Innovation and Welfare Impacts of Disclosure Regulation: A General Equilibrium Approach

Li Yang

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy under the Executive Committee of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024
Abstract

Innovation and Welfare Impacts of Disclosure Regulation:
A General Equilibrium Approach

Li Yang

I develop a general equilibrium model to examine the innovation and welfare effects of expanding mandatory financial disclosure to a broader set of firms. In the model, disclosure by relatively small firms reveals proprietary information about their local markets, which helps larger firms enter and compete. Consistent with previous empirical findings, the model predicts that mandatory disclosure encourages (discourages) innovation by larger (smaller) firms. More importantly, I identify conditions for when expanding the scope of disclosure regulation increases aggregate innovation and/or welfare. I structurally estimate the model using innovation data and plausibly exogenous variation in the extent of disclosure regulation in Europe. My estimates suggest that subjecting 15% more firms to full reporting requirements decreases aggregate innovation by around -0.26% but increases welfare by around 1%. This disparity is driven by the fact that production shifts to larger firms that innovate less but are more efficient in exploiting the fruits of innovations.
# Table of Contents

Acknowledgments ................................................................. v

Chapter 1: Introduction ......................................................... 1

Chapter 2: Model ................................................................. 9
  2.1 Final Goods Firm ......................................................... 9
  2.2 Intermediate Goods Firms ............................................. 10
  2.3 Representative Household .............................................. 15
  2.4 Equilibrium .............................................................. 16
  2.5 Reallocation Effects ................................................... 16
  2.6 Aggregate Comparative Statics ....................................... 19
  2.7 Welfare Analysis ....................................................... 20

Chapter 3: Estimation and Quantitative Analysis ....................... 24
  3.1 Institutional Background .............................................. 24
  3.2 Data ................................................................. 25
  3.3 Validation of Model Predictions ..................................... 26
  3.4 Estimation Strategy ................................................... 28
    3.4.1 Parameters ...................................................... 28
    3.4.2 Moments ......................................................... 29
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.3 Identification</td>
<td>30</td>
</tr>
<tr>
<td>3.5 Estimation Results</td>
<td>32</td>
</tr>
<tr>
<td>3.6 Counterfactuals</td>
<td>33</td>
</tr>
<tr>
<td>Conclusion</td>
<td>36</td>
</tr>
<tr>
<td>References</td>
<td>38</td>
</tr>
<tr>
<td>Appendix A: Figures and Tables</td>
<td>42</td>
</tr>
<tr>
<td>Appendix B: Alternative Model Assumptions</td>
<td>49</td>
</tr>
<tr>
<td>B.1 No Direct Competition</td>
<td>49</td>
</tr>
<tr>
<td>B.2 Cournot Competition</td>
<td>53</td>
</tr>
<tr>
<td>Appendix C: Model Proofs</td>
<td>59</td>
</tr>
<tr>
<td>C.1 Preliminaries</td>
<td>59</td>
</tr>
<tr>
<td>C.2 Proofs for Main Analysis</td>
<td>60</td>
</tr>
<tr>
<td>C.3 Proofs for Alternative Model Assumption Analysis</td>
<td>67</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Sequence of Events .................................................... 12
2.2 Total Welfare and Small-firm Disclosure .......................... 23
A.1 Comparative Statics .................................................... 42
A.1 Comparative Statics (continued) .................................. 43
A.2 Counterfactual Experiments ......................................... 44
List of Tables

A.1 Data Definitions ................................................. 45
A.2 Descriptive Statistics ........................................... 46
A.3 Innovation Impacts of Mandatory Financial Disclosure (by Group) ........ 47
A.4 Estimation Results ................................................ 48
Acknowledgements

I would like to thank my committee members–Tim Baldenius (co-chair), Thomas Bourveau, Matthias Breuer (co-chair), Hui Chen, and Jon Glover–for their continuous support and invaluable mentorship. I would like to thank Yue Chen, Qingkai Dong, Ilan Guttman, Jinhwan Kim, Anthony Le, Volker Laux, Xiaojing Meng, Stephen Penman, Phil Stocken, Laura Veldkamp, and workshop participants at the AAA/Deloitte Foundation Doctoral Consortium and Columbia Business School for helpful comments and suggestions. I acknowledge the financial support of the Lang Entrepreneurship Center, the Chazen Institute for Global Business, and the W. Edwards Deming Center at Columbia Business School.
Chapter 1: Introduction

There has been much debate about the extent to which mandatory disclosure should be imposed on corporations [1, 2]. European regulators, for example, require limited-liability companies to disclose financial statements, but differ in the extent to which smaller corporations are exempted from this requirement. The U.S. Securities and Exchange Commission limits its disclosure requirements to companies with dispersed investors bases (e.g., trading in public markets), but recently also considered extending its requirement to more closely held corporations [3, 4]. Research suggests that the extension of disclosure requirements tends to benefit larger firms, as disclosures by other firms serve as important sources of learning and can guide investment and innovation decisions.\(^1\) However, the flip side of this learning benefit is that the smaller firms face increased product market entry and competition, which potentially diminishes their innovation incentives.\(^2\) The countervailing effects on firms’ competitive positions and innovation incentives make it unclear a priori whether the expansion of disclosure regulation fosters or hurts aggregate innovation and welfare [14, 15].

I develop a general equilibrium model to examine the innovation and welfare effects of extending financial disclosure requirements to a broader set of firms, especially to smaller firms operating in local markets. In the model, disclosure by small firms enables larger ones to enter and better compete in these markets. Consistent with the literature, the model predicts that mandatory disclosure encourages innovation among large firms while discouraging it among smaller ones. My model elucidates conditions under which these changes enhance aggregate innovation and welfare. I further bring the model to data and examine the practical implications of the European disclosure regime. Leveraging detailed innovation data and plausibly exogenous variation in the extent

---

\(^1\)For example, [5], [6], [7], and [8].

\(^2\)[9], [10], [11], [12], and [13], among others, suggest that financial reporting affects competitive outcomes.
of disclosure regulations, I structurally estimate the model and evaluate the consequences of the financial reporting mandates in the European Union.

My model is a one-period abstraction of the endogenous growth models of [16] and [17]. As is common with this type of model, there is a representative household that consumes, provides skilled labor, and owns other input factors; a final goods firm that combines all intermediate goods; and many intermediate goods firms that are the focus of the model. Household consumption constitutes a measure of welfare in this economy.

Within the intermediate goods sector, there are small firms that each operate in a local/niche market and are local monopolists at the outset. There are also large firms that can enter the local markets. Each large firm enters a small firm’s market and then innovates via product development and competes with the incumbent small firm in the product market. The competition takes the form of Bertrand competition, where the firm with the lower marginal cost sets its price at the other firm’s marginal cost.

Each local market is affected by a market condition shock. This market shock represents unique developments in an industry that are critical for a firm’s success. It could pertain, for example, to differences in consumer tastes, key supply chain relationships, or applications of advanced technology. Small firms observe the local market conditions perfectly, which is meant to capture their knowledge advantage in their respective markets; while large firms do not observe the local conditions directly. How effectively a large firm can compete in the local product market depends on how well it can adapt to local conditions.

A priori, large firms have a cost disadvantage in local markets on account of their lack of local knowledge. They must rely on small firms’ disclosures and other public information to learn about market conditions. If large firms successfully adapt to the local conditions, their cost disadvantage turns into a cost advantage. Because of Bertrand competition and the ensuing limit pricing strategies, the cost advantage constitutes large firms’ profit margin upon successful adaptation.

---

3 I abstract away from an infinite horizon, because I focus on the cross-sectional effects of the disclosure regulations.  
4 In Appendix B.2, I examine an alternative case where firms engage in Cournot competition in the product market. All analytical results hold qualitatively, although the expressions of the exact conditions are different from the main paper.
For the small firms, their profit margin equals the monopoly margin if they remain local monop-
olists; otherwise it equals the large firms’ cost disadvantage if large firms, upon entering, fail in
the adaptation attempt. Therefore, firms’ expected profit margin is determined by both the success
probability of the adaptation attempt and the potential cost advantage (disadvantage).

The expected profit margin in the product market is a key factor that shapes firms’ innovation
incentives, because firms innovate in anticipation of the profit from developing a new product. The
expected profit margins are therefore the primary channel through which disclosure regulations
affect innovation. As large firms learn more through mandatory disclosure by small firms, their
adaptation is more likely to succeed, and they are more likely to obtain a cost advantage. The in-
crease in anticipated product market profits in turn encourages large firms to develop new products.
Conversely, for small firms, increased mandatory disclosure weakens their competitive position in
the product market, discouraging innovation. This is the learning channel through which disclo-
sure regulations play a role in reallocating innovation activities. What is a learning benefit for large
firms constitutes a business stealing cost for small firms. Furthermore, large firms’ expansion in
innovation activities puts an upward pressure on the equilibrium wage of the skilled labor, which
in turn makes it more expensive for small firms to innovate. This input price externality further
contributes to a reallocation of innovation from small to large firms.

These model predictions are well aligned with stylized findings from the recent empirical lit-
erature. For instance, [10], [18], and [19], among others, suggest that transparency regulations
induce product market entry, competition, and increased elasticity of consumer demand. The re-
allocation effect predicted by the model is also documented in reduced-form studies that suggest
that mandatory disclosure could induce resources, economic activities and innovation to flow to
less regulated or less affected firms [20, 21, 12, 8]. My model sheds light on the mechanism
leading to the reallocation.

Building on these results, I further identify conditions under which enhancing small firms’ dis-
closure requirements increases aggregate innovation expenditure and welfare. In general, aggre-
gate innovation expenditure can increase, decrease, or display a U-shaped relation with disclosure
regulations. Increased disclosure requirements redistribute innovation and production from small firms to large ones. Whether this redistribution has a positive aggregate effect depends on the expected product market profits, which, in turn, influence the innovation incentives of both groups of firms. If large firms’ potential cost advantage is low, their innovation incentives are muted, and the increase in their innovation cannot compensate for the loss in small firms’ innovation. In this case, aggregate innovation decreases with disclosure regulation. Conversely, when large firms have a significant potential cost advantage, aggregate innovation increases with disclosure regulation. Furthermore, with a moderate potential cost advantage, aggregate innovation tends to first decrease and then increase with disclosure requirements. This is because the expected profits are endogenous to disclosure: more accurate information can increase large firms’ expected profit, which adds to their innovation incentives.

The overall welfare in this economy is measured by the total consumption accruing from all innovations. To better understand the welfare impact of disclosure regulations, it is useful to decompose the effect into two channels: the total number of innovations and the consumption value delivered per unit of innovation. I refer to the latter as the consumption intensity of innovations. Intuitively, an innovation has higher consumption intensity if the firm has a lower marginal cost, so that it can produce cheaply, or if it faces competition, which compels it to lower prices and increase quantity. For the first channel, because innovations are an integral component of welfare, the impacts of disclosure regulations on aggregate innovation, discussed in the previously paragraph, carry over. Total number of innovations can again increase, decrease or display a U-shaped relation with disclosure requirements. However, in contrast to aggregate innovation, welfare has an additional dimension, which is the effect from the second channel: how much consumption can these innovations deliver? In terms of consumption intensity, reallocation to large firms has a positive effect, because production is reallocated to large firms that have lower cost and greater efficiency in exploiting the fruits of innovations than small firms. Combining the two effects, the composite effect on welfare can again go in either direction, depending on large firms’ potential cost advantage. But because of the positive consumption intensity effect, the cost advantage re-
quired for disclosure regulation to be welfare enhancing is generally lower than that required for it to increase aggregate innovation.

Next I use my model as a laboratory to examine the innovation and welfare consequences of the reporting mandates in Europe. The European setting offers two distinct advantages. First, the European Union mandates disclosure of audited financial statements for a substantial population of limited-liability firms (both public and private), and the mandate is also imposed on many small to medium-sized firms (e.g., those with more than 50 employees, 4 million EUR in total assets). Second, each country sets its own exemption thresholds (based on total assets, sales, and employee count), creating plausibly exogenous variation in regulatory intensity across countries and industries. I obtain confidential data on firms’ innovation activities from Eurostat’s Community Innovation Survey. The data covers 17 countries for the period from 2000 to 2018. Additionally, I obtain financial data from ORBIS and information about EU disclosure regulations from [11] for the corresponding countries and timeframe.

I estimate the model using the simulated method of moments approach. I use seven moments to estimate five parameters. Among the moments, mean and variance of both large and small firms’ profit margins broadly capture innovation incentives. Share of researchers employed by large firms reflects general presence of large firms. And the innovation elasticities of both large and small firms capture the impacts of mandatory disclosure by small firms. Innovation elasticity is the change in firms’ innovation expenditure when mandatory disclosure is imposed on small firms. Empirically, I proxy for the elasticities using the coefficients from a firm-level regression of innovation on regulation intensity. The regulation intensity measure calculates the percentage of firms subject to the disclosure mandates in a specific country-industry. This percentage is computed based on a hypothetical time- and country-invariant firm-size distribution for each industry in Europe. Thus, any variation in the regulation intensity measure stems only from the differences in exemption thresholds and firm size distributions across industries, which alleviates endogeneity concerns [23, 24]. Therefore, the estimation can recover a primitive parameter of disclosure precision to the

---

5See [22] for a comprehensive overview of accounting-related regulatory changes for each country of the European Union.
extent that the regression coefficients capture a causal response of firms’ innovation to disclosure regulation.

