = \left(\frac{\gamma_A(j)}{x(j)}\right)^{1-\alpha}, \forall j \in J$

(iii) Intermediate Producer Optimization: $P \frac{w}{\sigma_t(j)} V_i(j) = 0$

(iv) Sequential Innovator Optimization: $\frac{\partial}{\partial n_i} V_i(j) = 0$

(v) Free Entry: $V_i(j) = 0$ and $V_0 = 0$

(vi) Market Clearing (Goods): $y = \left(\int_0^J A^{1-\alpha}(j) \cdot x^\alpha(j) dj\right)^{\frac{1}{\alpha}}$

(vii) Market Clearing (Labor): $1 = \int_0^J x(j) dj$

To solve for the stationary equilibrium I start with the market clearing conditions and substitute each producer’s optimal output $x(j)$ to get:

$$y = yJ \frac{1}{\alpha} E(A)^\frac{1}{1-\alpha}, \text{ and } 1 = yJE(A) \left(\alpha \frac{\alpha}{w}\right)^{1-\alpha},$$

which can be solved for final-good output and the wage, each as functions of the average quality of producers $E(A)$ and the measure of product markets $J$:

$$y = J \frac{1}{1-\alpha} E(A)^{\frac{1}{1-\alpha}}, \text{ and } w = \alpha y.$$

As the measure of product markets is constant, the growth rate $g$ can be derived by using $A_{l+1}(j) = A_{l-1}(j) \cdot h_{l+1}$ to get:

$$\frac{y}{y_{l-1}} = 1 + g = E(h_{l+1})^{\frac{1}{1-\alpha}}.$$

The equilibrium number of sequential innovators per market $e$, level of research per innovation $n$, and measure of product markets $J$ can now be characterized using the free entry and optimal research conditions above:

free entry (new markets): $z_0 = \frac{\beta(1-\alpha)}{J[1-\rho\beta(1-\delta)]}$

free entry (existing markets): $z + \frac{n^\theta}{\theta} = \frac{(1-\rho)\beta(1-\alpha)}{eJ[1-\rho\beta(1-\delta)]}$

optimal research: $n^\theta = \frac{(1-\rho)\beta(1-\alpha)}{E(h_{l+1})J[1-\rho\beta(1-\delta)]} \cdot \frac{\partial}{\partial n_i} \text{Prob}(h_i = h_{l+1})E(h_{l+1}|\cdot) \bigg|_{n_i=n_{l+1}=n}$

9The model could presumably be extended to pin down the number of firms and number of products per firm in an industry as in Bernard, Redding, and Schott (2010), but such an extension is beyond the scope of the present chapter.
Given some distribution for $h_i$, the random variable determining each sequential innovator’s quality, all other variables are functions of $e$, $n$, and $J$.

### 2.2.5 Results

Here I examine how the stationary equilibrium of the economy depends on the strength of patent protection and the level of product-market regulation. I begin by considering equilibria associated with different levels of patent protection, measured here by the fraction $\rho$ of profits which infringers must pay in royalty fees. Consider first the free entry condition for new-product creators, above. It is immediately obvious that the denominator on the right-hand side must remain constant across different values of $\rho$. This implies that the measure of product markets $J$ must be increasing in patent protection. New-product creators never infringe on previous patents, so any increase in protection simply increases the present value of future royalty revenue, holding $J$ constant. But with free entry, the measure of firms creating new product markets each period increases until the increase in revenue is completely dissipated.

Now consider the free entry and optimal research conditions for sequential innovators. Given that $J[1 - \rho \beta (1 - \delta)]$ is constant, the net effect of $\rho$ on sequential innovators is equivalent to that of a simple tax on profits, reducing both the number of innovations $e$ and the level of research per innovation $n$.

Since the growth rate of the economy is increasing in both $e$ and $n$, it must be the case that growth is decreasing in patent protection.\footnote{From above, the growth rate is an increasing function of $E(h_{i1})$. That the best of $e$ draws is increasing in $e$ (conditional on $n$) should be obvious. In Appendix B.1 I show $E(h_{i1})$ must be increasing in the level of research $n$ (conditional on $e$).} Note that this is not simply a result of the increase in the measure of product markets $J$, which itself lowers the profits of innovators. Consider for a moment increasing $\rho$ while holding $J$ fixed. It is still the case that the profits of sequential innovators are decreasing in $\rho$, which continues to act as a tax on both the number of innovations and the level of research per innovation. The increase in $J$ simply magnifies the negative impact on growth. That growth is decreasing in patent protection is broadly consistent with the empirical evidence discussed in the introduction, and in Section 2.4 I provide some additional evidence in support of this conclusion. To my knowledge there has been no test of the hypothesis that the average value of innovations decreases with stronger patent protection. In Section 2.4 I provide evidence in support of this hypothesis.

Before discussing the effect of patent protection on welfare, it is useful to manipulate the three equilibrium conditions to obtain the investment rate of the economy $I$, the share of output devoted to sequential innovation in each existing market $e \left( z + \frac{n^\theta}{\theta} \right)$, and the relationship between the number of innovations $e$ and the level of research $n$;

\[
I = J \left[ \delta z_0 + (1 - \delta) e \left( z + \frac{n^\theta}{\theta} \right) \right] = \frac{\beta (1 - \alpha) [1 - \rho (1 - \delta)]}{1 - \rho \beta (1 - \delta)}
\]

\[
e \left( z + \frac{n^\theta}{\theta} \right) = (1 - \rho) z_0
\]

\[
n^\theta = \frac{\theta e n z \zeta (n_i, n - i, e)}{[\theta E (h_{i1}) - e n \zeta (n_i, n - i, e)]}
\]

where $\zeta (n_i, n - i, e) \equiv \frac{\partial}{\partial (n_i)} \text{Prob}(h_i = h_{i1}) E(h_{i1} | \cdot) \bigg|_{n_i = n - i = n}$. The first two equations show both total investment and investment in sequential innovation are decreasing in the strength of patent protection $\rho$. The last equation shows that while patent protection decreases investment in sequential innovation,
Chapter 2. Patent Protection as a Tax on Competition and Innovation

it does nothing to change the relationship between \( e \) and \( n \).

In Appendix B.3 I show a constrained social planner would choose an investment rate equal to \( 1 - \alpha \).

Equilibrium investment is therefore always lower than optimal, even given the inefficient relationship between the number of sequential innovations per existing market and the level of research per innovation. This all implies that in equilibrium it will never be the case that both the growth rate and the variety of products (which increases the level of productivity) are too high. Patent protection can therefore improve welfare only if the share of investment otherwise allocated to sequential innovation is so high that a relatively small increase in the resources allocated to creating new product markets increases welfare enough to offset a large decrease in resources devoted to sequential innovation. Note this is more likely in an economy with a low discount factor \( \beta \), since the optimal growth rate is increasing in \( \beta \) while the equilibrium growth rate is independent of \( \beta \).

To examine the effect of product-market regulation in the model, I interpret regulation as the imposition of higher costs on firms attempting to compete with an incumbent in an existing market. Consider the following free entry and optimal research conditions in an existing market where entrants must pay some multiple \( \tau > 1 \) of the normal costs of innovation, due to regulation:

- **Free entry (existing markets):**
  \[
  \tau \left( z + \frac{n^\theta}{\theta} \right) = \frac{(1 - \rho)\beta(1 - \alpha)}{eJ[1 - \rho \beta(1 - \delta)]}
  \]

- **Optimal research:**
  \[
  \tau_n^{\rho-1} = \frac{(1 - \rho)\beta(1 - \alpha)}{E(h_{[1]} J[1 - \rho \beta(1 - \delta)] \cdot \zeta(n_i, n_{-i}, e)}
  \]

where \( \zeta(n_i, n_{-i}, e) \equiv \frac{\partial}{\partial n} \text{Prob}(h_i = h_{[1]} E(h_{[1]} | \cdot)^n_i = n_{-i} = n) \). Clearly the effects of an increase in the burden of product-market regulation \( \tau \) on sequential innovation are identical to those of an increase in patent protection. An increase in \( \tau \) leads to fewer firms competing in each market, less research per firm, and lower growth. The only difference between increasing \( \tau \) and increasing \( \rho \) is that the greater costs incurred by entrants as a result of regulation are not transferred to previous patent holders, which implies the measure of product markets \( J \) is independent of \( \tau \). That the level of competition (proxied here by the number of competitors) and productivity growth both increase with product-market deregulation has become well established in recent years through both case studies and more systematic empirical studies, as discussed in the introduction.

To examine the interaction between product-market regulation and patent protection, again consider the free entry condition for existing markets. Note that \( e \left( z + \frac{n^\theta}{\theta} \right) \) is unit elastic with respect to \( \frac{(1 - \rho)}{\tau} \), so that any increase in the effective tax on profits will cause a corresponding decrease in \( e \left( z + \frac{n^\theta}{\theta} \right) \). Given that both the number of innovations \( e \) and the level of research per innovation \( n \) decrease with an increase in \( \rho \) or \( \tau \), it must be the case that \( e \cdot z \) falls slower than overall spending while \( e \cdot \frac{n^\theta}{\tau} \) (research intensity in an existing market) falls faster than overall spending. This implies that as the effective tax on profits increases, research intensity becomes a smaller and smaller portion of overall spending on innovation. This, in turn, implies that the elasticity of research intensity with respect to the tax is getting larger as the tax increases. All else equal, then, research intensity in an economy with strong patent protection (high \( \rho \)) will be a smaller portion of total spending on sequential innovation than in an economy with weak protection. This means product-market deregulation (a drop in \( \tau \)) should have a larger effect on research intensity in economies with strong patent protection. This result is consistent.
Chapter 2. Patent Protection as a Tax on Competition and Innovation

with the findings reported in Aghion, Howitt, and Prantl (2013), though the authors interpret their results in a very different way. While Aghion, Howitt, and Prantl argue their results are evidence of a previously undiscovered benefit of patent protection, the present model suggests that countries with high patent protection are simply deregulating from a much higher effective tax on innovation and thus enjoying bigger gains from the same change in policy.

2.3 Quantitative Analysis

In this section I calibrate the model to match several moments from U.S. manufacturing data. A quantitative analysis is useful for two reasons. First, to assess the quantitative effect of patent protection on growth in a model economy that behaves similarly to a real-world economy. Second, to determine whether welfare is increasing or decreasing with patent protection in the U.S.

An important parameter value for this analysis is the royalty rate $\rho$ associated with the strength of patent protection in the U.S. Over the last several decades, there has emerged a rule of thumb used by both license-negotiating firms and courts determining damages owed for infringement, such that a ‘reasonable royalty rate’ is about 25% of the operating profits of the infringer.\(^\text{12}\) For my calibration I assume the cost to sequential innovators of working around a patent is such that $\rho$ is equal to 25%.

For the purposes of this analysis, I assume sequential innovators take their quality draws from a one-parameter Kumaraswamy distribution, bounded by 1 and 1 + $n$, where $n$ continues to be chosen by the firm.\(^\text{13}\) This implies a firm-$i$’s random variable $h_i$ has the following cumulative distribution function:

$$F_i(h) = \text{prob}(h_i < h \mid n_i) = \left(\frac{h - 1}{n_i}\right)^\kappa,$$

where $\kappa$ is a shape parameter for the distribution. In Appendix B.2 I derive the expected value of the best of $e$ draws $E(h_{[1]})$, as a function of $e$ and the level of research per innovation $n$;

$$E(h_{[1]}) = 1 + \frac{\kappa en}{\kappa e + 1},$$

and show $\text{prob}(h_i = h_{[1]}) \cdot E(h_{[1]} \mid h_i = h_{[1]})$ can be expressed as;

$$\text{prob}(h_i = h_{[1]}) \cdot E(h_{[1]} \mid h_i = h_{[1]}) = \frac{n_i^{\kappa(e-1)}}{n_{-i}^{\kappa(e-1)}} \left(1 + \frac{\kappa(1 + n_i)}{e(\kappa e + 1)}\right).$$

With these distributions in hand, a quantitative analysis requires values for seven exogenous parameters; $\beta, \alpha, \theta, \delta, z, z_0$, and $\kappa$. Since $e$ is defined as the number of sequential innovations per existing product market, I can normalize the measure of product markets $J$ to 1. To obtain values for the seven required parameters, I therefore need to target six moments from U.S. manufacturing data while ensuring all three equilibrium conditions are satisfied. The first target I use is a real interest rate $r$ of 5%. Given some growth rate $g$, this will determine the value of $\beta = \frac{1 + g}{1 + r}$. In the model, the share of output retained by firms (and not paid to the factors of production) is $\alpha$. I set $\alpha$ to 0.85, a value commonly used.