The model fits well and matches the moments closely. I use the estimated model to study the effects of the EU reporting mandates. Specifically, I evaluate a counterfactual policy that extends full disclosure requirements to 15% more firms, which corresponds to an additional three percentage points of disclosing firms. This counterfactual experiment represents a policy change from requiring on average 20% firms to requiring 23% firms to disclose. This particular policy experiment is chosen because three percentage points is the average variation in the disclosure regulation intensity measure.\textsuperscript{6}

The estimation result suggests that this policy would improve external knowledge about the affected firms and reduce uncertainty regarding their local market conditions by 33%. This estimated magnitude of uncertainty reduction is similar to the findings regarding disclosure usefulness in the public stock market [e.g., 25, 26]. Moreover, the affected firms face an 11.9% increase in the probability of losing monopoly status and being out-competed, consistent with estimates from prior studies about the effect of disclosure on competitive entry [e.g., 18, 19]. Resulting from these competitive consequences, there is a reallocation of innovation from large to small firms in that the innovation expenditure of large firms increases by 7.7%, whereas that of small firms decreases by 5.1%, which translates to an increase of around 80 thousand Euro and 5.7 thousand Euro for an average large and small firm, respectively.\textsuperscript{7} Finally, combining the positive effect on large firms’ innovation and the negative effect on small firms, the policy decreases aggregate innovation by 0.26% while increases welfare by 0.96%. This result is driven by the fact that innovation and production are reallocated from small to large firms, which develop fewer products (innovate less) but better exploit the fruits of innovations.

My paper makes several contributions. It is a first attempt at building a general equilibrium model that incorporates heterogeneous effects of disclosure regulation and is suitable for welfare

\textsuperscript{6}Extrapolation beyond this level would subject to larger extrapolation error.

\textsuperscript{7}In the setting of patent disclosure, [21] find that less affected firms benefit from the mandatory disclosure in that their patents receive on average 33% more forward citations, while more affected firms’ patents receive 27% less citations, which is approximately of an economic value of $17.1 and $13.9 million over a 5 year period.
Analysis. Several review studies emphasize the importance of considering both positive spillovers and the potential proprietary costs on different groups of firms and evaluating the welfare consequences when it comes to evaluating the real effects of disclosure regulations [27, 28]. More recently, [15] calls for research on the welfare implications of accounting regulations. I provide a theoretical framework that elucidates potential first-order forces and key mechanisms in the interaction of disclosure regulation, product market competition, and firm innovation. This framework can be useful in evaluating potential policy implications and interpreting other empirical findings.

This paper also contributes to the literature on disclosure regulations and firm innovation. This is a booming research area, as intangible investments and innovation have become ever more important for the growth of developed economies [29, 30]. In the accounting literature, two types of disclosures have been of interest: innovation disclosure and financial disclosure. On a conceptual level, the two types of disclosures generate similar trade-offs: positive externalities and proprietary costs. In terms of the exact information disclosed, innovation disclosure reveals proprietary information about the innovation process itself (e.g., R&D disclosure, patent disclosure), while financial disclosure is informative about the business environment and the context of the innovations [31]. My model is built to study the effect of financial disclosure (e.g., the disclosure of an income statement or MD&A.), but it can also be applied more generally to examine other sources of information.

On the topic of financial disclosure and firm innovation, [12] suggest that mandatory financial reporting may redistribute innovation to a few large firms. [32] document that innovation expenditures and outcomes of young life-cycle firms are significantly damped by financial regulations (e.g., the Sarbanes-Oxley Act). [31] provide evidence that public firms’ financial disclosures provide context of commercialization for patents and facilitate future patent sales. [33] find that information disclosed in 10-Ks facilitates peer firms’ follow-on innovations. My paper synthesizes these recent empirical findings into a model, and adds to these reduced-form studies by going further and estimating the welfare impact of the disclosure regulations. Regarding innovation dis-

---

8In particular, [28] state: “(O)verall, we call for research that examines and quantifies the benefit of disclosure to peer firms and calibrates these benefits against the cost of disclosure to disclosing firms” (p. 19).
closure, [21] show a similar pattern of redistribution: there’s an increase in innovations for firms whose rivals reveal more information in patent disclosures and a decrease in innovation for firms whose own disclosures are divulged to competitors. Evidence that weak disclosing firms face proprietary cost is also shown by [34], who finds that firms’ innovation is deterred by disclosures of early-phase trial from strong rivals but encouraged by disclosures from weak rivals.

More generally, my paper relates to studies examining macroeconomic consequences of firm disclosures in structural models. [17] study how managers trade off intangible investments and earnings manipulation and the associated real consequences. My model is similar to theirs in the way firms innovate and produce. Additionally, [35] examines the macro impact of managers’ short-term incentives to cut R&D investments to meet or beat analysts’ forecasts. [36] structurally estimates a general equilibrium model that evaluates the resource allocation role of accrual accounting.
Chapter 2: Model

I develop a general equilibrium model that incorporates product market competition, innovation, learning, and proprietary costs, which are key aspects of concerns for evaluating disclosure regulation. The model includes a representative household, a representative final goods firm and small and large intermediate goods firms. The one-period model is an abstraction of the endogenous growth models such as in [16] and [17]. The purpose is to illustrate the main channels through which disclosure regulation affects product market competition and innovation.

2.1 Final Goods Firm

A representative final goods firm produces final output by combining some input factor (e.g., labor, land) and intermediate goods, with the following constant return to scale technology:

\[ Y = A^{1-\theta} \int_0^N x_j^\theta \, dj. \]  \hspace{1cm} (2.1)

\( A \) is the quantity of the input factor, with production share \( 1 - \theta, \theta \in (0, 1) \); \( x_j \) is the quantity of intermediate good variety \( j \); and \( N \) is the number of total product varieties developed. All else equal, the larger \( N \), the more varieties there are and the larger the final output. The total number of product varieties, \( N \), and the quantity of each individual product variety, \( x_j \), are two key dimensions that contribute to the final output, through which the welfare effects of disclosure regulation will be evaluated in this model. Throughout, final goods are the numeraire and their price is normalized to 1.

---

Within the category of endogenous growth models, this is an expanding variety model. See the review of the literature by [37].
The final goods firm solves the following maximization problem:

$$\max_{A,\{x_j\}} Y - P^A A - \int_0^N p_j x_j \, dj.$$  \hspace{1cm} (2.2)

$P^A$ is the price for input $A$ (e.g., labor wages, rent price for land), and $p_j$ is the price for product variety $j$. The quantity $x_j$ demanded by the final goods firm is thus:

$$x_j(p_j, A) = \left( \frac{\theta}{p_j} \right)^{\frac{1}{1-\theta}} A.$$  \hspace{1cm} (2.3)

Final goods are either consumed by a representative household, or used as inputs to produce intermediate goods. The resource constraint for the economy dictates that consumption plus production inputs make up total output:

$$Y = C + X,$$  \hspace{1cm} (2.4)

where $C$ denotes the part of final goods used as consumption, and $X$ that used as production inputs.

## 2.2 Intermediate Goods Firms

The intermediate goods firms are the main players of interest. I refer to them as just “firms" when it does not raise confusion. Intermediate goods firms engage in innovation and production activities. Innovation takes the form of new product development. Firms innovate by hiring researchers (skilled labor) to develop new product varieties, which bring in cash flow streams. It can also be interpreted as process innovation or general improvement in production efficiency. Each product developed and produced is sold to the final goods firm. Intermediate goods firms face a downward sloping demand curve for each product variety as specified by (2.3).

There are two exogenous types of intermediate goods firms: a measure 1 of small firms, indexed

\footnote{The construct that intermediate goods firms use final goods as inputs to produce, and also sell their products to the final goods firm is common in the literature, especially in expanding variety models [see 38]. It can be interpreted as intermediate goods are produced using the same technology as the final good.}
by \( s \in [0, 1] \); they each operate in one local market. In addition, there is a measure \( \mu \) of large firms, indexed by \( b \in [0, \mu] \) (\( b \) for big firms), who are potential entrants into the local markets. A generic intermediate goods firm is indexed by \( i \in \{s, b\} \). A local market is indexed by \( n \in [0, 1] \).

**Timeline.** Figure 2.1 summarizes the sequence of events. At the beginning of the period, an exogenous market condition shock \( \omega_n \sim N(0, \tau^{-1}) \) is realized for each market \( n \). This market shock is iid across markets and represents unique developments such as consumer tastes, key supply chain relationships, technology applications etc., that are critical in a particular market. Firms’ profitability from operating in a niche market depends on how well they can adapt to the local market conditions. Small firms are assumed to observe the shocks perfectly. This assumption is meant to capture that small, local firms usually have a knowledge advantage compare to large, nationally-oriented firms. Mandatory disclosure imposed on small firms imperfectly reveals the values of the respective market condition shock. Additionally, there is public information about the local market conditions.

The model has an innovation stage (t2) and a competition stage (t3). At the innovation stage, small firms innovate in their respective market, and large firms are looking to enter the local markets. Specifically, each large firm randomly draws a small firm’s market and enters.\(^3\) As a result, each market will have at most two firms. The large firm attempts to adapt to the local market condition \( \omega_n \) by taking an adaptation action \( a_{bn} \) and then innovates. How the adaptation action affects both types of firms’ profits is clarified later when the mechanism of competition is explained.

At the competition stage, large firms’ marginal costs are commonly observed before production takes place. In a local market \( n \), if a small firm faces no entry, it remains a monopolist and earns monopolistic profits. Otherwise, the two firms engage in Bertrand competition, namely, firms compete on prices and the firm with the lower marginal cost sets the limit price at the other firm’s marginal cost.\(^4\) All payoffs realize at the end of the period.

**Information structure.** Disclosure in this economy is mandatory. Disclosure regulations are

---

\(^3\)Because there is an infinite number of small firms/local markets, the probability of two large firms entering the same market is zero.

\(^4\)Appendix B.1 considers an alternative case where the large entrant develops *completely differentiated* products, such that firms do not compete and are monopolists of their products.
assumed to be already in place exogenously in order to protect different stakeholders. This model focuses on analyzing the implications of disclosure requirement imposed on relatively small firms. Therefore, only small firms’ disclosure is explicitly modeled. This modeling choice can be motivated by the fact that large firms with more dispersed shareholder base typically have stronger incentives of voluntary disclosure, or they may opt to be subject to high disclosure requirement by being public firms [39, 40]. I assume that small firms’ disclosure $d_n \sim N(\omega_n, \tau_d^{-1})$, reveals information about their respective market condition $\omega_n$, where $\tau_d$ is the disclosure precision required per disclosure regulation.

Large firms do not directly observe the local market conditions, rather they rely on a public signal $\chi_n \sim N(\omega_n, \tau_\chi^{-1})$, and disclosure by small firms to learn about market condition $\omega_n$. The public signal $\chi_n$ and small firm’s disclosure $d_n$ are conditionally independent. Given a public signal $\chi_n$, and small firm’s disclosure $d_n$, a large firm active in market $n$ updates its belief about the market condition to:

$$E_b[\omega_n | \chi_n, d_n] = \frac{\tau_\chi \chi_n + \tau_d d_n}{\tau_\omega + \tau_\chi + \tau_d} \equiv \hat{\omega}_n,$$

(2.5)

where subscript $b$ indicates that the expectation is taken by the large firm, and $\hat{\omega}_n$ denotes its posterior about market condition $\omega_n$. The large firm’s residual uncertainty about $\omega_n$ is the conditional variance:

$$\text{Var}_b[\omega_n | \chi_n, d_n] = \frac{1}{\tau_\omega + \tau_\chi + \tau_d} \equiv \hat{\sigma}^2.$$

(2.6)

The residual uncertainty depends on the precision of the public signal $\tau_\chi$ and the precision of
disclosure by small firms $\tau_d$. It can thus be written as a function of $\tau_x$ and $\tau_d$, i.e., $\hat{\sigma}^2(\tau_x, \tau_d)$, but for the most part of the paper, $\hat{\sigma}^2$ is used as a shorthand notation. Without learning from the disclosure by small firms, large firms’ residual uncertainty is $\hat{\sigma}_o^2 = \hat{\sigma}^2(\tau_x, 0) = \frac{1}{\tau_x + \tau_y}$.

**Product market competition.** In a local market $n$, equilibrium product price may be set by the small firm or the large firm, depending on the competition structure and the relative marginal costs of the two types of firms.

A small firm produces products at a marginal cost $\phi > 0$. If a small firm $s$ is a monopolist, it sets monopoly price $p_s^m$, to maximize:

$$
\max_{p_s^m} \quad p_s^m x(p_s^m, A) - \phi x(p_s^m, A),
$$

(2.7)

where $x(p, A)$ is the quantity demanded by the final goods firm given by equation (2.3).\(^5\) The optimal monopoly price is a constant markup over the marginal cost: $p_s^m = \frac{\phi}{\theta}$, which results in a monopoly quantity of $x_s^m = x(p_s^m, A)$, and a monopoly profit of $\pi_s^m$.

For a large firm entering market $n$, it has two potential cost realizations: $\phi_b \in \{\phi^L, \phi^H\}$, where $\phi^L < \phi < \phi^H$. The large firm’s adaptation action $a_{bn}$ determines how likely it is going to draw the low marginal cost. Specifically, with probability $z_n(\omega_n, a_{bn}) = 1 - \alpha(\omega_n - a_{bn})^2$, the large firm’s marginal cost is $\phi^L$; and with probability $1 - z_n(\omega_n, a_{bn}) = \alpha(\omega_n - a_{bn})^2$, the large firm’s cost is $\phi^H$, where $\alpha > 0$ represents the importance of information in firms’ strategic decisions. Here the probability itself is a random variable that depends on the realization of market condition shock $\omega_n$ and the large firm’s strategic action $a_{bn}$. The more closely the large firm’s strategic action matches the market condition, the more likely it will obtain a lower cost and secure a cost advantage over the small firm. Thus, a large firm’s optimal action is its best guess of the local market condition, given its information set, namely,

$$
a_{bn} = \mathbb{E}_b[\omega_n \mid \chi_n, d_n] = \hat{\omega}_n.
$$

(2.8)

---

\(^5\)Subscript $j$ that denotes individual product variety is suppressed, because a firm charges the same price for all its products.
If the large firm has perfect information about the local market condition, it obtains the low cost $\phi_L$ with certainty, which implies that the large firms are more efficient in production if they know as much as the small firms.

When the two firms engage in Bertrand competition, the firm with lower marginal cost is able to set a limit price for market $n$, and the firm with higher marginal cost earns zero profit. If the large firm has marginal cost $\phi_b = \phi_L$, it sets price at either the small firm’s cost or the monopoly price, whichever is lower, i.e., $p^c_b = \min\{\phi, \phi_L\}$, produces a quantity $x^c_b = x(p^c_b, A)$, and earns a profit $\pi^c_b$; if the large firm has marginal cost $\phi_b = \phi_H$, the small firm similarly sets price at $p^c_s = \min\{\phi_H, \phi\}$, produces a quantity $x^c_s = x(p^c_s, A)$, and earns a profit $\pi^c_s$.

In summary, the equilibrium prices and quantities in different markets are as follows,

if local monopolist: \[ p^m_s = \frac{\phi}{\theta}, \quad x^m_s = \left(\frac{\theta^2}{\phi}\right)^{\frac{1}{1-\theta}} A, \quad \pi^m_s = (p^m_s - \phi)x^m_s \]  \hspace{1cm} (2.9)

if large firm has $\phi_L$: \[ p^c_b = \min\{\phi, \frac{\phi_L}{\theta}\}, \quad x^c_b = \left(\frac{\theta}{p^c_b}\right)^{\frac{1}{1-\theta}} A, \quad \pi^c_b = (p^c_b - \phi_L)x^c_b \]  \hspace{1cm} (2.10)

if large firm has $\phi_H$: \[ p^c_s = \min\{\phi_H, \frac{\phi}{\theta}\}, \quad x^c_s = \left(\frac{\theta}{p^c_s}\right)^{\frac{1}{1-\theta}} A, \quad \pi^c_s = (p^c_s - \phi)x^c_s. \]  \hspace{1cm} (2.11)

**Innovation.** At time $t2$ intermediate goods firms innovate by developing new product varieties. I call one product variety a unit of innovation. Firms hire researchers for innovation, which is of fixed supply $L = 1$. For each $l$ unit of labor employed, $l^\gamma$ new product varieties can be developed, where $\gamma \in (0, 1)$. Firms need to compensate the labor at an equilibrium wage rate of $W$. A firm $i$ in market $n$ thus solves the following maximization problem:

\[ \max_{l_i} E[\pi_{in}] l^\gamma_i - W l_i, \]  \hspace{1cm} (2.12)

where $E[\pi_{in}]$ denotes the expected gross profit from developing a new product and it is a probability-weighted average of the potential product market profits as specified in equation (2.9) to (2.11). Correspondingly, $E[\pi_{in}] l^\gamma - W l$ is the net profit from innovation. This innovation function is consistent with the general macro-innovation literature such as [41] and [17], except that here the role
of existing products and hence the role of innovation spillover is abstracted away.\footnote{Positive spillovers are an important feature of innovations, however it does not directly interact with the effect of disclosure on competition and reallocation here.} Firms’ optimal innovation decision solves the maximization problem in (2.12),

\[
 l_i = \left( \frac{E[\pi_{in}]}{W} \right)^{\frac{1}{1-\gamma}},
\]

(2.13)

which indicates that firms’ innovation motive is driven by the expected gross profit they earn from the product market, \(E[\pi_i]\); and the innovation cost that is determined by the equilibrium wage paid to researchers, \(W\). The expected gross profit directly drives firms’ innovation demand (direct effect); the skilled labor wage is the \textit{indirect} general equilibrium effect that captures the impact of expansion or contraction of all firms in the economy (indirect effect). The equilibrium wage \(W\) solves the labor market clearing condition,

\[
\int_n l_s \, dn + \mu \int_n l_b \, dn = L = 1 \quad \Leftrightarrow \quad W = \gamma \left( E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}} \right)^{1-\gamma}.
\]

(2.15)

The value functions are obtained by substituting equation (2.15) into equation (2.12), together with (2.13). The respective ex-ante firm values are the net expected returns on innovation, given equilibrium innovation decision (2.13) and equilibrium wage (2.15):

\[
V_s = E[\pi_s]^{\frac{1}{1-\gamma}} l_s^\gamma - Wl_s = (1 - \gamma) E[\pi_s]^{\frac{1}{1-\gamma}} (E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}})^{1-\gamma},
\]

(2.16)

\[
V_b = E[\pi_b]^{\frac{1}{1-\gamma}} l_b^\gamma - Wl_b = (1 - \gamma) E[\pi_b]^{\frac{1}{1-\gamma}} (E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}})^{1-\gamma}.
\]

(2.17)

\section*{2.3 Representative Household}

A representative household purchases final goods for consumption \(C\) from the final goods firm and derives utility from the consumption. The representative household owns one unit of production input resource \(A = 1\) and is endowed with one unit of skilled labor \(L = 1\). They earn
rent $P^A A$ from supplying input $A$ to the final goods producer and earn wage $WL$ from providing labor to intermediate goods firms. Moreover, the representative household is the owner of all firms in the economy and receives total dividends $D$ from the firms.\footnote{The dividends are the net profits made by all intermediate goods firms. The final goods firm yields zero profit in equilibrium because of its constant return to scale technology.} Because this is a one-period model without aggregate uncertainty, the exact form of the utility function does not matter as long as it is monotonically increasing in $C$. Effectively, the household only needs to fulfill the budget constraint:

$$C = P^A A + D + WL.$$  \hspace{1cm} (2.18)

### 2.4 Equilibrium

**Definition 1** An equilibrium of this economy is a collection of equilibrium values:

$$\{\{p_n\}_n, \{x_n\}_n, \{a_{bn}\}_n, \{l_{sn}\}_n, \{l_{bn}\}_n, P^A, W, A, C, X\}$$

such that in each market $n$, (i) equilibrium prices and quantities are determined by equations (2.9) to (2.11); (ii) optimal action $a_{bn}$ is taken according to (2.8); (iii) taking wage $W$ and expected gross profit as given, research labor demand $l_s$ and $l_{bn}$ satisfy (2.13) for small and large firm respectively; (iv) wage $W$ is such that the labor market clears; (v) taking input price $P^A$ as given, final goods producer optimally choose quantity of input $A$, and $P^A$ is such that the input market clears; (vi) the market for final goods clears, i.e., (2.4) holds.

### 2.5 Reallocation Effects

This section examines how skilled labor demand, innovation expenditure and firm value change for each group of small and large firms as more disclosure is required from small firms. Firms’ innovation incentives are driven by (i) the expected gross profits from the product market and (ii) the innovation cost as indicated by equation (2.13). The impact of disclosure regulation is thus through both of these two channels.
First, firms’ expected gross profits are affected through the learning channel. It reflects two sides of the same coin: what is a learning benefit for large firms constitutes a business stealing cost for small firms. Large firm’s expected gross profit from developing a new product in market \( n \) is

\[
E[\pi_b] = E\left[z_n(\omega_n, a_{bn})\pi_b^c\right],
\]

\[
= E[1 - \alpha(\omega_n - \hat{\omega}_n)^2]\pi_b^c,
\]

\[
= \left(1 - \alpha \hat{\sigma}^2\right)\pi_b^c,
\]

(2.19)

(2.20)

(2.21)

where \( z_n(\omega_n, a_{bn}) \), the probability that the large firm achieves a low cost, is itself a random variable. It depends on how well the large firm’s strategic action \( a_{bn} \) matches the true market condition. Equation (2.20) spells out the probability explicitly. Ex ante, the average probability is determined by the quality of the large firm’s best estimate. In particular, \( E[(\omega_n - \hat{\omega}_n)^2] \) is the average squared distance between the true market condition \( \omega_n \) and the large firm’s estimate of it, \( \hat{\omega}_n \), which measures the residual uncertainty \( \hat{\sigma}^2 \) from the perspective of the large firm. Large firms thus base their innovation decision on the expectation of their information quality. Similarly, for a small firm, its expected gross profit from developing a new product is

\[
E[\pi_s] = E\left[(1 - \mu)\pi_s^m + \mu(1 - z_n(\omega_n, a_{bn}))\pi_s^c\right],
\]

\[
= (1 - \mu)\pi_s^m + \mu \alpha \hat{\sigma}^2 \pi_s^c.
\]

(2.22)

(2.23)

Because there is a measure \( \mu \) of large entrants, by the Law of Large Numbers, with probability \( \mu \) the small firm faces entry of a large firm. In the ensuing duopoly competition, with probability \( 1 - z_n(\omega_n, a_{bn}) \) the small firm has a cost advantage and earns profit \( \pi_s^c \). With probability \( 1 - \mu \), there is no entry, the small firm remains a monopolist and earns monopoly profit \( \pi_s^m \).

As small firms’ disclosure quality increases, the large firms can better adjust to local market conditions and are more likely to obtain a lower cost and out-compete the small firms. This increase in expected gross profit leads to more innovation by the large firms as per optimal innovation
function (2.13). On the flip side, small firms’ expected gross profit from developing a new product decreases as a result of the enhanced competition from large firms, which leads to a decrease in small firms’ innovation.

Second, there is an externality channel affecting firms’ innovation costs. This stems from firms competing for the same pool of skilled labor. On one hand, the increase in large firms’ innovation puts an upward pressure on the equilibrium wage $W$ of the researchers, and makes it more expensive for the small firms to innovate. On the other hand, the contraction of small firms reduces the equilibrium wage, which further increases large firms’ net profit from innovation. This externality channel is an equilibrium consequence resulting from the learning channel. The externality channel is important as it implies that a redistribution from small to large firms will take place even without direct product market competition between the two types of firms, i.e., no business stealing cost for the small firms. Even if the large firms use small-firm disclosure to develop completely differentiated products and do not encroach on small firms’ gross profit margin, small firms still bear an equilibrium cost and will innovate less as a result. Appendix B.1 provides detailed analysis of the case with no business stealing.

Combining both the learning and the externality channel, small firms bear a cost of enhanced disclosure through both the business stealing effect and the upward pressure on innovation cost. It therefore follows in a straightforward manner that small firms’ skilled labor demand, innovation expenditure and firm value all decrease in the disclosure quality. The opposite is true for large firms. Proposition 1 summarizes these results.

**Proposition 1 (i)** Large firms’ skilled labor demand, innovation expenditure, and firm value are all monotonically increasing in small firms’ disclosure, (ii) whereas the converse holds for small firms. That is, $\forall \tau_d$,

$$\frac{d l_b}{d \tau_d} > 0, \quad \frac{d W l_b}{d \tau_d} > 0, \quad \frac{d V_b}{d \tau_d} > 0$$

$$\frac{d l_s}{d \tau_d} < 0, \quad \frac{d W l_s}{d \tau_d} < 0, \quad \frac{d V_s}{d \tau_d} < 0.$$
2.6 Aggregate Comparative Statics

Given the increase in large firms’ innovation and decrease in small firms’ innovation, the question of central importance is how the aggregate levels of innovation and welfare change change with the small firms’ disclosure quality. Proposition 2 summarizes how total innovation expenditure and total firm value in the economy change as small firms are mandated to disclose more precise information. \( WL \) is the total innovation expenditure, because firms pay an equilibrium wage \( W \), to the skilled labor they hire for innovation. Total firm value in the economy is expressed by integrating over all individual firm values.

**Proposition 2** Given a combination of parameter values \( \{ \phi, \phi^L, \phi^H \} \), such that \( \phi^L < \phi < \phi^H \), one of the following cases hold.

a. If \( \phi^L \) is lower than some threshold \( \underline{\phi}^L \), total innovation expenditure and total firm value are monotonically increasing in small firms’ disclosure precision.

b. If \( \phi^L \) is higher than some threshold \( \bar{\phi}^L \), total innovation expenditure and total firm value are decreasing in small firms’ disclosure precision.

c. If \( \underline{\phi}^L < \phi^L < \bar{\phi}^L \), total innovation expenditure and total firm value first decrease then increase in small firms’ disclosure.