---

\(^{12}\)See Goldscheider, Jarosz, and Mulhern (2002) for a description of this ‘25 Per Cent Rule’ and a brief history of its emergence.

\(^{13}\)The Kumaraswamy distribution is similar to the beta distribution, but with simple closed-form density and distribution functions. See Kumaraswamy (1980) and Nadarajah (2008) for details.
Table 2.1: Calibrated Parameter Values

<table>
<thead>
<tr>
<th>parameter</th>
<th>description</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>time preference</td>
<td>0.97</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>profit share of output</td>
<td>0.85</td>
</tr>
<tr>
<td>$\theta$</td>
<td>elast. of upper bound w.r.t. R&amp;D</td>
<td>3.42</td>
</tr>
<tr>
<td>$\delta$</td>
<td>prob. of market destruction</td>
<td>0.02</td>
</tr>
<tr>
<td>$z$</td>
<td>fixed cost of sequential innovation</td>
<td>0.02</td>
</tr>
<tr>
<td>$z_0$</td>
<td>fixed cost of new-product creation</td>
<td>0.19</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>distribution parameter</td>
<td>0.14</td>
</tr>
</tbody>
</table>

in the growth and development literature.\textsuperscript{14} The remaining four targets are a growth rate $g$ of 2.2%, an R&D intensity (research share of output) of 3.1%, a failure rate of newly introduced products of 80%, and an elasticity of expected output with respect to R&D of 0.08. The growth target is the average annual growth rate of total factor productivity for U.S. manufacturing from 1977 to 2007, taken from O’Mahony and Timmer (2009), and R&D data is for U.S. manufacturing from 1991 to 2008 from the National Science Foundation.\textsuperscript{15} The failure rate of new products is from Daft (2004). Hall, Mairesse, and Mohnen (2010) survey a number of empirical studies of the elasticity of output with respect to R&D and suggest 0.08 as a plausible value. Estimates vary quite a lot, however, so below I discuss the robustness of my results with respect to a range of targets for the elasticity of output with respect to R&D.

In the model the growth rate is $g = E(h_{[1]}) - 1$, given above. Since new-product creators have no research decision, I assume $z$ accounts for the fixed cost of introducing a product to the market (as in existing markets) and the residual $z_0 - z$ is the fixed cost of research for new-product creators. R&D intensity is therefore equal to $(1 - \delta) e \left( \frac{n^\theta}{\theta} \right) + \delta (z_0 - z)$, multiplied by the measure of product markets $J$. In each period one product is introduced in each new market and $e$ introduced in existing markets. Only one product survives in each market, so the overall failure rate is equal to $1 - \frac{1}{\delta + (1 - \delta) e}$. To derive the elasticity of expected output with respect to research expenditure, note that the equilibrium output of the best firm in a market is a linear function of the firm’s quality draw $h_{[1]}$. The elasticity is therefore:

$$\frac{\partial^2 E(h_{[1]})}{\partial n^\theta / \partial n} \cdot \frac{n}{(n^\theta / \theta)}.$$

The parameter values obtained through the calibration are reported in Table 2.1.

Using the above parameter values, I can now solve for the competitive equilibrium in an identical economy without patent protection ($\rho = 0$). To estimate the effect of patent protection on welfare I assume $u(\cdot) = \log(\cdot)$ in the consumer’s lifetime utility function. Table 2.2 reports the values of several variables of interest for the economy with and without patent protection. Removing patent protection from the benchmark economy reduces the measure of product markets by 21% and the consumption share of output by a sliver, but substantially increases the growth rate from 2.2% to 2.7%. The net result on welfare is significantly positive. The increase in steady-state utility from removing patent protection is equivalent to the increase in utility consumers would gain from a permanent increase in

\textsuperscript{14}For example, see Restuccia and Rogerson (2008).

\textsuperscript{15}Due to changes in sample design, comparable R&D data is unavailable before 1991.
Table 2.2: Quantitative Results

<table>
<thead>
<tr>
<th>variable</th>
<th>with patent protection</th>
<th>without patent protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>product markets</td>
<td>1</td>
<td>0.79</td>
</tr>
<tr>
<td>number of sequential innovations</td>
<td>5.6</td>
<td>7.3</td>
</tr>
<tr>
<td>level of research per innovation</td>
<td>0.45</td>
<td>0.48</td>
</tr>
<tr>
<td>growth rate</td>
<td>2.2%</td>
<td>2.7%</td>
</tr>
<tr>
<td>research intensity</td>
<td>3.1%</td>
<td>3.0%</td>
</tr>
<tr>
<td>investment rate</td>
<td>14.5%</td>
<td>14.6%</td>
</tr>
<tr>
<td>number of patents</td>
<td>5.0</td>
<td>5.1</td>
</tr>
<tr>
<td>utility</td>
<td>23.9</td>
<td>29.5</td>
</tr>
</tbody>
</table>

consumption of 16%.

That the optimal growth rate in an economy depends on the consumer’s discount factor while the equilibrium growth rate does not, implies the welfare predictions of the model are sensitive to the interest rate target, which determines $\beta$. A low interest rate implies a high $\beta$, which in turn implies a high optimal growth rate. With a high optimal growth rate the model is more likely to predict large welfare gains from reduced patent protection. Calibrating the model with an interest rate of 10% (rather than 5%) confirms this, generating an increase in welfare from the elimination of patent protection equivalent to that from an increase in annual consumption of only 2.7%.\textsuperscript{16}

For robustness, I also calibrate the model using a wide range of target values for the elasticity of output with respect to R&D. If a target of 0.01 is used, eliminating patent rights causes a larger drop in research intensity while barely changing the results with respect to the change in growth or the measure of product markets. As a result, the estimated increase in welfare is higher. If a target of 0.2 is used, eliminating patent rights causes an increase in research intensity and a slightly lower increase in growth. Welfare still increases, but by slightly less than in the benchmark calibration.

### 2.4 Patents, Citations, and Patent Protection

The model developed in Section 2.2 generates two unambiguous and testable implications. First, an increase in the strength of patent protection should lead to a decrease in productivity growth. Second, greater protection should lead to a decrease in the average quality of innovations. In this section I provide evidence in support of both of these implications. To proxy productivity growth due to innovation I consider a citation-weighted patent count, as is standard in the literature. To proxy the average quality of innovations I consider the average number of citations per patent. Although the model developed in Section 2.2 is ambiguous with respect to the effect of patent protection on the (unweighted) number of patents, both the quantitative analysis in Section 2.3 and Lerner (2009) suggest the number of patents should drop with an increase in patent protection. I find this is indeed the case. My general strategy here is to examine the relationship between a country’s patenting activity in the U.S. and its level of patent protection at home.

\textsuperscript{16}The welfare effect of eliminating patent protection turns negative if the model is calibrated to match an interest rate higher than 15%.
2.4.1 Data

The explanatory variable of interest in this analysis is the strength of patent protection across countries and time. As a measure of patent protection I use the Ginarte-Park Patent Rights Index developed in Ginarte and Park (1997) and updated in Park (2008). Each country-year in the Index is given a value from zero to five, with the index value increasing in coverage (the number of product-types protected), duration of protection, the scope of patent rights, and membership in international intellectual property treaties, while decreasing in the number of restrictions on patent rights (like compulsory licensing). The Index reports values for each country every five years from 1960 to 2005.

All patent and citation data is from the NBER Patent Database, which contains observations for each patent applied for in the U.S. from 1976 to 2006. Each observation contains information about the origin country (or countries) of each patent, the application year, the number of citations received by the patent in subsequent years, and the unique 3-digit SIC code associated with the patent (I throw out observations missing any of this information). The analyses here will use country-industry-time observations. Given the five-year intervals in the Patent Rights Index, I pool observations across the five years between the years used in the Index (for example, 1980 to 1984). The index value I use for each interval in the analysis is an average of the index values bounding the interval (1980 and 1985). For each observation of each independent variable of interest $y_i$ (where $y$ could be patent counts, citation-weighted patent counts, or citations per patent) I also construct a control variable $y_{not-US}$ using all other patents not used to construct $y_i$. I do not include the U.S. in my analysis because the level of patent protection in the U.S. presumably affects the behavior of foreign firms patenting in the U.S., which means an appropriate control variable $y_{not-US}$ cannot be constructed. Finally, I drop all patents with multiple countries of origin.

Once each of the above adjustments have been made, I am left with an unbalanced panel of country-industry observations made up of 87 countries and 452 industries over 6 time periods. Table 2.3 provides descriptive statistics for the sample. The average index value across country-'years' is 3.0, with a great deal of variation across countries and over time. The average number of patents across country-industry-years is 5.8, with a distribution highly skewed to the left. I treat observations with zero patents

---

Table 2.3: Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>s.d.</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patent Index</td>
<td>3.0</td>
<td>0.90</td>
<td>0.7</td>
<td>4.67</td>
</tr>
<tr>
<td>number of unweighted patents</td>
<td>5.8</td>
<td>13.8</td>
<td>0</td>
<td>530</td>
</tr>
<tr>
<td>number of citation-weighted patents</td>
<td>38</td>
<td>89</td>
<td>0</td>
<td>2707</td>
</tr>
<tr>
<td>citations per patent</td>
<td>5.2</td>
<td>7.3</td>
<td>0</td>
<td>324</td>
</tr>
<tr>
<td>time periods per country</td>
<td>4.9</td>
<td>1.6</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>number of countries</td>
<td>87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of industries</td>
<td>452</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

17 See Hall, Jaffe, and Trajtenberg (2005) for further details.
18 For the first time interval, I use 1976 to 1979 and simply scale up the number of citations and patents by 5/4.
19 An alternative strategy is to pool observations across a five-year interval bounding each index year (for example, 1978 to 1982 for index year 1980). But due to the start date of the patent data, this would involve throwing out more observations and reducing the number of time periods.
20 For the remainder of this section I use ‘year’ to refer to a 5-year time period.
in the following way. Until a country-industry reports a positive number of patents, I assume no activity in that industry for that country, and treat the number of patents (unweighted or otherwise) as missing. Once a country-industry reports a positive number of patents, I treat any future zero observations for that country-industry as zeroes. Interpreted in this way, almost one quarter of all country-industry-years report zero patents.

### 2.4.2 Econometric Models

As is standard in the patent-count literature, I assume both patents and citation-weighted patents follow a Poisson process with the following hazard rate:

\[ \text{exp}\{\beta_0 + \beta_1 \text{index} + \beta_2 \text{controls}\}, \]

where controls for each observation \( y_{c,i,t} \) include the control variable described above, as well as a complete set of dummies for country, industry, and time.

Endogeneity is a potential problem for two reasons. First, the political decision to strengthen patent rights could very well be influenced by the level of innovation and growth (in particular, firms may lobby for protection when expenditure on innovation is high). Second, the Ginarte and Park Patent Rights Index allocates a higher index value to countries ratifying international property rights treaties. Ratification of these treaties is often accompanied by, and even a condition for, entry into the World Trade Organization or other trade agreements which could themselves increase the incentive for firms to innovate. Not accounting for these problems should lead to upwardly-biased estimates of the effects of patent protection. To address this potential endogeneity I instrument the index value for each country-year with the previous year’s value, using a control function approach as in Aghion et al. (2005).

For my test of the effect of patent protection on citations per patent, I use a weighted least squares estimator with the same controls and with observations weighted by the number of patents used to construct the observation. Faced with the same endogeneity issues, I use the same instrument described above with a two stage weighted least squares estimator.

### 2.4.3 Results

Table 2.4 reports the results of four regressions for the number of unweighted and citation-weighted patents, with and without controlling for endogeneity. Without controlling for the endogeneity of patent protection, both unweighted and citation-weighted patents are estimated to increase with an increase in patent protection. Once endogeneity is controlled for, however, patents are estimated to decrease when patent protection increases. These changes in the signs of the estimated effects, combined with the estimated coefficients for the control function in columns 2 and 4, suggest endogeneity indeed plays a role. The estimates in columns 2 and 4 suggest an increase in patent protection by one standard deviation from an average value of 3 is associated with a drop in the number of unweighted patents from an average of 5.8 to an average of 4.7, and a drop in citation-weighted patents from 38 to 25. The effect of patent protection on the number of patents is both statistically significant and large in magnitude.