\( \underline{\phi}^L \) and \( \bar{\phi}^L \) are functions of \( \phi \) and \( \phi^H \); the exact conditions that pin down \( \underline{\phi}^L \) and \( \bar{\phi}^L \) are provided in the Appendix. Proposition 2 expresses the thresholds in terms of an exogenous parameter \( \phi^L \). But essentially, what determines the different cases is the relation of \( \phi^L \), \( \phi^H \) and \( \phi \) jointly, because these three parameters determine the relative margin and thus the relative innovation incentives of the large and small firms. Proposition 2 highlights that whether enhanced small-firm disclosure increases total innovation expenditure and firm value depends on whether disclosure shifts innovation to the group of firms with higher expected gross profits. Proposition 2 part a discusses the case
when large firms’ have the higher expected profits. In particular, if large firms’ products deliver higher payoff than small firms even without learning from disclosure by small firms, then increasing $\tau_d$ always shifts innovation to the group of firms that have higher expected gross profit from innovation and thus are more motivated to innovate. This will be the case if public information is very precise, i.e., $\hat{\sigma}^2_o = \hat{\sigma}^2(\tau, \tau_d = 0)$ is low, or if the potential cost advantage of the large firm is large, i.e., $\phi^L$ is low. In these cases, total innovation expenditure and total firm value increase with better small-firm disclosure.

On the other hand, large firms can have very low expected gross profit from innovation even if they learn the market condition perfectly ($\tau_d \to +\infty$). This is the case when large firms have little potential cost advantage compared to small firms, i.e., $\phi^L$ is only marginally lower than $\phi$. Large firms earn thin profits in the Bertrand competition and thus do not have strong incentive to develop new products. Meanwhile, if small firms’ expected profit from innovation is relatively high, then enhanced small firm disclosure will decrease total innovation expenditure and firm value, because it shifts innovation to the less profitable firms who have less incentive to innovate.

In this model, gross profitability of the two groups of firms is endogenous to small firms’ disclosure quality. Better local disclosure helps large firms to become more profitable in the product market. It may be the case that disclosure is initially shifting innovation to the less profitable large firms, and hence leads to lower total innovation expenditure and firm value; but as disclosure quality further increases, it eventually renders large firms more profitable, and thus increases total innovation expenditure and firm value. This is the case discussed in Proposition 2, part $c$.

2.7 Welfare Analysis

While total innovation expenditure and total firm value may be measures that a regulator would be interested in, the ultimate measure for welfare in this economy is total consumption $C$, as it is the source of household utility. Consumption $C$ can be decomposed as contributions from small and large firms separately:
\[ C = (1 - \mu)((x^m_x) - \phi x^m_x)l^m_x + \mu \alpha \hat{\sigma}^2((x^c_x) - \phi x^c_x)l^c_x + \mu \left(1 - \alpha \hat{\sigma}^2\right) ((x^c_b) - \phi^L_x x^c_b)l^c_b \] (2.26)

from local monopolists from small firms, \( \phi_b = \phi^H \) from large firms, \( \phi_b = \phi^L \)

\[ = (1 - \mu)c^m_x l^m_x + \mu \alpha \hat{\sigma}^2 c^c_x l^c_x + \mu \left(1 - \alpha \hat{\sigma}^2\right) c^c_b l^c_b. \] (2.27)

Here \( l^r_i \) is the number of innovations developed by firm \( i \). \( c^m_i \) or \( c^c_i \) is the amount of consumption generated by each product variety of firm \( i \) under monopoly (\( m \)) or Bertrand competition (\( c \)), which I refer to as consumption intensities of an innovation. A measure \( 1 - \mu \) of small firms remain local monopolists and each of their product variety has consumption intensity \( c^m_x \). For the rest of the local markets (of measure \( \mu \)), there is competition. \( \alpha \hat{\sigma}^2 \) proportion of them have the small firms winning the competition and delivering \( c^c_x \) per unit of innovation; \( 1 - \alpha \hat{\sigma}^2 \) proportion of them have the large firms winning the competition and delivering \( c^c_b \) per unit of innovation.

As is evident from equation 2.27, both the total number of innovations and the consumption intensity play an important role in determining the welfare. Lemma 1 ranks the consumption intensities.

**Lemma 1** Consumption delivered per unit of innovation ranks as follows:

\[ c^m_x \leq c^c_x \leq c^c_b. \] (2.28)

Products developed by local monopolists have lower consumption intensity compared to those developed by firms under competition. This is intuitive, as monopolists suppress production to earn monopolistic rents, which is an inefficient distortion from a welfare perspective. In markets with a large firm entrant, competition drives down prices and firms produce more efficient quantities. Moreover, among competing firms, large firms’ products have higher consumption intensity, because given that large firms are producing, they must have lower marginal cost and hence be more
efficient.

More precise disclosure by small firms reduces residual uncertainty, \(\hat{\sigma}^2\), for large firms and increases the proportion of large firms with low marginal cost and high consumption intensity. Therefore, in terms of consumption intensity, disclosure regulation is always desirable, and the lemma below follows without the need of a formal proof:

**Lemma 2** For a given level of total number of innovations, i.e., given \(l_s^r\) and \(l_b^r\), more precise disclosure by small firms increases welfare.

On the other hand, how total number of innovations changes with enhanced small-firm disclosure is not as unambiguous. The effect is generally in line with the discussion in Proposition 2, where firms’ expected profits from product market competitions are key in determining the aggregate effect. Intuitively, firms that have higher consumption intensity should also innovate more from a social welfare perspective. However, when firms are making their innovation decisions, they are motivated by their expected gross profits \(E[\pi_i]\) rather than their expected consumption intensity \(E[c_i]\). For instance, a firm \(i\) can have a low marginal cost in an absolute sense and thus have a high consumption intensity. But on a relative scale, if firm \(i\)'s marginal cost is only slightly lower than its competitor’s, it will have little incentive to innovate because of the thin margin. This potential wedge plays a key role in determining the welfare impact of small firms’ disclosure. The following proposition identifies extreme cases for when welfare is increasing or decreasing as disclosure by small firms becomes more precise.

**Proposition 3**

- **a.** If \(\phi^L\) is sufficiently close to \(\phi\), welfare is decreasing in small firms’ disclosure precision.

- **b.** If \(\phi^H\) is sufficiently close to \(\phi\), and \(\phi^L\) is sufficiently small, welfare is increasing in small firms’ disclosure precision.

In Proposition 3 part a the potential cost advantage for large firms (in the case \(\phi_L\) is realized) is small. The marginal increase in large firms’ innovation in response to more disclosure by small
firms then is modest. Production is redistributed from more innovative small firms to less innovative large firms. This is a case where large firms take over small firms’ markets, but they do not innovate sufficiently in those markets, which leads to an overall decrease in total welfare. Figure 2.2(a) depicts such a case where welfare is decreasing with small firms’ disclosure precision. On the contrary, if small firms’ profit margin is thin when competing with a large firm with high cost \( \phi^H \) (Proposition 3 part b), reallocation to large firms has a positive effect on welfare. In this case, large firms have sufficient innovation incentives, and they also produce these products very efficiently (low \( \phi^L \)). Figure 2.2(b) presents a case where welfare is increasing with small firms’ disclosure precision. Finally, similar to the aggregate innovation expenditure case, total welfare can have a non-monotonic relation in the quality of small firms’ disclosure, as shown in Figure 2.2(c).

In summary, because innovation is an integral component of welfare, the effects of disclosure on welfare has similar patterns as those on aggregate innovation. However, innovations of large firms have higher consumption intensity. This potential production efficiency of large firms always positively contributes to welfare, which mitigates negative innovation consequences in some cases.
Chapter 3: Estimation and Quantitative Analysis

The model analysis shows that depending on parameter values the relation between small firms’ disclosure quality and total innovation or welfare can vary. I estimate the model with European data to pin down whether small-firm disclosure enhances or hampers total innovation and welfare in the current disclosure regime of the EU.

3.1 Institutional Background

In the EU, limited liability firms (both public and private) are required to publicly disclose a set of financial statements annually by the EU Accounting Directives. The required disclosure includes a balance sheet, an income statement and a management report that discusses firms’ economic and competitive environment, strategy, risk and opportunities. These requirements are enforced by each national government. Typically, firms below certain size thresholds are allowed to report an abbreviated version of the financial statements. The size-related thresholds include combinations of number of employees, total assets and sales. The EU sets maximum exemption thresholds that all countries must follow, which were around 50 employees, 4 million Euros in total assets and 8 million Euros in sales for most of the sample period. Within the confines of the restrictions set by the EU, each country can set more restrictive exemption thresholds, which results in variation in the exemption thresholds across European countries. Moreover, the country specific exemption thresholds are applicable to all industries in a country, generating variation in the proportion of firms that are subject to the disclosure regulation across industries. I take advantage of the reporting exemption thresholds to construct a measure that reflects the intensity of reporting regulation, and use its relationship with firms’ innovation behavior as the key empirical context to estimate the model.
### 3.2 Data

The data are primarily from two sources: data of companies’ innovation activities across Europe from Eurostat’s Community Innovation Survey (CIS), and firm financial data from BvD’s Orbis database. The Community Innovation Survey collects information about firms’ product/process innovation, innovation expenditure and other innovation-related information. The survey is conducted on a biennial basis and I use data from all available survey years spanning from 2000 to 2018, covering 17 countries.\(^1\) This data source has several advantages. The implementation of the Community Innovation Survey is required by EU regulation [42]; the survey questions are harmonized across countries and years; the participating firms are anonymized; and the access of the data is restricted to scientific use of researchers from recognized institutions. These features mitigate concerns over issues such as sample representativeness, comparability, self-selection and firms’ misreporting incentives. The primary variable I use from Community Innovation Survey is *innovation expenditure*, which is generally defined as in-house or contracted research and development (R&D) expenditures including labor costs, services and materials purchased and capital expenditures [43].\(^2\) Additionally, I complement the confidential CIS data with Eurostat’s public database for science, technology and innovation to get national statistics on headcounts of R&D personnel by firm size (number of employees).

I obtain firm financial data from BvD’s Orbis database for the same time period and countries. The financial data is used to construct non-innovation variables such as profit margins and extents of disclosure regulations. In particular, I construct a country-industry level measure of disclosure regulation intensity using the share of firms above a country’s reporting thresholds in that country-

---


\(^2\)R&D activities are defined in the CIS as “creative and systematic work undertaken in order to increase the stock of knowledge – including knowledge of humankind, culture and society – and to devise new applications of available knowledge”; an innovation is defined as “a new or improved product or process (or combination thereof) that differs significantly from the unit’s previous products or processes and that has been made available to potential users (product) or brought into use by the unit (process).” [43].
industry, following [11].

3.3 Validation of Model Predictions

Before estimating the model parameters, I present evidence consistent with the model prediction that indeed large (small) firms’ innovation is increasing (decreasing) with the intensity of small-firm disclosure. To this end, I run the following firm-level regression:

\[
Y_{ijt} = \beta_1 \text{Large}_{ijt} \times \text{RegIntensity}_{jt} + \beta_2 \text{RegIntensity}_{jt} \\
+ \beta_3 \text{Large}_{ijt} + FE + \varepsilon_{ijt},
\]

(3.1)

where \(Y_{ijt}\) is either \(\ln(\text{InnovExpense}_{ijt})\), the logarithm of innovation expenditure of firm \(i\), in industry \(j\) and year \(t\), or an indicator variable that equals one if the firm has positive innovation expenditure. \(\text{Large}_{ijt}\) is a dummy variable that takes the value one if firm \(i\) has more than 250 employees in year \(t\). \(\text{RegIntensity}_{jt}\) measures the share of firms above their country-level reporting thresholds in industry \(j\) and year \(t\). Country-year fixed effects and industry-year fixed effects are included to control for systematic differences across countries and industries.

The measure \(\text{RegIntensity}_{jt}\) captures the extent of mandatory disclosure regulation in a specific country-industry. It is constructed based on a simulated instrument approach, following [11]. Essentially, a standardized country-year-invariant firm-size distribution is used to calculate the share of firms above the reporting thresholds, instead of using the actual firm-size distribution. This approach mitigates various endogeneity concerns such as reverse causality or correlated omitted variables [23, 24]. The regression coefficients are thus plausibly causal estimates of firms’ innovation elasticities, and they can be interpreted as by how much firms’ expenditures increase/decrease.

---

3I cannot link firm level financial information in Orbis to the CIS firms, because of the anonymity of the CIS data. However, this issue is not of great concern, because (i) the variable regulation intensity is constructed on the country-industry level and can be matched with firms in CIS; (ii) both Orbis and CIS database have extensive coverage and good representation, the average statistics calculated from the two data sets can roughly represent the same underlying population of firms [44].