Table 2.5 reports the results of three regressions for the number of citations per patent. Controlling for the endogeneity of patent protection significantly increases its estimated impact on citations per patent. The estimates in column 3 suggest an increase in patent protection by one standard deviation
Table 2.4: Results

<table>
<thead>
<tr>
<th></th>
<th>unweighted patents</th>
<th>citation-weighted patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>instrument used:</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Patent Index</td>
<td>0.042*** (0.01)</td>
<td>-0.23*** (0.02)</td>
</tr>
<tr>
<td>control variable</td>
<td>0.001*** (0.00)</td>
<td>0.001*** (0.00)</td>
</tr>
<tr>
<td>significance of Index</td>
<td>112.41 (0.00)</td>
<td>1514.57 (0.00)</td>
</tr>
<tr>
<td>significance of instrument</td>
<td>21996 (0.00)</td>
<td>21962 (0.00)</td>
</tr>
<tr>
<td>control function</td>
<td>0.335 (0.03)</td>
<td>0.628 (0.01)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>$R^2$ of reduced form</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>observations</td>
<td>41548</td>
<td>40227</td>
</tr>
</tbody>
</table>

All regressions include dummies for country, industry, and year. Dummy and constant coefficients omitted. Estimates from a Poisson regression, with standard errors in parantheses. *** denotes significance at 1% level, no stars denotes lack of significance at 10% level. Standard errors in columns 2 and 4 have not been corrected for inclusion of control function. Significance tests show likelihood ratio test-statistics with $P$-values in parantheses.

Table 2.5: Results

<table>
<thead>
<tr>
<th></th>
<th>no</th>
<th>yes</th>
<th>yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>instrument used:</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>weights used:</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Patent Index</td>
<td>-0.42*** (0.15)</td>
<td>-1.18*** (0.31)</td>
<td>-1.19*** (0.22)</td>
</tr>
<tr>
<td>control variable</td>
<td>0.62*** (0.03)</td>
<td>0.58*** (0.04)</td>
<td>0.62*** (0.03)</td>
</tr>
<tr>
<td>$R^2$ (adjusted)</td>
<td>0.54</td>
<td>0.30</td>
<td>0.55</td>
</tr>
<tr>
<td>observations</td>
<td>32111</td>
<td>31187</td>
<td>31187</td>
</tr>
</tbody>
</table>

Dependent variable is citations per patent. All regressions include dummies for country, industry, and year. Dummy and constant coefficients omitted. All standard errors clustered at country-industry level. *** denotes significance at 1% level.
from an average value of 3 is associated with a drop in the number of citations per patent from an average of 5.2 to an average of 4.1. This effect is also both statistically significant and large in magnitude.

2.5 Conclusion

In this chapter I pull together ideas from the theoretical literatures on competition, innovation, growth, and patent rights to develop a comprehensive yet tractable model consistent with key facts documented in recent empirical studies. Given the tendency for cross-country intellectual property treaties to be tied to trade agreements and market reforms, a comprehensive framework of the kind developed here is necessary to disentangle the often contradictory effects of the resulting changes in policy. In the model, as in the data, an increase in competition driven by product-market deregulation is associated with more innovation and higher growth. An increase in the strength of patent rights results in less sequential innovation and lower growth while expanding the equilibrium variety of products. While the model is ambiguous with respect to the effect of patent rights on welfare, calibrating the model generates the prediction that eliminating patent rights should result in a substantial increase in welfare. The model generates an additional testable implication, that stronger patent protection should lead to a lower average quality of innovations. I offer evidence that this is indeed the case.

The theoretical and empirical results reported here support Boldrin and Levine’s (2008b) recent claim that evidence of a positive link between competition and innovation is itself evidence against the need for patent protection to encourage innovation. One must wonder at the foresight of a commenter who reached similar conclusions two centuries ago;

In some other countries [the granting of patent rights] is sometimes done, in a great case, and by a special and personal act, but, generally speaking, other nations have thought that these monopolies produce more embarrassament than advantage to society; and it may be observed that the nations which refuse monopolies of invention, are as fruitful as England in new and useful devices.

-Thomas Jefferson, letter to Isaac McPherson, 13 August 1813
Chapter 3

Niche Firms, Mass Markets, and Income Across Countries: Accounting for the Impact of Entry Costs

Abstract:
I develop a model of monopolistic competition in which I distinguish between niche markets and mass markets, in the spirit of Holmes and Stevens (2010). Firms choose between entering a small niche market with high markups or a large mass market with low markups. Entry costs have a much greater impact on output in the niche market as the gains to specialization are high, relative to the mass market where varieties are highly substitutable. Calibrated to match data from U.S. manufacturing, the model generates an elasticity of total factor productivity with respect to entry costs more than twice that in a model that abstracts from heterogeneous markets. I use data on entry costs across countries to show entry costs alone can account for 45% of the cross-country variation in productivity and income per worker. In comparison, empirical estimates of the explanatory power of entry costs are about 50%.
3.1 Introduction

Income differences across countries are enormous. GDP per capita in the poorest ten percent of countries averaged less than two percent of the U.S. level in 1996. Economists as far back as Adam Smith have considered barriers to entry for new firms to be a likely contributor to these differences across nations. But the last decade has seen a surge in studies of the effects of these barriers on aggregate outcomes, fueled in no small part by the efforts of Djankov et al. (2002) and the World Bank to measure, catalogue, and report standardized measures of the regulatory costs of doing business across countries. The story that has begun to emerge is one in which the (substantial) variation in entry costs across countries explain as much as half of the differences in both income and TFP.

While the empirical evidence suggests a large impact of entry costs on aggregate outcomes, quantitative models have consistently predicted a much lower impact. A number of papers have been written recently investigating a wide range of possible mechanisms through which high entry costs might affect both within-firm and aggregate TFP. These include reducing the number of firms and thereby inefficiently inflating the size of firms, allowing less productive firms to survive, allowing higher markups over marginal cost, and reducing incentives to adopt new technologies. Remarkably, this diverse group of theories share one common feature - a relatively low impact of entry costs on aggregate TFP and output.

To understand why entry costs have been consistently predicted to have a low impact, consider a simple constant-elasticity-of-substitution (CES) model of monopolistic competition (I will at times refer to this as the ‘standard’ model). An important parameter in the model is the elasticity of substitution between varieties. Estimating elasticities for a number of industries, Imbs and Méjean (2009) suggest a plausible elasticity between varieties of 6-7 in a model with constant elasticities across industries. In such a model, entry costs affect productivity by reducing the number of firms and thereby enlarging the average firm above its efficient size. The high substitutibility between varieties suggested by Imbs and Méjean implies small gains from specialization, so that more output from a smaller number of firms is almost as efficient as less output from a larger number of firms (in the limit, an infinite elasticity implies any number of firms is efficient). As a result, this model can only account for about 7% of the variation in TFP across countries. Although the mechanisms differ, this general result holds across all of the models developed in recent papers.

In this chapter, I develop a model of monopolistic competition in which I distinguish between niche markets and mass markets within the same industry. Incorporating these heterogeneous markets magnifies the impact of entry costs (and other distortions) significantly. In the spirit of Holmes and Stevens (2010), I allow entering firms in each industry to choose between producing a niche good or a mass-market good. Think here of McDonalds and Burger King versus a vegetarian burger chain, or a mass-market furniture manufacturer versus a manufacturer of Amish furniture. The market for niche goods is rela-
tively small and characterized by a low elasticity of substitution between varieties, as niche goods target very specific tastes. The demand for mass-market goods, on the other hand, is large and characterized by high substitutability between varieties. In essence, new firms choose between a small market with high markups, or a large market with small markups. Each market exhibits a different relationship between entry costs (via the number of firms) and aggregate productivity. In particular, the impact of entry costs on aggregate niche output is far greater than that on mass-market output. I can compare the model to the standard CES model by ensuring the average elasticity of substitution between all firms in an industry is the same in both models. I show that for any targeted elasticity, the impact of entry costs on aggregate TFP is higher when both niche firms and mass-market firms are allowed for.

Calibrating the model to match moments from U.S. manufacturing data, I perform a counterfactual experiment in which I increase entry costs and observe the resulting change in aggregate TFP and output per worker. The calibrated model generates an elasticity of TFP with respect to entry costs of 0.45, about two-and-a-half times that in the standard model. The corresponding elasticity for output per worker is 0.77. I perform a second experiment in which I generate cross-country TFP and output per worker by feeding cross-country data on entry costs from the World Bank into the model. I find the model explains 45% of the variation in TFP and output per worker, close to the estimate of one half reported in Herrendorf and Teixeira (2011). This implies the elasticities generated by the model are 95% of the actual elasticities in the data.

The idea to distinguish between niche and mass-market firms is from Holmes and Stevens (2010). They offer the idea as an alternative theory of the plant size distribution, complimenting the usual assumption that differences in size solely reflect differences in productivity across plants. The focus of their paper is on the location decision of firms and the differential effects of import competition on niche firms and mass-market firms.

Herrendorf and Teixeira (2011) develop a model that allows for a wide array of distortions, calibrate their model to match U.S. data, and then use a wide variety of cross-country data to isolate the variation in TFP and output per worker accounted for by entry costs. They estimate about half of the variation in TFP and output per worker is due to variation in entry costs. Barseghyan (2008) uses a number of different instruments to estimate the impact of regulatory entry costs and various proxies for institutional quality, and finds an increase in costs by one half of a standard deviation in his sample is associated with a 22% reduction in TFP.

This chapter’s focus on the aggregate impact of entry costs is related to the large recent literature analyzing the effects of various distortions in the prices faced by firms on the allocation of resources and aggregate outcomes. Although the focus of this chapter is on entry costs, I briefly explain how the model magnifies the impact of distortions more generally.

In the next section I present the model and solve for its steady-state equilibrium. In Section 3.3 I calibrate the model and show how the benchmark economy behaves when entry costs are increased. In Section 3.4 I construct a measure of entry costs for 136 countries and use the model to generate TFP and output per worker for each country in the sample. I then compare the model-generated outcomes to the data. Section 3.5 follows with a discussion about how accounting for heterogeneous markets affects the impact of other distortions. The final section concludes.

---

6This assumption is consistent with Hsieh and Klenow (2009), who report a negative relationship between establishment size and markups in the U.S. (p. 1436-37).

7For example, Guner et al. (2008), Restuccia and Rogerson (2008), and Hsieh and Klenow (2009).
3.2 The Model

My focus here is on the aggregate impact of entry costs rather than the various mechanisms through which entry costs may affect firm dynamics or firm-level TFP, so I abstract from any heterogeneity in firm productivity or fixed costs of operation.\(^8\) Total output in the economy is a CES aggregate of intermediate inputs, and all intermediate firms produce one-for-one using labor.\(^9\) As such, output per worker in this economy is equivalent to aggregate TFP. I further assume one representative industry, which contains both a market for niche varieties and a market for mass-market varieties. Firms pay a one-time fixed cost of entry and choose whether to produce a niche variety or a mass-market variety. All firms face an exogenous probability of death in each period, so there is firm turnover in steady state. I study the stationary competitive equilibrium of the economy in which firms take the economy-wide wage rate as given and free entry ensures zero expected profits for all entrants. I begin by describing the environment.

3.2.1 Environment

There is a representative consumer who supplies one unit of labor to intermediate firms. The consumer only values consumption \((C)\) and has a constant discount factor \(\beta \in (0, 1)\). Preferences over the stream of consumption in each period are described by the following utility function;\(^{10}\)

\[
\sum_{t=0}^{\infty} \beta^t u(C_t).
\]

The market for final output is perfectly competitive, with a representative firm using inputs from a representative intermediate industry to produce output according to the following production function;\(^{11}\)

\[
y = X_n^\phi X_m^{1-\phi},
\]

where \(X_n\) is a CES aggregate of niche output, \(X_m\) a CES aggregate of mass-market output, and \(\phi\) the weight on niche output. Under the above assumptions, aggregate output \(y\) is equivalent to both output per worker and aggregate TFP.

Niche output is aggregated in the following way;

\[
X_n \equiv \left( \int_0^{N_n} x_n^{\alpha_n} dn \right)^{\frac{1}{1-\alpha_n}},
\]

where \(N_n\) is the measure of niche firms/varieties, \(x_n\) the demand for the output of niche firm \(n\), and \(\frac{1}{1-\alpha_n}\) the constant elasticity of substitution between niche varieties. Mass-market output is aggregated in the same way, but with an elasticity of substitution between mass-market varieties of \(\frac{1}{1-\alpha_m} > \frac{1}{1-\alpha_n}\). Note that if the elasticity of substitution is made the same across all varieties, this model collapses to a standard model of monopolistic competition.

Intermediate firms produce output one-for-one with labor, so \(x_i\) also denotes labor demanded by firm

---

\(^8\)At the end of this section, I discuss the implications of allowing for both heterogeneous firms and fixed operating costs.

\(^9\)Henceforth I use ‘firm’ to mean an intermediate-good firm, unless otherwise indicated.

\(^{10}\)I focus my attention on the stationary equilibrium of the economy, so there is no need to specify the form of \(u(\cdot)\).