4In the Community Innovation Survey, firms are put into different size buckets by number of employees. In the CIS categorization, firms with 10-50 employees are labeled as small, 50-250 as medium, 250-500 as large, and 500+ as huge. But the huge category is only available since year 2010.
with an increase in mandatory disclosure regulation intensity. Importantly, because the exemption thresholds are size based, an increase in the share means subjecting more smaller firms to the full reporting requirements, which closely aligns with the policy change my model considers.

I also consider the following alternative specification, which further divides firms with fewer than 250 employees into two groups:

\[
Y_{ijt} = \beta_1 \text{Large}_{ijt} \times \text{RegIntensity}_{jt} + \beta_2 \text{Below50}_{ijt} \times \text{RegIntensity}_{jt} + \beta_3 \text{RegIntensity}_{jt} + \beta_4 \text{Large}_{ijt} + \beta_5 \text{Below50}_{ijt} + FE + \epsilon_{ijt},
\] (3.2)

where the additional variable \text{Below50}_{ijt} is a dummy variable that equals one if firm \(i\) has fewer than 50 employees. This specification highlights that the impact of mandatory disclosure may be especially pronounced for firms in the smallest size group.

Table A.3 presents the regression results. Consistent with model predictions, subjecting more smaller firms to mandatory disclosure is significantly positively associated with large firms’ innovation expenditure (column 1) and large firms’ propensity to innovate (column 3). For firms with fewer than 250 employees, the coefficients of \text{RegIntensity} in columns 1 and 3 are negative—consistent with expectation, but not significant. However, decomposing the size group further, I find that the magnitudes of effects become larger and statistically significant. Firms with fewer than 50 employees reduce innovation expenditure significantly and are significantly less likely to innovate with more extensive disclosure requirements (coefficients of \text{RegIntensity} \times \text{Below50} in column 2 and 4).

In terms of economic magnitude, the estimates suggest that a 10% increase (relative to its mean) in the share of disclosing firms increases large firms’ innovation expenditure by around 2%, and increases large firms’ likelihood of innovation by around 0.4%. Roughly, a 10% increase (relative to its mean) increase in the share of disclosing firms decreases innovation expenditure of firms with fewer than 250 employees by around 0.95%, and decreases their likelihood of innovation by around 0.14%.\(^5\) These estimates are also useful later as moments to identify information-related

---

\(^5\) Increasing the share of reporting firms by 10 percent corresponds to a 2 percentage points increase in the share
parameters in the structural estimation exercise.

3.4 Estimation Strategy

3.4.1 Parameters

There are 10 parameters to be estimated/calibrated: \{\theta, \tau_\omega, \tau_\chi, \tau_d, \alpha, \phi, \phi^L, \phi^H, \gamma, \mu\}. I calibrate generic parameters that are common in the literature from prior literature. Following the endogenous growth literature [e.g., 45, 41, 17], the production share of intermediate goods, \theta, is set to 0.5, and the elasticity of innovation expenditure, \gamma, is set to 0.7. Without loss of generality, I normalize the marginal cost of production of small firms, \phi, such that their monopoly profit of each innovation is equal to one, in accordance with [17].

For the information-related parameters, \{\tau_\omega, \tau_\chi, \tau_d, \alpha\}, the model does not allow me to separately identify precision of the prior (of the market condition shock) \tau_\omega, and precision of the public information \tau_\chi. I therefore combine these two parameters by estimating \tau_o = 1/\hat{\sigma}_o^2 = \tau_\omega + \tau_\chi, which represents the conditional variance given all available information except for small-firm disclosure. Moreover, I normalize \alpha to 1, since the level of the precision parameters are only pinned down jointly with \alpha, the significance of information precision. To this end, I note that the levels of the precision parameters do not have a standalone economic meaning, the precision of small-firm disclosure \tau_d can only be interpreted in relation to the precision of prior information \tau_o. For example, \(1 - \frac{\hat{\sigma}^2}{\hat{\sigma}_o^2}\), is the reduction of uncertainty about the market conditions given small firm-disclosure, where \(\hat{\sigma}^2 = \frac{1}{\tau_o + \tau_d}\), is the posterior variance after small-firm disclosure.

I estimate the remaining five parameters \{\tau_o, \tau_d, \phi^L, \phi^H, \mu\}, using the simulated method of moments (SMM) approach [46, 47]. Formally, I search over the parameter space \((\tau_o, \tau_d, \phi^L, \phi^H, \mu)\) to find the parameter combination that minimizes the weighted distance between model-generated of reporting firms (0.21 \times 10\% = 2\%), which increases large firms’ innovation expenditure by \((1.505 - 0.477) \times 2\% = 2.06\%\) (column 1) or \((1.216 - 0.200) \times 2\% = 2.03\%\) (column 2); increases large firms’ innovation propensity by \((0.128 - 0.0238) \times 2\% = 0.39\%\) (column 3) or \((0.104 - 0.00067) \times 2\% = 0.39\%\) (column 4), where 0.53 is the average propensity to innovate among large firms. It decreases small firms’ innovation expenditure by \(-0.477 \times 2\% = 0.95\%\) and decreases small firms’ innovation propensity by \(-0.0238 \times 2\% / 0.34 = 0.14\%\), where 0.34 is the average propensity to innovate among small firms.
statistics (e.g., mean, variance, regression coefficients) and the corresponding empirical statistics. I use the inverse of the covariance matrix of the empirical moments as the weight matrix [48].

### 3.4.2 Moments

The SMM procedure targets seven moments: the mean and the variance of margins for large and small firms, respectively, share of all R&D labor in an industry who are employed by large firms, and innovation elasticity with respect to small-firm disclosure regulation for large and small firms, respectively. To obtain the model moments, I simulate 100,000 local markets (which corresponds to 100,000 small firms and $\mu \times 100,000$ large entrants), and compute the implied statistics given a set of parameters. To obtain the data moments, I categorize a firm as “large" if it has more than 250 employees and “small" if it has fewer than 250 employees, and compute the corresponding moments. I use this dichotomy because the 250 employees cutoff is widely used by the European national statistical offices and the Community Innovation Survey. More importantly, most disclosure regulations also categorize firms with more than 250 employees as large, and these firms are always subject to the full disclosure requirements. In the CIS sample, around 10% of the firms have more than 250 employees.

The first five moments are straightforward statistics. I elaborate on the last two moments. The innovation elasticity refers to how much large (small) firms’ innovation expenditure increases (decreases) as small firms become subject to a higher level of mandatory disclosure requirements. In the model, this is captured by the average log innovation expenditure change induced by a particular $\tau_d$ compared with the case where $\tau_d = 0$. That is,

$$InnovElast_i = E[\ln(l_i) | \tau_d = \hat{\tau}_d] - E[\ln(l_i) | \tau_d = 0],$$

where $i \in \{s, b\}$. The data counterparts can be obtained from the coefficients of the regression
model in equation (3.1), which I reproduce here:

\[ Y_{ijt} = \beta_1 Large_{ijt} \times RegIntensity_{jt} + \beta_2 RegIntensity_{jt} + \beta_3 Large_{ijt} + FE + \epsilon_{ijt}. \]

\( \beta_1 + \beta_2 \) is interpreted as the percentage change of large firms’ innovation expenditure, when increasing the share of firms subjecting to the disclosure mandate from 0% to 100%. Similarly, \( \beta_2 \) is the small firms’ percentage change in innovation expenditure from no-firm to all-firm disclosure. I consider a 10% increase (relative to the mean) in the share of firms subjecting to the reporting mandates, which is equivalent to 3 percentage point increase in the share of reporting firms. A 3 percentage point increase in the share of reporting firms is considered, because the average identifying variation of the regression model is around 3 percentage points. Considering a larger increase may be prone to extrapolation bias.

The standard errors of the data moments are computed by bootstrap. Specifically, I draw samples of the same size as the original sample from the data sets and recalculate the moments such as mean/variance of margins, innovation elasticities, etc., for each bootstrap sample. Standard errors are then computed based on the distribution of the recalculated moments.

3.4.3 Identification

Although all five parameters are jointly pinned down by all seven moments using SMM, some moments are more relevant to, and have a monotonic relationship with, certain parameters. The monotonic relationships are particularly helpful for identifying the parameters. This section presents these relationships by showing the comparative statics. I set the parameters to the baseline estimates and vary each parameter one at a time for each simulated model moment. The most relevant comparative statics are shown in Figure A.1.

I start with the information-related parameters. Figure A.1a–f show comparative statics of small firms’ disclosure precision, \( \tau_d \). As discussed in the model analysis, an increase in \( \tau_d \) induces
a decrease in the mean profit margin of small firms and an increase in that of large firms, because of the learning benefit for large firms and the business stealing cost for small firms. The variance of large firms’ profit margins decreases with $\tau_d$, as more precise information allows them to secure the low cost with higher probability. The opposite is true for small firms, their profit margins vary less if disclosure precision is low, because they can enjoy large/monopoly profit with higher probability. The comparative statics for innovation elasticities are also intuitive. Small (large) firms’ innovation expenditure decreases (increases) more with more precise small-firm disclosure.

The precision of prior information, $\tau_o$, plays a similar role as $\tau_d$, when it comes to profit margins. Better information helps (hurts) the profit margins of large (small) firms, regardless of whether this information is attributable to prior information or to disclosure by small firms. Therefore, Figure A.1g, A.1h are similar to Figure A.1a, A.1b. However, the effects of $\tau_o$ on innovation elasticities (Figure A.1i, A.1j) are contrary to those of $\tau_d$. With better prior information (higher $\tau_o$), additional small-firm disclosure is not as useful. As a result, small (large) firms’ innovation expenditure decreases (increases) less with more precise prior information. At the extreme, when the precision of prior information is sufficiently large, the innovation elasticities go to zero, as small-firm disclosure provides almost no additional information.

Next, I consider the marginal cost parameters $\phi^H$ and $\phi^L$. Note that the domains $\phi^H$ and $\phi^L$ are constrained by the relation $\phi^L < \phi < \phi^H$. As $\phi^H$ increases, small firms are able to charge a higher limit price, and their mean margin increases (Figure A.1k). But the mean margin has an upper bound—when $\phi^H$ is so high that the small firms are better off charging the monopoly price, they do not further increase prices. As $\phi^L$ increases, large firms bear a higher cost and thus earn a lower margin (Figure A.1m). Moreover, the share of R&D labor employed by large firms decrease with $\phi^H$ and $\phi^L$, respectively (Figure A.1l, A.1n). As large firms’ marginal costs increase, they are effectively less efficient in production, and they contract as a group by employing fewer researchers. Finally, the measure of potential large entrants, $\mu$, drives the share of researchers employed by large firms. The more potential entrants, more researchers will be employed by large firms.
3.5 Estimation Results

I first discuss the parameter estimates and their implications. Panel A of Table A.4 reports the parameter estimates from the SMM procedure. The standard errors of the parameter estimates are obtained from the asymptotic covariance matrix of the parameters.\(^6\)

The estimated parameter values \(\tau_o \approx 3.8\) and \(\tau_d \approx 1.9\) reveal relative informativeness of existing information versus small-firm disclosure. It indicates that while disclosure from small firms is informative about local market conditions, information from other sources also plays an important role in helping large firms to learn and adapt to local markets. This result is not surprising since firms have been increasingly using information technology, including artificial intelligence, to collect information and conduct market intelligence research.

The cost-related parameters can be interpreted through the lens of the production efficiency of large firms relative to small firms. \(\phi^L \approx 0.046\) implies that large firms that are successful in entering local markets, are on average 26.4% more efficient operationally than the small firms \((\phi - \phi^L \approx 0.264)\). Similarly, \(\phi^H = 0.082\) suggests that the large firms that fail to enter local markets are 31.2% less efficient in production in that particular market. The cost parameters also directly affect firms’ profit margins. The estimates imply that successful large entrants earn a profit margin of 35.9%, which means there is still prospect to earn substantial profit for large firms, and their incentive to innovate is preserved to some extent.

Lastly, the mass of potential large entrants is estimated to be 0.632, which reflects large firms’ intention to expand. Large firms seek to expand and grow in general. This estimate is consistent with the recent studies examining concentration and growth of large firms [49, 50, 51]. In particular, [49] document growth of 59.9% for the top 25% largest firms in Compustat from 2000 to 2005, in contrast to 7.3% growth for the bottom half of firms.

Panel B of Table A.4 reports the data moments targeted (together with their standard errors), the predicted values from the model, and t-statistics for the null that the data and model moment

\(^6\)Computing the asymptotic standard errors relies on computing numerical derivatives, where I move each individual parameters by 1% and calculate the estimated gradient vector.
are equal. Although almost all moment pairs, except for the last two, are statistically significantly different from each other, they are economically similar. Given the large sample sizes, standard errors are quite small, therefore even a minor deviation of the simulated moment from data moment would constitute a rejection of equality. Overall, the model replicates main features of the data well, especially for the innovation elasticity moments. Moreover, the model also reproduces a lower average margin for large firms than for small firms, consistent with what is observed in the data. Despite the overidentification of matching 7 moments with 5 parameters, the fit is good.