\(^{11}\)Throughout the paper, I omit the time subscript unless clarity requires it.
i. In each period a new firm can choose to enter as either a niche firm or a mass-market firm, but must pay a fixed entry cost equal to a multiple \( c_e \) of output per worker \( y \). Entry costs are financed by issuing equity to consumers and I assume an unlimited mass of potential entrants. Finally, all firms face the same exogenous probability of death \( \lambda \) each period.

### 3.2.2 Competitive Equilibrium

I focus on the stationary competitive equilibrium of the model. In such an equilibrium, the real wage \( w \) and real interest rate \( r \) will be constant, as will firm entry and exit. I begin by describing the decision problems of each agent, and then define and solve for the stationary equilibrium.

#### Consumer

In each period the consumer chooses both consumption and savings, and the only vehicle for savings is the purchase of equity in intermediate firms, earning a rate of return of \( r \). The consumer’s problem is therefore to choose consumption \( C \) and savings \( S \) in each period \( t' \), given \( w \) and \( r \), to maximize:

\[
\sum_{t'=0}^{\infty} \beta^{t'} u(C_t), \quad \text{s.t. } C_t + S_t \leq w + S_{t-1}(1 + r).
\]

The first order conditions for this problem imply a constant real interest rate:

\[ r = \frac{1}{\beta} - 1. \]

#### Final-Good Producer

In each period, the final-good producer takes the prices of all intermediate inputs as given, and demands inputs from each intermediate firm to maximize profits:

\[
X_n^{\phi} X_m^{1-\phi} - \int_0^{N_n} P_n x_n dn - \int_0^{N_m} P_m x_m dm,
\]

where \( X_i = \left( \int_0^{N_i} x_i^{\alpha_i} di \right)^{1/\alpha_i}, \ i \in \{n,m\} \), \( P_n \) is the price of niche input \( n \), and \( P_m \) the price of mass-market input \( m \). The first order conditions for the final-good firm’s problem imply the following inverted demand functions for intermediate inputs:

\[ P_n = \frac{\phi g x_n^{\alpha_n - 1}}{X_n^{\alpha_n}}, \quad \text{and} \quad P_m = \frac{(1 - \phi) g x_m^{\alpha_m - 1}}{X_m^{\alpha_m}}, \]

for each niche input \( n \) and mass-market input \( m \).

#### Intermediate Firms

In each period, incumbent niche firms face the downward-sloping demand curves implied by the final-good firm’s problem above, and demand labor \( x_n \) given the measure, prices, and output of other firms, to maximize profits:

\[ \pi_n = P_n x_n - w x_n. \]
First order conditions imply an optimal price equal to:

\[ P_n = \frac{w}{\alpha_n}, \]

and optimal output equal to:

\[ x_n = \left( \frac{\alpha_n \phi y}{wX_n} \right)^{\frac{1}{1-\alpha_n}}. \]

Analogous to a niche firm’s choices, the first order conditions for mass-market firm \( m \) imply an optimal price of:

\[ P_m = \frac{w}{\alpha_m}, \]

and optimal output of:

\[ x_m = \left( \frac{\alpha_m (1 - \phi) y}{wX_m} \right)^{\frac{1}{1-\alpha_m}}. \]

In equilibrium, total expected discounted profits for incumbent niche and mass-market firms are:

\[ \frac{\pi_n}{1 - \rho} \quad \text{and} \quad \frac{\pi_m}{1 - \rho}, \]

where \( \rho = \frac{1 - \lambda}{1 + r} \) is each firm’s discount rate, \( \lambda \) is the exogenous probability of firm death, and \( r \) is the real interest rate. Free entry ensures that new firms will enter as niche firms as long as:

\[ \frac{\pi_n}{1 - \rho} \geq c_e y, \]

and new firms will enter as mass-market firms as long as:

\[ \frac{\pi_m}{1 - \rho} \geq c_e y. \]

**Steady-State Equilibrium**

Here I take into account that all niche firms face the same problem and all mass-market firms also face the same problem. A stationary competitive equilibrium is a real wage \( w \), a real interest rate \( r \), a niche price \( P_n \), a mass-market price \( P_m \), a measure of niche firms \( N_n \), a measure of mass-market firms \( N_m \), niche firm output \( x_n \), mass-market firm output \( x_m \), and final-good output \( y \) such that:

(i) Consumer Optimization: \( r = \frac{1}{z} - 1 \)

(ii) Final-Good Firm Optimization: \( P_n = \frac{\phi y}{N_n x_n} \quad \text{and} \quad P_m = \frac{(1 - \phi) y}{N_m x_m} \)

(iii) Intermediate Firm Optimization: \( P_n = \frac{w}{\alpha_n} \quad \text{and} \quad P_m = \frac{w}{\alpha_m} \)

(iv) Free Entry: \( c_e y = \frac{(1 - \alpha_n) w x_n}{\alpha_n (1 - \rho)} \quad \text{and} \quad c_e y = \frac{(1 - \alpha_m) w x_m}{\alpha_m (1 - \rho)} \)

(v) Market Clearing (Goods): \( y = \left( \frac{N_n^{1/\alpha_n} x_n}{x_n} \right)^{\phi} \left( \frac{N_m^{1/\alpha_m} x_m}{x_m} \right)^{1 - \phi} \)

(vi) Market Clearing (Labor): \( 1 = N_n x_n + N_m x_m \)

where \( \rho = \frac{1 - \lambda}{1 + r} \).

These equilibrium conditions can be used to solve for the steady-state measures of niche and mass-market firms:

\[ N_n = \frac{(1 - \alpha_n) \phi}{c_e (1 - \rho)} \quad \text{and} \quad N_m = \frac{(1 - \alpha_m)(1 - \phi)}{c_e (1 - \rho)}, \]
as well as aggregate TFP in this economy;

$$y = \Delta / c_e^{\phi \alpha_n + 1 - \phi \alpha_m - 1},$$

$$\Delta = \frac{[\alpha_n^\alpha (1 - \alpha_n)^{1-\alpha_n \phi}]^{\frac{\alpha_n}{\alpha_n}} [\alpha_m^\alpha (1 - \alpha_m)^{1-\alpha_m (1 - \phi)}]^{\frac{1-\phi}{\alpha_m}}}{[\alpha_n \phi + \alpha_m (1 - \phi)](1 - \rho)^{\frac{1-\phi}{\alpha_n} + \frac{1-\phi}{\alpha_m} - 1}}.$$

The elasticity of $y$ with respect to entry costs $c_e$ is equal to;

$$\frac{\phi}{\alpha_n} + \frac{1 - \phi}{\alpha_m} - 1.$$

This elasticity is increasing in $\phi$, and decreasing in both $\alpha_n$ and $\alpha_m$. As either of the $\alpha$’s increase, the gains from specialization decrease, lowering the impact of fewer varieties induced by higher entry costs. An increase in $\phi$ increases the relative importance of the niche market, where entry costs have a greater impact.

To see this elasticity is greater than that in the standard model, let $\frac{1}{1-\alpha}$ denote the homogeneous elasticity between all varieties in the standard model. The elasticity of $y$ with respect to entry costs in the standard model is therefore $\frac{1}{1-\alpha} - 1$. To discipline the comparison, I restrict $\alpha_n$ and $\alpha_m$ such that the average elasticity across firms is equal to $\frac{1}{1-\alpha}$. This implies $\phi \alpha_n + (1 - \phi) \alpha_m = \bar{\alpha}$. This is equivalent to forcing the share of aggregate output paid to labor to be equal to $\bar{\alpha}$ in both models. Given this restriction, entry costs have a greater impact on $y$ in the present model than in the standard model if;

$$\frac{\phi}{\alpha_n} + \frac{1 - \phi}{\alpha_m} > \frac{1}{\bar{\alpha}} = \frac{1}{\phi \alpha_n + (1 - \phi) \alpha_m},$$

or equivalently, if $\frac{\alpha_n}{\alpha_m} + \frac{\alpha_m}{\alpha_n} > 2$, which holds for any $\alpha_n \neq \alpha_m$.

### 3.2.3 Heterogeneous Firms

Here I discuss how a model of niche and mass markets can be extended to include heterogeneity in productivity and fixed operating costs. Consider first an environment identical to that described above, but where each firm draws its productivity from a distribution upon entry. Assuming all firms draw from the same distribution and realize their productivity after entry (as is standard in this class of models), the distribution of productivity across niche firms will be identical to that across mass-market firms. Heterogeneity in productivity implies that higher productivity draws will be associated with larger firms. Whereas the employment size of each mass-market firm in the benchmark model is larger than each niche firm by a factor of $\frac{x_m}{x_n} = \frac{(1 - \alpha_m)}{\alpha_m (1 - \alpha_n)}$, heterogeneity in productivity implies that for any given productivity draw, a mass-market firm will be larger (by the same factor) as a niche firm with the same productivity. Although this extension provides a useful framework to think about multiple drivers of the firm size distribution (which is interesting in and of itself), it does not change the aggregate impact of entry costs. The measure of firms in each market will continue to be (inversely) proportional to the cost of entry, and the impact of entry costs will continue to be determined by the elasticity of substitution in the two markets as well as the weight given to each market.

Now consider an identical environment, but where firms face a fixed cost in production as in Hopenhayn (1992). Assume this fixed cost is modeled in the same fashion as entry costs, as a fixed fraction of output per worker. Now after entering and realizing their productivity, firms must choose whether
to produce or to exit the industry. Firms with a realized productivity under some threshold will exit, so that only the most productive firms will produce. In this class of models higher entry costs reduce the total measure of firms, while higher operating costs reduce the productivity threshold at which the marginal entrant is indifferent between exiting and producing. It turns out that free entry together with the assumption that all firms face the same costs ensures that the productivity threshold is the same for niche and mass-market firms. As a result, the aggregate impact of these fixed costs remains the same. The only difference is that the total elasticity of TFP with respect to these fixed costs will be split between entry costs and operating costs.\textsuperscript{12} Accounting for niche and mass markets within an industry magnifies the aggregate impact of these costs (relative to a model without heterogeneous markets) by the same factor as that reported in Section 3.2.2.4.

### 3.3 Calibration

To quantify the impact of entry costs on TFP, I need values for three parameters: $\alpha_n$, which determines the elasticity of substitution between niche varieties; $\alpha_m$, which does the same for mass-market varieties; and $\phi$, which determines the total revenue share of niche firms. To obtain these parameter values, I calibrate the model to match three moments from U.S. manufacturing data. The first target is the CES suggested by Imbs and Méjean (2009) for models with homogeneous elasticities between varieties. They suggest a value between 6 and 7, so I use 6.5.\textsuperscript{13} Following the same procedure outlined in Section 3.2.2.4, I restrict $\alpha_n$ and $\alpha_m$ to satisfy:

$$\phi \alpha_n + (1 - \phi) \alpha_m = \frac{\overline{CES} - 1}{CES},$$

where $\overline{CES} = 6.5$ is the target.

The remaining targets are from Hsieh and Klenow (2009). They report moments from the within-industry distribution of total factor revenue productivity (TFPR) across U.S manufacturing establishments in 1996. Here I assume (as is standard) the U.S. is an undistorted economy, so that the authors’ measure of TFPR can be interpreted simply as a firm’s markup over marginal cost (multiplied by a constant). The two moments I target are the gap between the 90th and the 10th percentile of deviations of log(TFPR) from industry means, and the standard deviation of the same. Hsieh and Klenow report a 90/10 gap equal to 1.19 and a standard deviation equal to 0.49 for the U.S. in 1997. In the model, the markup of a niche firm is $1/\alpha_n$, while that of a mass-market firm is $1/\alpha_m$. For any firm $i$, the deviation of $\log(1/\alpha_i)$ from the average value of $\log(1/\alpha_i)$ across all firms is:

$$\log \left( \frac{1}{\alpha_i} \right) - \frac{N_n}{N_T} \log \left( \frac{1}{\alpha_n} \right) - \frac{N_m}{N_T} \log \left( \frac{1}{\alpha_m} \right),$$

where $N_T$ denotes the mass of all firms $N_n + N_m$. I assume (and then verify) that the 90th percentile firm in the model is a niche firm, while the 10th percentile firm is a mass-market firm, so matching the

\textsuperscript{12}This split is determined by the productivity distribution from which entrants draw.