3.6 Counterfactuals

Using the estimated model, I conduct counterfactual analysis to assess the impact of extending mandatory disclosure regulations. I focus on a scenario where an additional 15% of firms are required to disclose full financial statements. Currently, approximately 20% of firms fall beyond exemption thresholds and are subject to disclosure requirements. This counterfactual policy entails a shift from the current requirement of 20% to a requirement of 23% of firms to disclose.\(^7\)

The estimates of \(\tau_o\) and \(\tau_d\) imply that the disclosure requirements imposed on the additional firms reduce uncertainty about the local market conditions by 33% \((1-\frac{\hat{\sigma}_d^2}{\hat{\sigma}_o^2} \approx 0.33)\). This magnitude of uncertainty reduction is comparable with the findings regarding disclosure usefulness in the public stock market. For instance, [25] finds that the market already knows 76.5% of current earnings, which can be interpreted as the revelation of earnings increases market participants’ knowledge by about 31% \((\frac{100-76.5}{76.5} \approx 0.31)\); similarly, [26] also finds that relative to the total amount of information contained in an earnings announcement, about one-third of this information is new to the public.

With the reduction of uncertainty, the disclosure requirements increase the probability that those small firms loose their monopolistic position and face strong competition by about 5.5 percentage points or 11.9%–from 0.47 to 0.52.\(^8\) These estimates are largely consistent with findings

---

\(^7\)This counterfactual experiment is built-in in the estimation procedure since the innovation elasticity is obtained by setting a particular value for \(\tau_d\). Without the innovation elasticity moments, \(\tau_d\) and \(\tau_o\) cannot be separately identified.

\(^8\)To see this, note that the probability that a small firm faces successful entry changes from \(\mu(1 - \alpha \frac{1}{\tau_o})\) to \(\mu(1 - \)
from previous studies regarding disclosure and competitive entry in different contexts. For example, [18] find that moving from the median to the 75th percentile of public firm presence in an industry increases subsequent foreign import competition by about 1.3 percentage points, which corresponds to a 10% increase relative to the median. In the setting of health care industry, [19] estimate that a quality disclosure mandate increases the probability of a rival opening a new facility nearby a low-quality incumbent by 27%—from 7.8% to about 10%. My estimates predict a higher baseline probability of competition, because my framework considers a general form of entry, which can include both domestic and foreign competitors, new sales in a market with or without a new local establishment, and entry in the form of M&A.

Panel (a) in Figure A.2 shows the probability of successful entry for varying disclosure precisions, with the circle indicating the baseline estimate. Intuitively, more precise disclosure leads to more successful entry. This relationship is concave, because as information gets more precise an additional unit of precision becomes less valuable. Similar concave relation between firm disclosure and competitive entry is also documented in [18].

Panel (b) and (c) of Figure A.2 show that requiring more disclosure from small firms increases aggregate welfare but decreases aggregate innovation (total product varieties). This disparity arises because on one hand, large firms earn lower profit margin and thus have less incentive to innovate, which leads to fewer total product varieties. On the other hand, the successful entrants produce with a lower marginal cost, and they charge a lower price due to the competition effect, resulting in higher consumer welfare for each product variety. My estimates show that in terms of aggregate welfare, the second effect dominates. Subjecting 15% more firms to full disclosure requirements thus decreases aggregate innovation by -0.26%, but increases welfare by about 1%.

In summary, the counterfactual analysis illustrates the different channels through which disclosure regulation impacts welfare. The insight that the disclosure regulation reduces total product varieties but increases consumption intensity is more informative than the exact point estimates of welfare change. Here the definition of welfare is total consumption rather than total household

\[ \alpha \frac{1}{\tau_o + \tau_d}. \]
utility derived from consumption. An alternative welfare measure that incorporates love for variety may indicate a welfare decrease induced by the reporting mandates.
Conclusion

In this paper, I examine the innovation and welfare impacts of extending mandatory disclosure to smaller firms in a general equilibrium model. Such extensive disclosure regulations are in place in different jurisdictions[2]. The model provides a framework to study the mechanism through which disclosure regulation jointly affects the competitive landscape, innovation incentives and aggregate outcomes. In the model, large firms can use small firms’ disclosure to compete with them in the local markets. Whether mandatory disclosure helps or hampers aggregate innovation and welfare depends on whether the large firms, who take over small firms’ markets, earn sufficiently large profits on their products and have sufficient incentives to innovate. I use the model to evaluate the consequences of subjecting more firms to reporting mandates in Europe. Specifically, I estimate the model using simulated method of moments on detailed innovation data from the Community Innovation Survey and the EU financial reporting regulations. The model fits the key moments from data, and the estimates are consistent with the findings in the previous literature. The estimated model implies that the reporting mandates in Europe leads to fewer total product varieties (a decrease in total innovation), because innovation is reallocated to large firms that have a relatively low profit margin and thus innovate less. But large firms produce each product variety more efficiently. On the whole, the disclosure regulation in the EU increases total consumption, as a measure of welfare, by about 1%. There are several caveats in the paper due to my modeling choices. To focus the analysis on the competitive consequences of disclosure regulation, I leave out information asymmetry between managers and shareholders that could give rise to adverse selection and other agency issues.
Moreover, with a one period model, I do not model the long run consequence of the disclosure induced reallocation. There are potential welfare costs associated with the expansion of large firms, as concentration can come with issues such as anti-trust, regulatory capture or decrease in business dynamism. For instance, [52] find that larger firms are incentivized toward political connections to preempt competition. Results from this paper should be considered jointly with these aspects to have a more comprehensive evaluation of the disclosure regulation impacts. The mechanism elucidated in this paper is also applicable to other settings where large firms may potentially obtain more information about niche markets. The increasingly available tracking/satellite data combined with large firms’ increasing capability in exploiting data assets constitutes a similar situation as mandating more precise disclosure from small firms [53, 54]. My paper sheds light on a potential mechanism of how such economic trends affect resource allocation and aggregate innovation of the economy.
References


Appendix A: Figures and Tables

Figure A.1: Comparative Statics

(a) Mean margin – small

(b) Mean margin – large

(c) Variance margin – small

(d) Variance margin – large

(e) Innov. elasticity – small

(f) Innov. elasticity – large

(g) Mean margin – small

(h) Mean margin – large

(i) Innov. elasticity – small
Figure A.1: Comparative Statics (continued)

(j) Innov. elasticity – large

(k) Mean margin – small

(l) Large firm researcher share

(m) Mean margin – large

(n) Large firm researcher share

(o) Large firm researcher share

Note: The figure plots selected simulated moments as a function of various parameters. Each panel plots the moment values on the $y$-axis and varying one parameter on the $x$-axis, keeping all other parameters at their baseline estimates.
Figure A.2: Counterfactual Experiments

(a) Probability of small firm facing successful entry

(b) Total product varieties

(c) Welfare

Note: The figure plots selected outcomes of interest with varying degrees of disclosure requirement on small firms, $\tau_d$. Each point on the curve represents the outcome of a counterfactual policy. The red circle indicates the baseline estimates of the outcomes.
Table A.1: Data Definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Regression Variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Innov. Expenditure (Log)</td>
<td>Logarithm of firm innovation expenditure</td>
<td>CIS</td>
</tr>
<tr>
<td>Innov. Propensity</td>
<td>A dummy variable that equals one if firm $i$ has positive innovation expenditure</td>
<td>CIS</td>
</tr>
<tr>
<td>Large</td>
<td>A dummy variable that equals one if the firm has more than 250 employees.</td>
<td>CIS</td>
</tr>
<tr>
<td>Below50</td>
<td>A dummy variable that equals one if the firm has fewer than 20 employees.</td>
<td>CIS</td>
</tr>
<tr>
<td><strong>Panel B: SMM Variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profit margin</td>
<td>Earnings before tax and amortization divided by sales (ebta/turn)</td>
<td>Orbis</td>
</tr>
<tr>
<td>Large firm</td>
<td>Firms with more than 250 employees</td>
<td>Orbis &amp; CIS</td>
</tr>
<tr>
<td>Small firm</td>
<td>Firms with fewer than 250 employees</td>
<td>Orbis &amp; CIS</td>
</tr>
<tr>
<td>Share of researchers employed by large firms</td>
<td>Percentage of R&amp;D personnel employed by firms with more than 250 employees</td>
<td>Eurostat</td>
</tr>
<tr>
<td>Innovation elasticity of large firms</td>
<td>Regression coefficient $\beta_1 + \beta_2$ from Model (3.1)</td>
<td>Orbis &amp; CIS</td>
</tr>
<tr>
<td>Innovation elasticity of small firms</td>
<td>Regression coefficient $\beta_2$ from Model (3.1)</td>
<td>Orbis &amp; CIS</td>
</tr>
</tbody>
</table>

Note: The table presents the definitions and data sources for variables used in regression analysis (Panel A) and estimation exercise (Panel B).
Table A.2: Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>SD</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All firms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regulation Intensity</td>
<td>560,882</td>
<td>0.160</td>
<td>0.228</td>
<td>0.018</td>
<td>0.032</td>
<td>0.067</td>
<td>0.165</td>
<td>0.470</td>
</tr>
<tr>
<td>InnovExpenditure</td>
<td>573,554</td>
<td>221,380.806</td>
<td>934,425.816</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>30,000.000</td>
<td>334,944.656</td>
</tr>
<tr>
<td>InnovExpenditure (Log)</td>
<td>573,554</td>
<td>4.077</td>
<td>5.607</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>10.309</td>
<td>12.722</td>
</tr>
<tr>
<td>Innov. Propensity</td>
<td>573,554</td>
<td>0.364</td>
<td>0.481</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Large</td>
<td>573,554</td>
<td>0.118</td>
<td>0.322</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Sales (in thousand Euro)</td>
<td>573,554</td>
<td>51,018.940</td>
<td>658,245.273</td>
<td>327.590</td>
<td>948.288</td>
<td>2,916.841</td>
<td>11,132.000</td>
<td>43,060.000</td>
</tr>
<tr>
<td>Number of Employees</td>
<td>29,697,946</td>
<td>27.133</td>
<td>971.103</td>
<td>1.000</td>
<td>1.000</td>
<td>3.000</td>
<td>9.000</td>
<td>25.000</td>
</tr>
<tr>
<td>Profit Margin</td>
<td>23,060,964</td>
<td>0.193</td>
<td>0.222</td>
<td>0.020</td>
<td>0.045</td>
<td>0.105</td>
<td>0.247</td>
<td>0.528</td>
</tr>
<tr>
<td><strong>Large firms (over 250 employees)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>InnovExpenditure</td>
<td>67,530</td>
<td>1,043,320.463</td>
<td>2,163,239.626</td>
<td>0.000</td>
<td>0.000</td>
<td>8,288.295</td>
<td>700,000.000</td>
<td>4,404,106.000</td>
</tr>
<tr>
<td>InnovExpenditure (Log)</td>
<td>67,530</td>
<td>6.883</td>
<td>6.703</td>
<td>0.000</td>
<td>0.000</td>
<td>9.023</td>
<td>13.459</td>
<td>15.298</td>
</tr>
<tr>
<td>Innov. Propensity</td>
<td>67,530</td>
<td>0.529</td>
<td>0.499</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Sales (in thousand Euro)</td>
<td>67,530</td>
<td>354,840.000</td>
<td>1,877,143.614</td>
<td>7,180.045</td>
<td>19,793.098</td>
<td>53,879.752</td>
<td>154,100.000</td>
<td>536,855.094</td>
</tr>
<tr>
<td>Number of Employees</td>
<td>321,065</td>
<td>1,580.359</td>
<td>9,205.560</td>
<td>278.000</td>
<td>339.000</td>
<td>493.000</td>
<td>942.000</td>
<td>2,171.000</td>
</tr>
<tr>
<td>Profit Margin</td>
<td>257,226</td>
<td>0.111</td>
<td>0.104</td>
<td>0.023</td>
<td>0.045</td>
<td>0.083</td>
<td>0.141</td>
<td>0.227</td>
</tr>
<tr>
<td><strong>Small firms (under 250 employees)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>InnovExpenditure</td>
<td>506,024</td>
<td>111,691.176</td>
<td>512,825.533</td>
<td>0.000</td>
<td>0.000</td>
<td>17,376.697</td>
<td>209,666.672</td>
<td></td>
</tr>
<tr>
<td>InnovExpenditure (Log)</td>
<td>506,024</td>
<td>3.702</td>
<td>5.334</td>
<td>0.000</td>
<td>0.000</td>
<td>9.763</td>
<td>12.253</td>
<td></td>
</tr>
<tr>
<td>Innov. Propensity</td>
<td>506,024</td>
<td>0.342</td>
<td>0.474</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Sales (in thousand Euro)</td>
<td>506,024</td>
<td>10,473.361</td>
<td>83,140.415</td>
<td>281.246</td>
<td>817.464</td>
<td>2,261.356</td>
<td>6,860.000</td>
<td>19,343.000</td>
</tr>
<tr>
<td>Number of Employees</td>
<td>29,376,881</td>
<td>10.157</td>
<td>23.078</td>
<td>1.000</td>
<td>1.000</td>
<td>3.000</td>
<td>8.000</td>
<td>22.000</td>
</tr>
<tr>
<td>Profit Margin</td>
<td>22,803,738</td>
<td>0.194</td>
<td>0.223</td>
<td>0.020</td>
<td>0.045</td>
<td>0.105</td>
<td>0.249</td>
<td>0.531</td>
</tr>
<tr>
<td>Share of Researchers employed by Large Firms (country level)</td>
<td>257</td>
<td>0.458</td>
<td>0.243</td>
<td>0.238</td>
<td>0.366</td>
<td>0.438</td>
<td>0.531</td>
<td>0.631</td>
</tr>
<tr>
<td>Regulation Intensity (country-industry level)</td>
<td>10,350</td>
<td>0.212</td>
<td>0.283</td>
<td>0.016</td>
<td>0.032</td>
<td>0.081</td>
<td>0.233</td>
<td>0.846</td>
</tr>
</tbody>
</table>