\textsuperscript{13}This is consistent with Broda and Weinstein (2006). This value also implies payments to factors of production equal to 85% of aggregate output, which is a value commonly used as a target in the recent misallocation literature (for example, see Restuccia and Rogerson (2008)).
90/10 gap requires that the values of $\alpha_n$ and $\alpha_m$ satisfy:

$$\frac{\alpha_m}{\alpha_n} = e^{1.19} = 3.3.$$  

The standard deviation of the above statistic is:

$$\left( \frac{N_n}{N_T} \right)^{\frac{1}{2}} \left( \frac{N_m}{N_T} \right)^{\frac{1}{2}} \log \left( \frac{\alpha_m}{\alpha_n} \right),$$

Using the solutions for the measures of niche and mass-market firms from Section 3.2.2.4, the standard deviation can be expressed as a function of parameters. The parameter values must satisfy:

$$\frac{\phi(1 - \phi)(1 - \alpha_n)(1 - \alpha_m)}{\phi(1 - \alpha_n) + (1 - \phi)(1 - \alpha_m)} \log \left( \frac{\alpha_m}{\alpha_n} \right) = 0.49.$$  

With three equations in hand, I can now obtain the required parameter values. These are:

- $\phi : 0.17$, implying a revenue share for niche firms of 17%
- $\alpha_n : 0.29$, implying an elasticity of substitution across niche varieties of 1.4
- $\alpha_m : 0.96$, implying an elasticity of substitution across mass-market varieties of 25

A complication here is that the three targets give two sets of possible parameter values. The above values imply niche firms account for 78% of all firms, while the alternative set of values implies 22%. To choose between them, I note Holmes and Stevens (2010) estimate that niche firms account for 66% of all firms. Given the more narrow definition the authors use for ‘niche’, I choose the set of values above.\textsuperscript{14} The calibrated values also imply mass-market firms are much larger in terms of employment (by a factor of 58). For comparison, I note establishments in U.S. manufacturing industries in 1997 with less than 50 employees accounted for 83% of all establishments with payroll, and accounted for 12% of value-added.\textsuperscript{15}

The elasticity of TFP with respect to entry costs reported in Section 3.2.2.4 is:

$$\frac{\phi}{\alpha_n} + \frac{1 - \phi}{\alpha_m} - 1.$$  

Using the calibrated parameter values, I obtain an elasticity of 0.45. Given the distribution of entry costs across countries reported in the next section, the impact of entry costs is significant. An elasticity of 0.45 implies that an increase in entry costs of one half of the standard deviation of entry costs across countries above the mean will cause TFP to drop by 24%. To compare this elasticity to that in the standard model, I remind the reader that the elasticity of TFP with respect to entry costs in the standard model is $\frac{1}{\sigma} - 1$, where $\sigma = \frac{\text{CES} - 1}{\text{CES}}$. Using the same value for $\text{CES}$ targeted in the above calibration (6.5), I calculate an elasticity in the standard model of 0.18. Allowing for niche and mass-market firms increases the elasticity by a factor of 2.5.

Note that this calibration strategy is robust to the extensions discussed in Section 3.2.3, where I consider the implications of allowing for firm heterogeneity in productivity and fixed operating costs. The elasticities of substitution across firms in an industry, the markups chosen by each firm, and the

\textsuperscript{14}Holmes and Stevens (2010) define a niche firm as a producer that requires face-to-face interaction with customers, and so tends not to ship long distances. The definition I use here is much broader, as it allows for firms that cater to niche markets around the country.

\textsuperscript{15}Calculated using U.S. Census Bureau data.
ratio of niche firms to mass-market firms are all independent of the productivity distribution in each market. Using the same calibration strategy for the extended model would therefore produce the same parameter values.

### 3.4 Cross-Country Experiment

In this section I construct a measure of entry costs for 136 countries and use the model to generate values of TFP and output per worker for economies that differ only with respect to entry costs. I then compare the performance of the model to cross-country data. I begin by explaining how I measure entry costs.

#### 3.4.1 Constructing Entry Costs

I assume entry costs consist of two components: nonregulatory costs which are constant across countries (as a fraction of output per worker), and regulatory costs. The World Bank Doing Business Survey reports measures of the various regulatory costs of doing business for a broad sample of countries. For my measure of regulatory costs, I start by summing the monetary costs reported by the World Bank. These include the cost of starting a business (registering the business, getting the necessary licenses, etc...), the cost of construction permits, the cost of registering property, the cost of obtaining an electricity connection, and the wage cost of time spent doing taxes. The World Bank also reports the delays in opening a business associated with each of the above costs. I add to my measure of regulatory costs the discounted value of lost profits and wages to the business owner due to these delays.

To obtain a measure of nonregulatory costs, I start with the average monetary startup cost reported by Acs et al. (2005) for the U.S. of 86% of output per worker. From this number I subtract the monetary regulatory costs in the U.S. and find nonregulatory costs are 46% of output per worker. I assume this fraction is constant across countries.

Total entry costs are calculated by combining the nonregulatory and regulatory costs in each country. Table 3.1 reports some descriptive statistics of total entry costs and output per worker, each relative to the U.S. level.

---

16 In the appendix, I describe this process in more detail.

17 I assume owners pay themselves the same wage as any other employee. Without delays in opening the business this income could be ignored, as it is offset by the wage income the owner could earn elsewhere.
3.4.2 Experiment

I can now compare the model to the data. I construct measures of TFP for 116 countries for 1996 following Klenow and Rodríguez-Clare (2005), using data from Penn World Tables v7.0 and Barro and Lee (2012). I use the U.S. as a benchmark economy, so all variables are relative to the U.S. Using the entry cost data constructed above, the model generates predicted levels of TFP for all 136 countries in the sample. To compare the model to the cross-country output per worker data, I need an estimate of the elasticity of output per worker with respect to TFP. To this end, I perform a simple regression of output per worker on TFP and use the resulting estimated elasticity of 1.73. This all implies model-generated values for TFP in each country $i$ equal to

$$TFP_i = \left( \frac{c_i}{c_{US}} \right)^{-0.45},$$

and output per worker $y$ in each country $i$ equal to

$$y_i = \left( \frac{c_i}{c_{US}} \right)^{-0.77}.$$

Table 3.2 compares TFP from the data to that generated by the model, while Table 3.3 does the same for output per worker. The third column in each table reports statistics generated by the same experiment using the standard CES model. ‘90/10 ratio (costs)’ refers to average TFP or output per worker of countries in the bottom decile of entry costs relative to the average of countries in the highest decile. ‘90/10 ratio (output)’ refers to average TFP or output per worker of countries in the highest decile of output per worker relative to the average of countries in the bottom decile.

The model accounts for 45% of the log-variation in TFP across countries and 44% of the log-variation in output per worker, while the standard model accounts for just 7% of each. These magnitudes are close to those reported by Herrendorf and Teixeira (2011), and consistent with Barseghyan (2008). In the appendix I report regression estimates of elasticities with respect to entry costs that are also consistent with these magnitudes.

Figures 3.1 and 3.2 provide a visual comparison of the model to the standard CES model, plotting TFP generated by each model against TFP in the data.

---

18In the appendix I explain in more detail how TFP is constructed.
### Table 3.2: Cross-Country Experiment: TFP

<table>
<thead>
<tr>
<th></th>
<th>data</th>
<th>model</th>
<th>CES model</th>
</tr>
</thead>
<tbody>
<tr>
<td>90/10 ratio (costs)</td>
<td>6.6</td>
<td>5.1</td>
<td>1.9</td>
</tr>
<tr>
<td>90/10 ratio (output)</td>
<td>9.0</td>
<td>3.7</td>
<td>1.7</td>
</tr>
<tr>
<td>variation in log(TFP)</td>
<td>0.52</td>
<td>0.24</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Includes 116 countries with TFP data. Variables relative to U.S.

### Table 3.3: Cross-Country Experiment: Output per Worker

<table>
<thead>
<tr>
<th></th>
<th>data</th>
<th>model</th>
<th>CES model</th>
</tr>
</thead>
<tbody>
<tr>
<td>90/10 ratio (costs)</td>
<td>29.5</td>
<td>17.9</td>
<td>3.3</td>
</tr>
<tr>
<td>90/10 ratio (output)</td>
<td>52.4</td>
<td>9.8</td>
<td>2.6</td>
</tr>
<tr>
<td>variation in log(output per worker)</td>
<td>1.71</td>
<td>0.76</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Includes all 136 countries. Variables relative to U.S.

## 3.5 Niche Firms, Mass Markets, and Distortions

Here I show how allowing for niche and mass-market firms amplifies the impact of distortions on aggregate outcomes. Consider the same model developed in Section 3.2, but now let all intermediate firms face a proportional tax such that firms keep only a fraction \( \tau \) of output. The problem now facing any firm \( i \) is to take the wage \( w \) as given and choose labor \( x_i \) to maximize

\[
\tau P_i x_i - wx_i,
\]

where \( P_i \) is a function of \( x_i \) as before. The profit-maximizing price for firm \( i \) is now

\[
P_i = \frac{w}{\tau \alpha_i}.
\]

Changing the intermediate firm optimization condition that must be satisfied in equilibrium to reflect this new optimal price, I can now follow the same steps described in Section 3.2.2.4 to solve for the equilibrium measure of niche and mass-market firms;

\[
N_n = \frac{(1 - \alpha_n) \phi \tau}{c_e (1 - \rho)} \quad \text{and} \quad N_m = \frac{(1 - \alpha_m)(1 - \phi) \tau}{c_e (1 - \rho)},
\]

as well as aggregate TFP;

\[
y = \Delta \cdot \left( \frac{\tau}{c_e} \right)^{\frac{\phi}{\alpha_n} + \frac{1 - \phi}{\alpha_m} - 1}.
\]

The elasticity of TFP with respect to \( \tau \) is equal to

\[
\frac{\phi}{\alpha_n} + \frac{1 - \phi}{\alpha_m} - 1,
\]

which is the same as the elasticity with respect to entry costs. Using the same argument described in Section 3.2.2.4, it follows that this elasticity is higher than that in a standard model without heterogeneous markets. Using the parameter values obtained in Section 3.4, this elasticity is about 2.5 times
that in the standard model (0.45 versus 0.18).

The intuition for this result is the same as for the impact of entry costs. In equilibrium with free entry, a tax on output increases the profits required by each firm to cover the cost of entry. Profits are increased via a reduction in the measure of firms, and the impact of fewer firms is governed by the elasticity of substitution. The end result is an elasticity of TFP with respect to $\tau$ equal to the elasticity with respect to entry costs in both the standard and the present models.

\section*{3.6 Conclusion}

I have shown how incorporating niche and mass-market firms into a standard model of monopolistic competition amplifies the impact of distortions, and entry costs in particular, on aggregate outcomes. The calibrated model amplifies elasticities with respect to these distortions by a factor of 2.5. Using cross-country data on entry costs the model can account for almost half of the log-variation in TFP and output per worker, consistent with recent estimates of a large impact of entry costs.

It is generally the case that the market guides agents in an economy to mitigate the effects of any tax or regulation by diverting resources to activities less affected by these policies. This is what leads to a smaller impact of bad policy in general equilibrium than what might be predicted by partial equilibrium analysis. One way around this is to tax all activities, as in Chapter 6 from Parente and Prescott (2000). Another way is to increase the cost of switching to the less-taxed activity. This second way is the approach used in Restuccia and Rogerson (2008), where taxes are tied to productivity and thus resources are diverted to much less productive firms. This chapter also uses the second approach. Higher entry costs lead to a higher concentration of resources in a smaller number of firms, since a large firm can spread a fixed cost over more output. In a standard model with a relatively high CES across all varieties, the aggregate cost of this lower level of specialization is small, since the gains to specialization are themselves small. But when some portion of expenditure is devoted to niche output, where the gains to specialization are high, entry costs have a much greater impact by forcing consumers to settle for varieties much less suited to their needs.
Figure 3.1: Model vs Data

Figure 3.2: Standard CES Model vs Data
Appendix A

Appendix to Chapter 1

A.1 Density Functions

A.1.1 General

To solve the competitive model, it is necessary to derive the joint density function of $h_{[1]}$, $h_{[2]}$, and $h_i$. If $h_{\ell}$ is distributed according to a continuously differentiable distribution $F_{\ell}(h)$ between 0 and $n_{\ell}^\theta$, the relevant joint density function is:

$$f(h_{[1]} = u_1, h_{[2]} = u_2, h_i = h_{[1]} | e, n_i, n_k \neq i) =$$

$$\int_0^{u_2} \cdots \int_0^{u_{e-1}} (e-1)! f(h_i = u_1) \prod_{\ell=2}^e f(h_k = u_\ell) du_\ell \cdots du_3,$$

where $e$ is the number of firms (draws), $n_i$ is firm-$i$’s level of research, and $n_k$ is every other firm’s level of research. Integrating over all $u_{\ell>2}$, this simplifies to:

$$f(h_{[1]} = u_1, h_{[2]} = u_2, h_i = h_{[1]} | e, n_i, n_k \neq i) = (e-1) f_i(u_1) f_k(u_2) F_k(u_2)^{\epsilon-2},$$

defined over all $u_1$ and $u_2$, such that $u_1 \leq n_1^\theta$, $u_2 \leq n_k^\theta$, and $u_1 \geq u_2 \geq 0$.