**Note:** The table presents descriptive statistics for the regression and estimation variables. Corresponding variable definitions are in Table A.1. The sample is based on Community Innovation Survey, Orbis and Eurostat. The sample spans from 2000 to 2018, covering firms from 17 countries. This table presents statistics for all firms, firms over and under 250 employees, respectively. All continuous variables are winsorized at the 1% and 99%. All variables are at the firm level, except for *Share of Researchers employed by Large Firms*, which is a country level variable obtained from Eurostat.
Table A.3: Innovation Impacts of Mandatory Financial Disclosure (by Group)

<table>
<thead>
<tr>
<th></th>
<th>(1) InnovExpenditure (Log)</th>
<th>(2) InnovExpenditure (Log)</th>
<th>(3) Innov. Propensity</th>
<th>(4) Innov. Propensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation Intensity</td>
<td>-0.477 (0.61)</td>
<td>-0.200 (0.63)</td>
<td>-0.024 (0.05)</td>
<td>-0.001 (0.05)</td>
</tr>
<tr>
<td>Large</td>
<td>2.436*** (0.17)</td>
<td>1.675*** (0.14)</td>
<td>0.128*** (0.01)</td>
<td>0.080*** (0.01)</td>
</tr>
<tr>
<td>Regulation Intensity ×</td>
<td>1.505*** (0.43)</td>
<td>1.216*** (0.35)</td>
<td>0.128*** (0.03)</td>
<td>0.104*** (0.02)</td>
</tr>
<tr>
<td>Above 250</td>
<td>-1.196*** (0.09)</td>
<td>-0.075*** (0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below 50</td>
<td>-1.039*** (0.20)</td>
<td>-0.078*** (0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>556,051</td>
<td>556,051</td>
<td>556,051</td>
<td>556,051</td>
</tr>
<tr>
<td>Country-Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry-Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Clusters (Country-Industry)</td>
<td>965</td>
<td>965</td>
<td>965</td>
<td>965</td>
</tr>
<tr>
<td>Adjusted R-Squared</td>
<td>0.333</td>
<td>0.344</td>
<td>0.301</td>
<td>0.308</td>
</tr>
</tbody>
</table>

Note: The table reports the results of regression model (3.1) and (3.2), regressing firm innovation expenditure/innovation propensity on regulation intensity and its interaction of size indicator variables. In Column 1 and 2, the dependent variable, *Innov. Expenditure (Log)*, is the logarithm of firm innovation expenditure, obtained from the Community Innovation Survey (CIS) data. In Column 3 and 4, the dependent variable, *Innov. Propensity*, is a dummy variable that equals 1 if firm $i$ has positive innovation expenditure according to CIS data. *Reg. Intensity* is the share of firms above reporting-related exemption thresholds in firm $i$’s country, industry, and year using a standardized firm-size distribution per industry (across countries). *Large* is a dummy variable that takes the value of 1 if the firm has more than 250 employees. *Below 50* is a dummy variable that takes the value of 1 if the firm has fewer than 50 employees. The regression include industry-year fixed effects. Standard errors (in parentheses) are clustered at the country-industry level. All industries are defined using two-digit NACE classifications. *, **, and *** indicate statistical significance levels of 10%, 5% and 1%, respectively.
### Table A.4: Estimation Results

#### Panel A: Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Standard Error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{\tau_o}$</td>
<td>Variance of prior information</td>
<td>0.263</td>
<td>0.064</td>
<td>4.11</td>
</tr>
<tr>
<td>$\frac{1}{\tau_d}$</td>
<td>Variance of small-firm disclosure</td>
<td>0.526</td>
<td>0.132</td>
<td>3.97</td>
</tr>
<tr>
<td>$\phi_L$</td>
<td>Low marginal cost of large firms</td>
<td>0.046</td>
<td>0.002</td>
<td>28.35</td>
</tr>
<tr>
<td>$\phi^H$</td>
<td>High marginal cost of large firms</td>
<td>0.082</td>
<td>0.009</td>
<td>8.97</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Mass of potential large entrants</td>
<td>0.632</td>
<td>0.002</td>
<td>267.94</td>
</tr>
</tbody>
</table>

#### Panel B: Moments

<table>
<thead>
<tr>
<th></th>
<th>Simulated moments</th>
<th>Data moments</th>
<th>Standard Error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of researchers employed by large firms</td>
<td>0.492</td>
<td>0.458</td>
<td>0.015</td>
<td>2.20</td>
</tr>
<tr>
<td>Mean profit margin of large firms</td>
<td>0.138</td>
<td>0.111</td>
<td>0.000</td>
<td>131.97</td>
</tr>
<tr>
<td>Variance of profit margin of large firms</td>
<td>0.017</td>
<td>0.011</td>
<td>0.000</td>
<td>83.32</td>
</tr>
<tr>
<td>Mean profit margin of small firms</td>
<td>0.210</td>
<td>0.194</td>
<td>0.000</td>
<td>342.81</td>
</tr>
<tr>
<td>Variance of profit margin of small firms</td>
<td>0.054</td>
<td>0.050</td>
<td>0.000</td>
<td>210.82</td>
</tr>
<tr>
<td>Innovation elasticity of large firms</td>
<td>0.027</td>
<td>0.031</td>
<td>0.009</td>
<td>-0.50</td>
</tr>
<tr>
<td>Innovation elasticity of small firms</td>
<td>-0.014</td>
<td>-0.014</td>
<td>0.008</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

*Note:* The table reports the estimation results of the simulated method of moment procedure. The procedure searches over the parameter space and finds the parameter values that minimize the distance of a vector of data moments and their corresponding model moments from simulation. Panel A presents the estimated structural parameters with standard errors and t-statistics. The standard errors are computed from the asymptotic covariance matrix, and the t-statistics are for the null that the parameters are equal to zero. Panel B reports the simulated and actual data moments. The standard error of the data moment is computed from bootstrap. Samples that are of the same size to the original sample are drawn from the data sets, and the corresponding moments are recalculate. Standard errors are then computed based on the distribution of the recalculated moments. The t-statistics are for the null that the simulated moments and data moments are equal.
Appendix B: Alternative Model Assumptions

B.1 No Direct Competition

In this section, I consider the case where the large firms enter local markets, learn from small firms’ disclosure, but they use this information to develop completely differentiated products from the small firms. In this case, when both firms are monopolists of their products, i.e., disclosure has no business stealing effect. This benchmark case is considered to better differentiate different channels through which increased disclosure from small firms can have an effect.

Small firms’ disclosure can reveal market specific information such as local customers’ general willingness to pay, local infrastructure condition etc. that is useful for large entrants, even if they don’t directly engage in head-to-head competition with small firms. In this version of the model, better information still helps a large firm to achieve lower cost, in that the probability that a low marginal cost $\phi_L$ realizes is still determined by the best estimate of the market condition, i.e., $z_n(\omega_n, a_{bn}) = 1 - \alpha(\omega_n - a_{bn})^2$. But the small firm’s profits from product innovation will not be encroached because both firms are monopolists of their own products.

In this case, the small firm earns $\pi_s^m$ for the products it develops, i.e., $E[\pi_s] = \pi_s^m$. The large firm charges a monopoly price of $p_b^{Lm} = \frac{\phi_L}{\theta}$ and earns a monopoly profit of $\pi_b^{Lm}$, if its cost realization is $\phi_b = \phi_L$; it charges a monopoly price of $p_b^{Hm} = \frac{\phi_H}{\theta}$ and earns a profit of $\pi_b^{Hm}$, if its cost realization is $\phi_b = \phi_H$.\footnote{The exact expressions are $\pi_b^{Lm} = (\phi_L) \frac{\nu}{\phi_L} (1 - \theta) \theta \frac{\nu}{\phi_L} A$, $\pi_b^{Hm} = (\phi_H) \frac{\nu}{\phi_H} (1 - \theta) \theta \frac{\nu}{\phi_H} A$.} Note that $\pi_b^{Lm} > \pi_b^{Hm}$, which means the large firm’s monopoly profit
is higher when its marginal cost is lower. The expected gross profit of innovation of a large firm is

\[
E[\pi_b] = E \left[ z_n(\omega_n, a_{bn}) \pi^L_m + (1 - z_n(\omega_n, a_{bn})) \pi^H_m \right],
\]

(B.1)

\[
= \alpha E[(\omega_n - \hat{\omega}_n)^2] \pi^L_m + \left(1 - \alpha E[(\omega_n - \hat{\omega}_n)^2]\right) \pi^H_m,
\]

(B.2)

\[
= \left(1 - \alpha \hat{\sigma}^2\right) \pi^L_b + \alpha \hat{\sigma}^2 \pi^H_b.
\]

(B.3)

(B.3) shows that the expected gross profit in is decreasing in \(\hat{\sigma}^2\), which means a large firm’s expected gross profit of innovation is increasing in accuracy of its information.

Given this payoff structure, Proposition 4 summarizes how small and large firms’ labor employment, innovation expenditure and firm value change with small firms’ disclosure quality.

**Proposition 4** Without business stealing, (i) large firms’ labor employment, innovation expenditure, and firm value are monotonically increasing in small firms’ disclosure; (ii) whereas small firms’ labor employment, innovation expenditure, and firm value are monotonically decreasing in small firms’ disclosure, i.e., \(\forall \tau_d\),

\[
\frac{dl_b}{d\tau_d} > 0, \quad \frac{dWl_b}{d\tau_d} > 0, \quad \frac{dV_b}{d\tau_d} > 0 \quad \text{(B.4)}
\]

\[
\frac{dl_s}{d\tau_d} < 0, \quad \frac{dWl_s}{d\tau_d} < 0, \quad \frac{dV_s}{d\tau_d} < 0 \quad \text{(B.5)}
\]

The proof is provided in Appendix C. The first result is intuitive. As disclosure quality increases, the large firms can better adjust to local market conditions and are more likely to obtain a lower cost and earn higher monopoly profit. This increase in expected gross profit leads to more innovation by the large firms as per optimal innovation function (2.13).

The second result may be more surprising: innovation and value of the small firms decrease even when the large firms do not directly use disclosure to take away small firms’ markets. This is because the increase in large firms’ innovation drives up the equilibrium wage \(W\) in the economy, and makes it more expensive for the small firms to innovate. Therefore, although the products
deliver as much gross profit for the small firms as when there’s lower level of disclosure, they innovate less and have lower firm values.

This result highlights the interconnection between firms through the general equilibrium channel. Strengthening of one group may hurt another group, because they are competing for the same resources. Empirically, we would expect to observe a decrease in small firms’ innovation investment and firm value post enhanced disclosure regulation, even in absence of direct competition and business stealing; it is sufficient that the regulation would increase profitability of the large firms.

**Proposition 5** Without business stealing, total innovation investment and total firm value are monotonically increasing in small firms’ disclosure quality, i.e., \( \forall \tau_d \),

\[
\frac{dWL}{d\tau_d} > 0, \quad \frac{dV}{d\tau_d} > 0. \tag{B.6}
\]

The proof is provided in Appendix C. The reason why total innovation expenditure \( WL \) increases is straightforward: \( L \) is constant for a fixed amount of labor supply and equilibrium wage \( W \) increases in disclosure \( \tau_d \), as more innovation from large firms drives up equilibrium wage. Total firm value \( V \) also increases with disclosure, despite the decrease in small firm value. This is because on aggregate, large firms are becoming more profitable with better disclosure, i.e., \( \frac{dE[\pi_b]}{d\tau_d} > 0 \); while the small firms’ gross profitability remains the same, i.e., \( \frac{dE[\pi_s]}{d\tau_d} = 0 \). Though there is reshuffling on innovation and firm value between the two groups of firms as shown in proposition 4, total firm value for the whole economy is increasing with disclosure.

For the welfare measure – total consumption, write consumption \( C \) as contributions from small and large firms separately:

\[
C = \left( (x_s^m)^\theta - \phi x_s^m \right) l_s^\gamma_s + \mu (1 - \alpha \hat{\sigma}^2) \left( (x_b^L)^\theta - \phi x_b^L \right) l_b^\gamma_b + \mu \alpha \hat{\sigma}^2 \left( (x_b^H)^\theta - \phi x_b^H \right) l_b^\gamma_b \tag{B.7}
\]

\[
= c_s^m l_s^\gamma_s + \mu \left( 1 - \alpha \hat{\sigma}^2 \right) c_b^L l_b^\gamma_b + \mu \alpha \hat{\sigma}^2 c_b^H l_b^\gamma_b \tag{B.8}
\]
where $c^m_s$, $c^{Lm}_b$, and $c^{Hm}_b$ denote the consumption generated from each new product innovation by small firms, and by large firms with low and high marginal cost respectively; superscript $m$ denotes that the consumption is generated by monopolists. The proportion of large firms with low cost realization is $1 - \alpha \hat{\sigma}^2$. The following lemma summarizes consumption per innovation generated by different types of firms.