When calculating the expected value of an equilibrium variable, the relevant joint density function is similar to the one above, but without the qualification that firm-$i$ has the highest draw. Taking into account $n_i = n_k = n$ in equilibrium, the required density function is:

$$f(h_{[1]} = u_1, h_{[2]} = u_2 | e, n) = \int_0^{u_2} \cdots \int_0^{u_{e-1}} e! \prod_{\ell=1}^e f(h = u_\ell) du_\ell \cdots du_3,$$

or

$$f(h_{[1]} = u_1, h_{[2]} = u_2 | e, n) = e(e-1) f(u_1) f(u_2) F(u_2)^{\epsilon-2},$$

defined over all $u_1$ and $u_2$, such that $n^\theta \geq u_1 \geq u_2 \geq 0$. 

50
A.1.2 Uniform

If each firm-\(i\)'s quality draw is from a uniform distribution bounded by \(0\) and \(n_i^\ell\), then \(f_\ell(u) = \frac{1}{n_i^\ell}\) and \(F_\ell(u) = \frac{u}{n_i^\ell}\). It follows that the two joint density functions derived in Appendix A.1.1 become:

\[
f(h_1 = u_1, h_2 = u_2, h_i = h_{[1]} \mid e, n_i, n_{k \neq i}) = \frac{(e - 1)u_2^{e-2}}{n_i^\ell n_{k}^{(e-1)}},
\]

and

\[
f(h_1 = u_1, h_2 = u_2 \mid e, n) = \frac{e(e - 1)u_2^{e-2}}{n^\ell}.
\]

A.2 Competitive Equilibrium

Firm-\(i\)'s expected discounted profits (1.6) are:

\[
E_{t-1} \left( \frac{\pi_i}{R_i} \right) = \frac{1 - \alpha}{R_t} \psi_{t-1}(1 + g_t)^{1-\alpha} \text{Prob} \left[ h_{it} = h_{[1]} \geq \frac{h_{[2]} - 1}{\alpha} \right] \cdot E_{t-1} \left( \frac{\pi_i}{R_i} \right) h_{it} \geq \frac{h_{[2]} - 1}{\alpha}
\]

\[
+ \psi_{t-1}(1 + g_t)^{1-\alpha} \text{Prob} \left[ h_{it} = h_{[1]} \lt \frac{h_{[2]} - 1}{\alpha} \right] \cdot E_{t-1} \left( \frac{\pi_i}{R_i} \right) h_{it} \lt \frac{h_{[2]} - 1}{\alpha}
\]

\[- \psi_{t-1}((z + mn_i, t-1),
\]

where \(h_{[\ell]}\) denotes the \(\ell\)th-highest realized draw in period \(t\), \(R\) is the gross interest rate, \(g_t \equiv \frac{y_t}{y_{t-1}} - 1\), and \(\psi_{t-1} \equiv A_{t-1}\{j\}L^\alpha g_{t-1}^{1-\alpha}\). Using the joint density function from Appendix A.1.2, profits can expressed as:

\[
E_{t-1} \left( \frac{\pi_i}{R_i} \right) = \frac{1 - \alpha}{R_t} \psi_{t-1}(1 + g_t)^{1-\alpha} \int_0^{n_i^\ell} \int_{0}^{\alpha e^\ell} \frac{(e_{t-1} - 1)u_1^\alpha e_2^{e_{t-1}-2}}{\alpha_1 u_2^{1-\alpha} n_{t}^{(e_{t-1}-1)}} \text{du}_2 \text{du}_1
\]

\[
+ \psi_{t-1}(1 + g_t)^{1-\alpha} \int_0^{n_i^\ell} \int_{\alpha u_1}^{\alpha e^\ell} \frac{(e_{t-1} - 1)(u_1 - u_2)u_2^{e_{t-1}-2}}{u_1^{1-\alpha} n_{t}^{(e_{t-1}-1)}} \text{du}_2 \text{du}_1 - \psi_{t-1}(z + mn_i, t-1).
\]

Since it is now obvious that decisions are time-independent, expected discounted profits can be expressed as:

\[
E \left( \frac{\pi_i}{R} \right) = \frac{\alpha^{e-1}(1 - \alpha)\psi(1 + g)^{1-\alpha} n_i^{\theta(e+\alpha-1)}}{R(e + \alpha)n_k^{\theta(e-1)}}
\]

\[
+ \frac{\psi(1 + g)^{1-\alpha} n_i^{\theta(e+\alpha-1)}}{Re(e + \alpha)n_k^{\theta(e-1)}} \left[ 1 - \alpha^e \alpha^{e-1}(1 - \alpha)e \right] - \psi(z + mn_i),
\]

or

\[
E \left( \frac{\pi_i}{R} \right) = \frac{(1 - \alpha^e)\psi(1 + g)^{1-\alpha} n_i^{\theta(e+\alpha-1)}}{Re(e + \alpha)n_k^{\theta(e-1)}} - \psi(z + mn_i).
\]

\(^1\)To be precise, the above is true only if \(n_{it} \leq n_{k\ell}\). But if one were to instead assume \(n_{it} \geq n_{k\ell}\), the same free-entry and research conditions would result in equilibrium, where \(n_{it} = n_{k\ell}\).
Appendix A. Appendix to Chapter 1

Each firm-i chooses its level of research \( n_i \) to maximize \( E\left( \frac{\pi_i R}{n_i} \right) \), given \( n_{k \neq i} \) and the number of firms \( e \). Given identical firms, this results in the research condition (1.8):

\[
    n = \left[ \frac{(1+g)^{1-\alpha}(1-\alpha^e)\theta(e+\alpha-1)}{mRe(e+\alpha)} \right]^{\frac{1}{1-\alpha}}.
\]

Free entry ensures the number of firms \( e \) adjusts until \( E\left( \frac{\pi_i R}{n_i} \right) = 0 \). Given that \( n_i = n_k = n \), the free-entry condition (1.9) must also hold in equilibrium:

\[
    z + mn = \frac{(1+g)^{1-\alpha}(1-\alpha^e)n^{\alpha\theta}}{Re(e+\alpha)}.
\]

The growth rate of the economy \( g_t \) is equal to:

\[
    g_t = \frac{y_t}{y_{t-1}} - 1 = \left( \frac{\int_0^1 A_{t-1}[j]h_{t[1]}(j) \, dj}{\int_0^1 A_{t-1}[j] \, dj} \right)^{\frac{1}{\theta}} - 1,
\]

where \( A_{t-1}[j] \) is the highest quality in market-j in period \( t-1 \), and \( h_{t[1]}(j) \) is the realization of the best quality draw in market-j in period \( t \). Given arbitrary \( A_0(j) \), \( g_t \) can be expressed as:

\[
    g_t = \left( \frac{\int_0^1 A_0(j) \Pi_{s=1}^t h_{s[1]}(j) \, dj}{\int_0^1 A_0(j) \Pi_{s=1}^{t-1} h_{s[1]}(j) \, dj} \right)^{\frac{1}{\alpha}} - 1.
\]

In a symmetric economy where all parameter values including \( A_0 \) are common across markets and industries, the above expression is equal to:

\[
    g_t = \left( \frac{E\left[ A_0 \Pi_{s=1}^t h_{s[1]} \right]}{E\left[ A_0 \Pi_{s=1}^{t-1} h_{s[1]} \right]} \right)^{\frac{1}{\theta}} - 1.
\]

With independence of draws across time periods, the growth rate can be conveniently expressed as:

\[
    g_t = \left( \frac{A_0^\theta E\left[ h_{[1]}^{\alpha} \right]^{t}}{A_0^\theta E\left[ h_{[1]}^{\alpha} \right]^{t-1}} \right)^{\frac{1}{\theta}} - 1 = E\left[ h_{[1]}^{\alpha} \right]^{\frac{1}{\theta}} - 1.
\]

Using the joint density function from A.1.2, this becomes equation 1.10:

\[
    g = \left( \frac{\int_0^n \int_0^{u_1} e(e-1)u_2^{\alpha-2} \, du_2 du_1}{n\theta e} \right)^{\frac{1}{\theta}} - 1 = \left( \frac{en^{\alpha\theta}}{e+\alpha} \right)^{\frac{1}{\theta}} - 1,
\]

where the number of firms \( e \) and the level of research per firm \( n \) are common across markets and industries.

Given the interest rate \( R \) is equal to \( \frac{1+g}{\alpha} \) from the household’s problem, (1.10) can be combined with the equilibrium conditions for a single market (1.8) and (1.9) to characterize equilibrium in a symmetric economy (equations 1.12 and 1.13).
A.3 Measuring Competition

Competition (as measured in Aghion et al. (2005)) in an industry with a continuum of product markets is equal to:

\[
1 - \frac{1}{e} E \left( \frac{\text{Price} - \text{Marginal Cost}}{\text{Price}} \right).
\]

Using the pricing strategy given by (1.5), \( \frac{p(j) - MC(j)}{p(j)} \) in market-\( j \) is equal to:

\[
\begin{cases} 
1 - \alpha, & \text{if } \frac{h_2(j)}{h_1(j)} \leq \alpha \\
1 - \frac{h_2(j)}{h_1(j)}, & \text{if } \frac{h_2(j)}{h_1(j)} > \alpha.
\end{cases}
\]

Competition in an industry with a continuum of product markets is therefore;

\[
\text{Competition} = e^{\frac{e - 1}{e}} + \int_0^u \int_0^{\alpha u_1} \frac{\alpha f(h_{[1]} = u_1, h_{[2]} = u_2 | e, n) du_2 du_1}{e} + \int_0^u \int_0^{\alpha u_2} \frac{u_2 f(h_{[1]} = u_1, h_{[2]} = u_2 | e, n) du_2 du_1}{u_1 e}.
\]

Substituting in the density function from Appendix A.1.2 results in equation (1.16);

\[
\text{Competition} = e^{\frac{e^2 + \alpha e - 1}{e^2}}.
\]

An alternative measure of competition might give weight only to the Lerner Index of producing firms, so that competition would be measured as \( 1 - E(\text{Lerner}) \), or;

\[
\text{Competition (alternative)} = 1 - E \left( \frac{\text{Price} - \text{Marginal Cost}}{\text{Price}} \right) = e^{\frac{e^2 + \alpha e - 1}{e}}.
\]

To see that this alternative measure of competition is increasing in \( e \), note that;

\[
\frac{\partial}{\partial e} \left( e + \frac{\alpha e - 1}{e} \right) = \frac{\alpha e \ln(\alpha e)}{e^2} + \frac{1 - \alpha e}{e^2},
\]

which is greater than zero for all \( e > 0, \alpha \in (0, 1) \).

A.4 Proofs of Results

For convenience, I repeat the research and free-entry conditions;

research: \[ n = \left[ \frac{(1 + g)^{1-\alpha}(1 - \alpha^e)\theta(e + \alpha - 1)}{m Re(e + \alpha)} \right]^{\frac{1}{-\alpha^e}} \]

and

free-entry: \[ z + mn = \frac{(1 + g)^{1-\alpha}(1 - \alpha^e)n^{\alpha^e}}{Re(e + \alpha)}. \]
A.4.1 Number of Firms Decreasing in z

I start by showing that the equilibrium number of firms is decreasing in the fixed cost of introducing an innovation $z$. The right-hand side of the free-entry condition is simply the expected discounted profits of an innovating firm (divided by a constant), and so must be decreasing in the number of firms $e$, given $n$. Since $n$ maximizes expected profits, any increase in $n$ (given $e$) must increase $mn$ more than it increases expected profits. Since the left-hand side of the free-entry condition increases more than the right-hand side when $z$ and $n$ increase, the number of firms $e$ must decrease to satisfy the condition. Again because $n$ is maximizing expected profits, any decrease in $n$ (given $e$) must decrease expected profits more than it decreases $mn$. In this case, too, $e$ must drop to satisfy the free-entry condition.

A.4.2 Average Innovation Has Inverted-U Relationship with $z$

The fixed cost of introducing an innovation $z$ does not enter into the research condition, and so research per firm $n$ only depends on $z$ through its effect on the number of firms $e$. Since $e$ is monotonically decreasing in $z$, I simply focus on the relationship between $n$ and $e$. It is clear from the research condition than $n$ is increasing if $(1 - \alpha e)(e + \alpha - 1) - \alpha e \ln(\alpha) < 0$, and the denominator is increasing in $e$ at an increasing rate. It follows that there exists some $e^*$, such that research is decreasing in $e$ for all $e > e^*$. To prove that $n$ has an inverted-U relationship with $e$, I further show that research per firm is always higher when $e = 2$ than when $e = 1$:

$$n(e = 2) = \phi \left[ \frac{(1 - \alpha^2)(1 + \alpha)}{2(2 + \alpha)} \right]^{\frac{1}{1 - \alpha}} > \phi \left[ \frac{(1 - \alpha)\alpha}{(1 + \alpha)} \right]^{\frac{1}{1 - \alpha}} = n(e = 1),$$

for all $\alpha \in (0, 1), \theta \in (0, 1)$. By solving $\frac{\partial^2}{\partial e^2} \left( \frac{(1 - \alpha^2)(e + \alpha - 1)}{e(e + \alpha)} \right) = 0$ for $e$, I can also show that for values of $\alpha$ less than 0.5 (implying an elasticity of substitution equal to 2), $e^*$ is equal to about 2. As $\alpha$ increases from 0.5, $e^*$ also increases, as shown in Figure A.1. Finally, note that average innovation per firm $1.15$ is a simple increasing function of research per firm, and so its relationships with $e$ and $z$ are the same as those for $n$.

A.4.3 Competition Decreasing in $z$

To show measured competition is decreasing in $z$, I take for granted that the number of firms $e$ is decreasing in $z$, and show that competition is increasing in $e$. Measured competition for an industry with $e$ innovating firms $1.16$ is;

$$\frac{e^2 + \alpha e - 1}{e^2} = 1 - \frac{1 - \alpha^2}{e^2}.$$

Using the same argument made above, there must exist some $e^*$ such that competition is increasing in $e$ for all $e > e^*$. To prove competition is always increasing in $e$ (for all $e \geq 1$), I further show that competition is always higher when $e = 2$ than when $e = 1$;

$$\text{Competition (} e = 2 \text{)} = 1 - \frac{1 - \alpha^2}{2} > \alpha = \text{Competition (} e = 1 \text{)}.$$
A.4.4 Growth Decreasing in $z$

Denoting the industry-level growth rate as $g$, $1 + g = \left(\frac{n^\gamma}{e + \alpha}\right)$. To prove growth is decreasing in $z$, given that the number of firms $e$ is decreasing in $z$, I show that:

$$\frac{d[(1 + g)\alpha]}{de} = \frac{\partial[(1 + g)^\alpha|n]}{\partial e} + \frac{\partial[(1 + g)^\alpha | e]}{\partial n} \cdot \frac{\partial n(e)}{\partial e} > 0.$$  

Using the research condition above to obtain $n$ as an explicit function of $e$, the above inequality becomes:

$$\frac{\alpha n^\alpha}{(e + \alpha)^2} + \frac{e\theta n^\alpha}{(e + \alpha)n} \cdot \frac{\partial n(e)}{\partial e} > 0,$$

which holds if $\frac{n}{e\theta(e + \alpha)} > -\frac{\partial n(e)}{\partial e}$. The partial derivative $\frac{\partial n(e)}{\partial e}$ is equal to:

$$\frac{\partial n(e)}{\partial e} = n^\alpha \theta (1 + g_y)^{1 - \alpha} \cdot \frac{-\alpha\ln(\alpha e) + (1 - \alpha^e)(1 - \alpha)(2e + \alpha) - e^2}{e^2(e + \alpha)^2},$$

where $g_y$ is the economy-wide growth rate (held constant while $z$ is varied for only one industry). Using the expression for $n(e)$ given by the research condition, the inequality that needs to hold can now be expressed as:

$$\frac{(1 - \alpha^e)(e + \alpha - 1)}{\theta} > \alpha^e \ln(\alpha e) - (1 - \alpha^e)(1 - \alpha)(2e + \alpha) - e^2.$$  

Note that for any given $e$ and $\alpha$, the left-hand side of the inequality is decreasing in $\theta$. Assume for the moment that $\theta$ is constrained to be no higher than $\frac{1}{e + \alpha - 1}$. It would then be sufficient to show that the inequality holds for $\theta = \frac{1}{e + \alpha - 1}$. Under this assumption, the inequality becomes:

$$(1 - \alpha^e)(e + \alpha - 1)^2 > \alpha^e \ln(\alpha e) - (1 - \alpha^e)(1 - \alpha)(2e + \alpha) - e^2,$$
or
\[(1 - \alpha^e)(1 - \alpha) > \alpha^e \ln(\alpha^e),\]
which holds for all \(\alpha \in (0, 1)\) and \(e \geq 1\). To see that \(\theta\) can be no higher than \(\frac{1}{e + \alpha - 1}\), combine the free-entry and research conditions to obtain:
\[z = \frac{m n [1 - \theta (e + \alpha - 1)]}{\theta (e + \alpha - 1)}.\]
This expression can be interpreted as characterizing the fixed cost of introducing an innovation \(z\) needed to accommodate an equilibrium with a number of firms equal to \(e\) and a level of research per firm equal to \(n\), given all other parameter values. Assuming \(z \geq 0\) it is clear that \(\theta\) can be no higher than \(\frac{1}{e + \alpha - 1}\), for any given \(e\).

**A.4.5 Average Distance to Frontier Increasing with Competition**

To show that the distance between the quality of the average firm and the best firm is increasing with competition, I need to relax the assumption that each industry contains a continuum of product markets. With a continuum of markets, the best firm in an industry would always have a quality equal to the upper bound of the distribution from which firms draw, irrespective of the number draws in each market. Consider an industry made up of \(N\) products, with \(e\) innovating firms in each product market. The expected gap between the average and the best firm within this industry is;
\[\text{gap} = E \left( \frac{h_{[1]} - \frac{1}{e} \sum_{\ell=1}^{e \cdot N} h_{[\ell]}}{h_{[1]}} \right) = 1 - \frac{1}{e \cdot N} \sum_{\ell=1}^{e \cdot N} E \left( \frac{h_{[\ell]}}{h_{[1]}} \right),\]
where \(h \sim U(0, n^\theta)\) as before, and \(h_{[\ell]}\) now denotes the \(\ell\)th-best of \(e \cdot N\) draws. Using the same steps followed in Appendix A.1, one can derive the expected value of \(\frac{h_{[\ell]}}{h_{[1]}}\):
\[E \left( \frac{h_{[\ell]}}{h_{[1]}} \right) = \frac{e \cdot N - \ell + 1}{e \cdot N},\]
which results in an expected gap equal to;
\[\text{gap} = 1 - \frac{1}{e \cdot N} \sum_{\ell=1}^{e \cdot N} \left( \frac{e \cdot N - \ell + 1}{e \cdot N} \right) = \frac{1}{2} \left( 1 - \frac{1}{e \cdot N} \right),\]
which is increasing in \(e\) for any finite \(N\). Since measured competition is also increasing in \(e\), it follows that the gap between the average and the best firm in an industry is increasing in competition.

**A.5 Calibration**

To illustrate the relationships generated by the model in Section 1.2.4, I need values for \(\alpha, \theta,\) and \(m\), as well as the economy-wide growth rate and interest rate. I obtain these parameter values (as well as a value for \(z\)) by calibrating the model to match the following targets; industry-level productivity growth of 0.3%, an R&D to output ratio of 0.025, a level of competition (as measured in Aghion et al. (2005)) equal to 0.95, and a real interest rate of 5%. I further ensure that the number of firms are such that research
per firm is at its maximum (given all other parameter values). The productivity growth target is the median industry-level growth rate of five-factor-productivity across 473 6-digit NAICS manufacturing industries in the U.S., where each industry’s growth rate is an average of annual growth rates from 1959 to 2005 reported in the NBER-CES Manufacturing Industry Database. R&D intensity is from the National Science Foundation. Aghion et al. (2005) report a median level of measured competition for their sample equal to 0.95, and further report the peak of their inverted-U between firm-level innovation and competition as occurring at the median value of competition. I derive the equilibrium values of these variables as I discuss them in Section 1.2.4. Through this procedure I obtain the following values; \( \alpha = 0.816 \) (implying an elasticity of substitution of 5.4), \( \theta = 0.06 \), \( m = 6 \cdot 10^{-5} \), and \( z = 0.04 \). In Figures 1.1 through 1.3, I illustrate the effects on equilibrium variables from varying \( z \) (around the calibrated value) for one industry while keeping the interest rate and growth rate of the economy fixed. The growth rate of an economy with a representative industry and a population of 1 is identical to the industry-level productivity growth rate, so I assume the growth rate of the economy is also equal to 0.3%. Finally, I choose the range of \( z \) to ensure the model generates a range in measured competition from 0.87 to 0.995, which is the range reported for the sample of industry-years in Aghion et al. (2005).

### A.6 Monopoly-Only

In this section I consider each firm’s optimal level of research when the winning firm can always charge a monopoly price, regardless of how close its competitors may be.

Deviating from (1.6), any innovating firm-\( i \)'s expected discounted profits are now:

\[
E_{t-1} \left( \frac{\pi_i}{R_t} \right) = (1 - \alpha) \frac{\psi_{t-1}(1 + g_t)^{1-\alpha}}{R_t} \operatorname{Prob} \left[ h_{it} = h_{i[1]} \right] \cdot E_{t-1} \left[ h_{i[1]} \right] \left[ h_{it} = h_{i[1]} \right] - \psi_{t-1}(z + mn_{i,t-1}),
\]

where \( h_{i[1]} \) denotes the highest realized draw in period \( t \), \( R \) is the gross interest rate, \( g_t \equiv \frac{y_t}{y_{t-1}} - 1 \), and \( \psi_{t-1} \equiv A_{t-1}^{\alpha} L_t^{\alpha} y_t^{1-\alpha} \). Using the joint density function from Appendix A.1.2 and dropping the time subscript, profits can be expressed as:

\[
E \left( \frac{\pi_i}{R} \right) = \frac{(1 - \alpha)\psi(1 + g)^{1-\alpha}n_i}{R} \int_0^{n_i^e} \int_0^{u_1^e} \frac{(e - 1)u_1^\alpha u_2^{e-2}}{n_k^{\alpha\theta\theta}(\theta - 1)} du_2 du_1 - \psi(z + mn_i),
\]

or

\[
E \left( \frac{\pi_i}{R} \right) = \frac{(1 - \alpha)\psi(1 + g)^{1-\alpha}n_i^{\theta\theta(e + \alpha - 1)}}{R(e + \alpha)n_k^{\theta\theta(e - 1)}} - \psi(z + mn_i).
\]

Each firm-\( i \) chooses its level of research \( n_i \) to maximize \( E \left( \frac{\pi_i}{R} \right) \), given \( n_k \neq i \) and the number of firms \( e \), so the optimal level of research must satisfy:

\[
n = \left[ \frac{(1 + g)^{1-\alpha}(1 - \alpha)\theta(e + \alpha - 1)}{mR(e + \alpha)} \right]^{\frac{1}{1+\alpha\theta}}.
\]  

(A.1)

Comparing equations 1.8 and A.1, research in the monopoly-only model \( n^M \) is always higher than
research in the full model $n^C$, given a number of firms greater than one:

$$n^M = \phi(e)(1 - \alpha) \frac{1}{1 - \alpha \theta} > \phi(e) \left( \frac{1 - \alpha e}{e} \right) \frac{1}{1 - \alpha \theta} = n^C,$$

where $\phi(e) \equiv \left[ \frac{(1+g)^{1-\alpha}(e+\alpha-1)}{mR(e+\alpha)} \right] \frac{1}{1 - \alpha \theta}$. Further, $n^M$ is always increasing in the number of firms:

$$\frac{\partial n^M}{\partial e} = \frac{(n^M)^{\frac{\alpha g}{\alpha - \theta}}(1 + g)^{1-\alpha} (1 - \alpha) \theta}{(1 - \alpha \theta) mR(e + \alpha)^2} > 0.$$
Appendix B

Appendix to Chapter 2

B.1 Expected Value of the Best Draw: General

In existing markets a sequential innovator receives a quality \( A \) equal to \( A_{i-1}[1] \cdot h_i \), where \( A_{i-1}[1] \) is the best quality of the previous period and \( h_i \) is the realization of a random variable drawn from a distribution with support 1 and \( 1 + n_i \). Let \( f_i(h) \) and \( F_i(h) \) denote the pdf and cdf of an innovator-\( i \)'s draw, conditional on \( n_i \). It will be useful to transform these functions in the following way. Define a new random variable \( \hat{h} \in (0, 1) \) such that \( h_i = 1 + n_i \cdot \hat{h} \), so \( \hat{f}(\hat{h}) \) and \( \hat{F}(\hat{h}) \) are independent of firm-\( i \)'s level of research \( n_i \). Each firm chooses the same level of research \( n \) in equilibrium, so deriving the expected value of the best draw \( h_{i[1]} \) conditional on \( n \) and the number of draws \( e \) requires only a density function for \( \hat{h}_{i[1]} \):

\[
f(\hat{h}_{i[1]} = v_1 \mid e, n) = \int_0^{v_1} \cdots \int_0^{v_{e-1}} e! \prod_{\ell=1}^e \hat{f}(\hat{h} = v_\ell) du_\ell \cdots du_2,
\]
or

\[
f(\hat{h}_{i[1]} = v_1 \mid e, n) = e \hat{f}(v_1) \hat{F}(v_1)^{e-1},
\]
defined over \( v_1 \in (0, 1) \). The expected value of the best draw is therefore;

\[
E(h_{i[1]}) = 1 + n \cdot e \int_0^1 v \hat{f}(v_1) \hat{F}(v_1)^{e-1} dv_1.
\]

The above expression is obviously increasing in the level of research per innovation \( n \).

B.2 Expected Value of the Best Draw: Kumaraswamy

For the purposes of the calibration in Section 2.3, I assume sequential innovators take their quality draws from a one-parameter Kumaraswamy distribution, bounded by 1 and \( 1 + n \), where \( n \) is chosen by the firm. Each firm-\( i \)'s random variable \( h_i \) therefore has the following cumulative distribution function;

\[
F_i(h) = \text{prob}(h_i < h \mid n_i) = \left( \frac{h - 1}{n_i} \right)^\kappa,
\]

59
and the following probability density function:

\[ f_i(h) = \text{prob}(h_i = h \mid n_i) = \frac{\kappa(h - 1)^{\kappa - 1}}{n_i^\kappa}, \]

where \( \kappa \) is a shape parameter for the distribution.

To solve the model, it is necessary to derive the joint density function of \( h_i \) and \( h_{[1]} \):

\[
\int_1^{v_1} \cdots \int_1^{v_{e-1}} (e - 1)! f(h_i = v_1) \prod_{\ell=2}^e f(h_{-i} = v_\ell) dv_2 \cdots dv_e,
\]

where \( e \) is the number of firms (draws), \( n_i \) is firm-\( i \)'s level of research, and \( n_{-i} \) is every other firm's level of research. Integrating over all \( v_{\ell>1} \) and using the density function given above, this simplifies to:

\[
f(h_{[1]} = v_1, h_1 = h_{[1]} \mid e, n_i, n_{-i}) = \frac{\kappa(v_1 - 1)^{\kappa e - 1}}{n_i^\kappa n_{-i}^\kappa (e - 1)^{\kappa - 1}}.
\]

The density function of \( h_{[1]} \) is derived in a similar way, but without the qualification that firm-\( i \) has the highest draw:

\[
f(h_{[1]} = v_1 \mid e, n) = \int_1^{v_1} \cdots \int_1^{v_{e-1}} e! \prod_{\ell=1}^e f(h = v_\ell) dv_2 \cdots dv_e,
\]
or

\[
f(h_{[1]} = v_1 \mid e, n) = \frac{\kappa e(v_1 - 1)^{\kappa e - 1}}{n^{\kappa e}}.
\]

The density functions above can now be used to derive the following:

\[
\text{prob}(h_i = h_{[1]} \mid h_i = h_{[1]}) \cdot E(h_{[1]} \mid h_i = h_{[1]}) = \frac{n_i^{\kappa(e-1)}}{n_{-i}^{\kappa(e-1)}} \left( \frac{1 + \kappa e (1 + n_i)}{e(\kappa e + 1)} \right),
\]

and

\[
E(h_{[1]}) = 1 + \frac{\kappa en}{\kappa e + 1}.
\]

**B.3 Constrained Social Planner’s Problem**

Here I solve a constrained social planner’s problem to characterize the optimal investment rate and optimal amount of investment in sequential innovation, where the social planner is bound by the constraint that the relationship between the number of innovations per market \( e \) and the level of research per innovation \( n \) is identical to that in the decentralized economy. As in Section 2.3 I assume log-utility for the consumer. Given my focus on the stationary equilibrium of the economy, the constrained social planner’s problem is to choose the investment rate \( I \) and the per-market level of investment in sequential
innovation $X$ to maximize;\footnote{I normalize $A_0$ to 1 for all markets in period 0.}

$$U = \sum_{t=0}^{\infty} \beta^t u(C_t) = \sum_{t=0}^{\infty} \beta^t \log \left[ J^{\frac{1-\alpha}{\alpha}} \cdot (1 + g)^t \cdot (1 - I) \right],$$

or

$$U = \frac{\log(1 - I)}{1 - \beta} + \frac{(1 - \alpha)}{\alpha (1 - \beta)} \log(J) + \frac{\beta}{(1 - \beta)^2} \left( \frac{1 - \alpha}{\alpha} \right) \log(E(h_{[1]}),$$

s.t. $J = \frac{I}{\delta z_0 + (1 - \delta)X},$

where $X = e^{z + \frac{n^\theta}{\theta}}$.

The first-order conditions for the above problem can be manipulated to find that the optimal investment rate $I$ is equal to $1 - \alpha$. It is immediately clear that the investment rate in the decentralized economy is lower without patent protection, and even lower with patent protection. Optimal investment in sequential innovation per market $X$ is characterized by the following equation;

$$\left( \frac{\partial}{\partial X} E(h_{[1]}) \right) = \frac{(1 - \delta)(1 - \beta)}{\beta (1 - \alpha)}.$$
Appendix C

Appendix to Chapter 3

C.1 Sensitivity to Parameter Values

The elasticity of TFP with respect to entry costs generated by the model in Section 3.2 depends on the values of three parameters: $\alpha_n$, which determines the elasticity of substitution between niche varieties; $\alpha_m$, which does the same for mass-market varieties; and $\phi$, which determines the total revenue share of niche output. Here I consider the sensitivity of the elasticity of TFP (and output per worker) to changes in the target values used to calibrate the model in Section 3.3. The benchmark targets are an average elasticity of substitution ($CES$) of 6.5, a difference in the TFPR statistic between the 90th and 10th percentile establishments (90/10 gap) of 1.19 (implying a 90/10 ratio for TFPR of 3.3), and a standard deviation of the same TFPR statistic across firms within an industry of 0.49. These imply elasticities of TFP and output per worker with respect to entry costs of -0.448 and -0.775.

I start with an increase in $CES$ to 7.5 (above the range recommended by Imbs and Méjean (2009) of 6-7). As discussed in the introduction, this lowers the gains to specialization and so reduces the impact of entry costs in any model. The resulting elasticity of TFP with respect to entry costs is -0.385, while that of output per worker is -0.667. A decrease in $CES$ to 5.5 implies elasticities of -0.534 and -0.925.

I now consider an increase in the dispersion of TFPR by targeting a standard deviation of 0.6 and a 90/10 ratio of 4. An increase in the 90/10 ratio implies a bigger gap between $\alpha_m$ and $\alpha_n$, while a higher standard deviation implies a lower $\phi$. The implied elasticity of TFP with respect to entry costs is -0.524, while that of output per worker is -0.907. I consider a decrease in the dispersion of TFPR by targeting a standard deviation of 0.4 and a 90/10 ratio of 2.5. These values imply an elasticity of TFP with respect to entry costs of -0.342, and an elasticity for output per worker of -0.592.

Each of the different targets above reflect a change of approximately 15% to 25% from the benchmark targets. These different targets imply elasticities 19% to 24% higher or lower than the benchmark elasticities. In all cases, the implied elasticities remain at least 1.9 times larger than in the standard $CES$ model.
C.2 Constructing Entry Costs

Data for the regulatory cost of doing business comes from the World Bank Doing Business Survey.\footnote{World Bank (2012) describes each variable and the survey methodology.} I use the monetary cost and time required to start a business (registering the business, getting the necessary licenses, etc...), obtain construction permits, register property, and obtain an electricity connection, as well as the time required to comply with tax-related requirements. I use Penn World Tables v7.0 data to transform all entry cost measures into fractions (or multiples) of output per worker, making these measures comparable to those used in the model. To the same end, I use the present discounted value of all costs. This requires a value for the intertemporal discount rate of firms $\rho$, which in turn requires values for the probability of firm death $\lambda$ and the real interest rate $r$. I use $\lambda = 0.1$ and $r = 0.05$, which are common in cross-country calibrations. These values imply a $\rho$ equal to 0.86.

To ease exposition, let $c_e$ denote total entry costs, $c_{reg}^M$ monetary regulatory startup costs, $c_{reg}^W$ the wage costs of both startup delays and the post-startup time to do taxes, and $c_{nonreg}$ nonregulatory startup costs. I assume the wage cost of one year is a fraction $\frac{2}{3}$ of output per worker, consistent with a model that includes capital and in which capital receives one third of all factor payments. Acs et al. (2005) report an average monetary startup cost in the U.S. in 2005 equal to 86% of output per worker. I calculate nonregulatory costs $c_{nonreg}$ (assumed constant across countries) equal to $0.46 = 0.86 - c_{reg}^M$.

In the model, free entry requires $c_e = \pi \cdot \left(\frac{1}{1-\rho}\right)$, where $\pi$ denotes per-period profits relative to output per worker. If $c_e$ includes the cost of lost profits, then this condition can be rewritten as:

$$\tilde{c}_e = \pi \cdot \left(\frac{1}{1-\rho} - delay_0 - \rho \cdot delay_1 - \rho^2 \cdot delay_2 - \ldots\right),$$

where $\tilde{c}_e = c_{nonreg} + c_{reg}^M + c_{reg}^W$ is total costs without lost profits, and $delay_t$ denotes the fraction of the $t^{th}$ year included in total delays.\footnote{For example, a delay of 547 days would imply $delay_0 = 1$, $delay_1 = 0.5$, and $delay_{t>1} = 0$.} This all implies total entry costs $c_e$ can be calculated as:

$$c_e = \frac{\tilde{c}_e \cdot \left(\frac{1}{1-\rho}\right)}{\left(\frac{1}{1-\rho} - delay_0 - \rho \cdot delay_1 - \rho^2 \cdot delay_2 - \ldots\right)}.$$

C.3 Constructing TFP

I construct TFP following Klenow and Rodríguez-Clare (2005). TFP for a country $i$ relative to the U.S. in 1996 is calculated as:

$$\text{TFP}_i = \frac{y_i}{k^*_i h^*_i},$$

where $y$ is output, $k$ is capital, $h$ is human capital, and all variables are per worker and relative to the U.S.

Penn World Tables v7.0 contains annual data on output per worker and per capita in constant PPP dollars, population, and the investment share of output for a large number of countries as far back as 1950. I drop all observations after 1996 and then construct an initial capital-output ratio $K_0/Y_0$ for the
first year each country $i$ reports the required data;

$$\frac{K_{i,0}}{Y_{i,0}} = \frac{\bar{I}_i}{g + \delta + \bar{n}_i},$$

where $\bar{I}_i$ is the average investment share in country $i$ over the sample, $\bar{n}_i$ is its average population growth rate, $\bar{g}$ is the average world growth rate in output per capita which I assume to be 0.02, and $\delta$ is the depreciation rate of capital which I assume to be 0.08. I drop estimates for countries without at least ten years of continuous data, and then construct a capital-output ratio for each year up to 1996 using the usual capital accumulation equation.

I calculate human capital per worker as $\exp(0.085 \cdot \text{years of schooling})$, where years of schooling refers to the average years of schooling of people at least 25 years of age. This last is from Barro and Lee (2012). The value 0.085 represents the return to schooling, and is common in the literature (including Klenow and Rodríguez (2005)).

### C.4 Empirical Estimates

The results of the cross-country experiment in Section 3.4.2 are not directly comparable to papers with different measures of entry costs, so I report here the empirical results of regressions using the entry costs constructed in this paper. I regress TFP and output per worker on total entry costs, controlling for institutional quality using two variables from Barseghyan (2008). The first control variable is the property rights index reported by the Heritage Foundation and described in Heritage Foundation (2012). The second control is the debt recovery rate for each country reported by the World Bank Doing Business Survey and described in World Bank (2012), for which I use the first reported value for each country. This is a measure of the fraction of debt recovered by claimants in bankruptcy proceedings. Table C.1 displays the results of the two regressions.

The elasticities for TFP and output per worker generated by the model are 98% and 97% of the estimated elasticities, which are in turn consistent with entry costs accounting for about half of the log-variation in TFP and output per worker in the data. These estimated elasticities are not comparable to those in Barseghyan (2008) due to different measures of entry costs. But Barseghyan reports a decrease in TFP of 22% associated with an increase in one half of the standard deviation of entry costs from the mean. My estimates imply a decrease of 25%.

---

3 As in Barseghyan (2008), I use the average value of all reported years from 1987 to 1996 for each country. I also rescaled the reported values to match those used by Barseghyan.
Bibliography