**Lemma 3** Without business stealing, consumption generated per innovation ranks as follows:

$$c^{Hm}_b < c^m_i < c^{Lm}_b.$$  \hfill (B.9)

The proof is provided in Appendix C. An innovation from a low-cost large firm generates more consumption than that from a small firm, which in turn generates more consumption than a high-cost large firm. Because all firms are producing their respective monopoly quantities, the ranking of consumption per innovation is solely determined by firms’ marginal production costs. The lower a firm’s marginal cost, the more consumption would be generated from a newly developed product.

The derivative of welfare with respect to disclosure $\tau_d$ is,

$$\frac{dC}{d\tau_d} = \gamma c^m_s l_s^{\gamma-1} \frac{dl_s}{d\tau_d} + \gamma \mu E[c^m_b l_b^{\gamma-1}] \frac{dl_b}{d\tau_d} + \mu (c^{Lm}_b - c^{Hm}_b) l_b^{\gamma} \frac{d(1 - \alpha \hat{\sigma}^2)}{d\tau_d}. \hfill (B.10)$$

Increasing small firms’ disclosure quality has again a innovation reallocation effect and a production reallocation effect on welfare. The production reallocation effect plays an unambiguously positive role on welfare, since low-cost large firms can generate higher consumption per innovation than high-cost ones.

For the innovation reallocation effect, as in the main model, it is positive if and only if the large firms have a higher expected consumption contribution to expected gross profit ratio, i.e., $A' > 0$ if and only if $\frac{E[c^m_b]}{E[\pi_b]} > \frac{c^m_s}{\pi^m_s}$. In the case without business stealing, these two ratios are always
equal because both types of firms charge monopoly prices and produce monopoly quantities. For large firms, their expected consumption contribution and expected gross profits move in tandem and the ratio remains unchanged, as small-firm disclosure quality increases. For small firms, their consumption contribution and profit per innovation are deterministic and the ratio is identical to that of the large firms. This means the innovation reallocation effect is always zero.

In summary, as small firms’ disclosure becomes more accurate, large firms earn higher expected gross profit and also generate more consumption in expectation. The increase in large firms’ expected gross profit translates into increase in innovation, which exactly matches their increased consumption contribution. Therefore, the reduction in small firms’ innovation is exactly offset by large firms’ increase in innovation, leaving only the positive production reallocation effect at work. Proposition 6 summarizes this result. The proof is provided in Appendix C.

**Proposition 6** Without business stealing, total welfare is monotonically increasing in small firms’ disclosure quality, i.e., \( \forall \tau_d, \)

\[
\frac{dC}{d\tau_d} > 0. \tag{B.11}
\]

**B.2 Cournot Competition**

This section considers a case where firms engage in Cournot instead of Bertrand competition in the product market, i.e., firms compete on the quantity rather than price of the products that they produce. Changing the competition form changes the ex-post production and product market profits of the firms. It does not change the relation between ex-ante expected profits and firms’ innovation incentives. Therefore, I start with the analysis of ex-post product market competition. In the following analysis, I use \( x_s, x_b, \) and \( \pi_s, \pi_b \) to denote the ex-post quantity of products and product market profit by small and large firms, respectively, and the superscripts that denote the form of competition is omitted when doing so does not raise confusion.
Given demand function from equation (2.3), the equilibrium price of a product \( j \) is determined by the total quantity of products produced, namely,

\[
p_j = \theta (x_s + x_b)^{\theta -1}.
\]  

(B.12)

The ex-post product market profit for each firm is determined by their choice of quantity, equilibrium product price and their respective marginal cost:

\[
\pi_s = (p_j - \phi_s)x_s = (\theta (x_s + x_b)^{\theta -1} - \phi_s)x_s,
\]  

(B.13)

\[
\pi_b = (p_j - \phi_b)x_b = (\theta (x_s + x_b)^{\theta -1} - \phi_b)x_b,
\]  

(B.14)

where \( \phi_s = \phi \) and \( \phi_b \in \{ \phi^L, \phi^H \} \). As each firm chooses quantity to maximize their respective profit, the first-order conditions are as follows,

\[
\frac{\partial \pi_s}{\partial x_s} = \theta (x_s + x_b)^{\theta -1} - \phi + \theta(\theta - 1)(x_s + x_b)^{\theta - 2}x_s = 0,
\]  

(B.15)

\[
\frac{\partial \pi_b}{\partial x_b} = \theta (x_s + x_b)^{\theta -1} - \phi_b + \theta(\theta - 1)(x_s + x_b)^{\theta - 2}x_b = 0,
\]  

(B.16)

and each firm’s optimal quantity can be solved from the above system of equations. In particular, the equilibrium quantities are:

\[
x_s + x_b = \left( \frac{\phi + \phi_b}{\theta^2 + \theta} \right)^{\frac{1}{\theta - 1}},
\]  

(B.17)

\[
x_s = \frac{(\phi_b - \theta \phi)(\theta^2 + \theta)^{\frac{1}{\theta - 1}}}{(1 - \theta)(\phi + \phi_b)^{\frac{1}{\theta - 1}}},
\]  

(B.18)

\[
x_b = \frac{(\phi - \theta \phi_b)(\theta^2 + \theta)^{\frac{1}{\theta - 1}}}{(1 - \theta)(\phi + \phi_b)^{\frac{1}{\theta - 1}}},
\]  

(B.19)

In the case of Bertrand competition, the firm with higher marginal cost is priced out of the market and makes zero profit. With Cournot competition, the optimal quantities \( x_s \) and \( x_b \) imply that when the large firm obtains a low cost realization, i.e., \( \phi_b = \phi^L \), the small firm still produces a positive
amount and makes a positive profit if $\phi^L > \theta \phi$; it makes zero profit if $\phi^L \leq \theta \phi$. Similarly, when $\phi_b = \phi^H$, the large firm still produces a positive quantity and makes a positive profit if $\phi > \theta \phi^H$.

Here I focus on the parameter ranges where both firms produce a positive amount, i.e., $\phi^L > \theta \phi$ and $\phi > \theta \phi^H$, since these are the distinct cases that arise in Cournot competition. Moreover, it is worth noting that the large firms produce a positive amount and earn a positive profit, as long as $\phi^L < \frac{\phi}{\theta}$, which implies that it continues to be profitable in the product market even if $\phi^L = \phi$.

Similarly, the small firms earn a positive profit even when $\phi^H = \phi$, and they continue to do so as long as $\phi^H > \theta \phi$. Therefore, we do not need to restrict to the case where $\phi^L < \phi < \phi^H$, and we can consider a generic parameter combination of $\{\phi^L, \phi, \phi^H\}$ as long as $\phi^L < \phi^H$.

Substituting (B.17-B.19) into (B.13) and (B.14), the equilibrium product market profit for small and large firms are

$$
\pi_s = (\phi_b - \theta \phi)^2 \frac{\theta^{\frac{1}{\gamma}} (1 + \theta)^{\frac{\alpha}{\gamma}}}{(1 - \theta)(\phi + \phi_b)^\frac{2-\gamma}{\gamma}},
$$

$$
\pi_b = (\phi - \theta \phi_b)^2 \frac{\theta^{\frac{1}{\gamma}} (1 + \theta)^{\frac{\alpha}{\gamma}}}{(1 - \theta)(\phi + \phi_b)^\frac{2-\gamma}{\gamma}}.
$$

where the values of $\pi_s$ and $\pi_b$ depend on the realization of $\phi_b$. Denote $\pi^L_s$, $\pi^H_s$, $\pi^L_b$, and $\pi^H_b$, small and large firm’s ex-post product market profit if the large firm’s marginal cost realization is $\phi^L$ and $\phi^H$ respectively. The expected profits of small and large firms are

$$
E[\pi_s] = (1 - \mu)\pi^m_s + \mu \alpha \hat{\sigma}^2 \pi^H_s + \mu \left(1 - \alpha \hat{\sigma}^2\right) \pi^L_s,
$$

$$
E[\pi_b] = \left(1 - \alpha \hat{\sigma}^2\right) \pi^L_b + \alpha \hat{\sigma}^2 \pi^H_b.
$$

The expressions of expected product market profits under Cournot competition (equation (B.22) and (B.23)), differ from those under Bertrand competition (equation (2.21) and (2.23)). This is the most critical difference between the two cases. Given expected product market profits, the subsequent analysis follows a similar manner as in the main paper.

First, because Proposition 1 is proven based on firms’ expected product market profits, to show
that Proposition 1 continues to hold with Cournot competition, it is sufficient to show that the expected product market profit of small firms is decreasing in disclosure and vice versa for the large firms. The following proposition shows that this is indeed the case.

**Proposition 7** If firms engage in Cournot competition in the product market, the expected product market profit of a small firm is decreasing in disclosure and expected product market profit of a large firm is decreasing in disclosure i.e.,

\[
\frac{dE[\pi_s]}{d\tau_d} < 0, \quad \frac{dE[\pi_b]}{d\tau_d} > 0.
\] (B.24)

And Proposition 1 continues to hold, i.e., \(\forall \tau_d\),

\[
\frac{dl_b}{d\tau_d} > 0, \quad \frac{dWl_b}{d\tau_d} > 0, \quad \frac{dV_b}{d\tau_d} > 0 \quad \frac{dl_s}{d\tau_d} < 0, \quad \frac{dWl_s}{d\tau_d} < 0, \quad \frac{dV_s}{d\tau_d} < 0.
\] (B.25)

The proof is provided in Appendix C. This result is intuitive, as the reallocation effect of disclosure hinges on the mechanism that large firms earning higher expected profits from the product market with more disclosure and thus innovate more, while small firms expect lower profits and innovate less. As long as this mechanism is preserved, the exact form of competition in the product market is not critical.

Second, the effect of disclosure on aggregate innovation, presented in Proposition 2, is also derived based on firms’ expected product market profits. Therefore, the general argument of the possible cases still applies in the case of Cournot competition, albeit the exact values of thresholds differ.

**Proposition 8** If firms engage in Cournot competition in the product market, given a combination of parameter values \(\{\phi, \phi^L, \phi^H\}\), such that \(\phi^L < \phi^H\), one of the following cases hold.
a. If \( \phi^L \) is lower than some threshold \( \phi^* \), total innovation expenditure and total firm value are monotonically increasing in small firms’ disclosure precision.

b. If \( \phi^L \) is higher than some threshold \( \phi^* \), total innovation expenditure and total firm value are decreasing in small firms’ disclosure precision.

c. If \( \phi^* < \phi^L < \phi^* \), total innovation expenditure and total firm value first decrease then increase in small firms’ disclosure.

Here * indicates that the threshold values are different under Cournot competition. The conditions that pin down the threshold values and the proof of Proposition 8 are provided in Appendix C.

Finally, to examine the impact of disclosure on welfare, it is again useful to first analyze the extreme cases. Under Bertrand competition, as \( \phi^L \) goes very close to \( \phi \), the large firms earns a profit margin that equals to zero regardless whether its cost realization is \( \phi^L \) or \( \phi^H \), and it thus has no incentive to innovate. Under Cournot competition, the large firm still earns a positive profit even when \( \phi^L = \phi \), because Cournot competition is a less intense form of competition. However, if \( \phi^L \) is so high such that \( \phi^L > \frac{\phi}{\theta} \), the large firm will be earning zero profit from the product market even with a low cost realization and has no incentive to innovate. Similarly, the small firm still earns a positive profit even when \( \phi^H = \phi \), but if \( \phi^H < \theta \phi \), the small firm stops innovating. Incorporating the feature of Cournot competition, the effect of disclosure on welfare can be determined in the following extreme cases.

**Proposition 9** In the case where firms engage in Cournot competition in the product market,

a. if \( \phi^L \) goes sufficiently close to \( \frac{\phi}{\theta} \) from below and \( \phi^H > \frac{\phi}{\theta} \), welfare is decreasing in small firms’ disclosure precision;

b. if \( \phi^H \) goes sufficiently close to \( \theta \phi \) from above and \( \phi^L < \theta \phi \), welfare is increasing in small firms’ disclosure precision.
The proof of Proposition 9 follows step by step the proof of Proposition 3 and is thus omitted. With the same logic that if the large firms marginal costs are high and do not earn high product market profits, they do not innovate much but yet drive the small firms out of market, and welfare decreases as a consequence. In the opposite case, if large firms’ marginal costs are sufficiently low, such that the small firms do not earn a high profit even when the large firms’ marginal cost is $\phi^H$, then more disclosure helps large firms to obtain a cost advantage and increases overall welfare.
Appendix C: Model Proofs

C.1 Preliminaries

• In all the following proofs, value of equilibrium land quantity $A = 1$ is already substituted
into the equations to reduce clutter.

• Useful expressions: substituting equilibrium wage (2.15) into optimal innovation function
(2.13), and value functions (2.16) and (2.17).

\[
\begin{align*}
  l_s &= \frac{E[\pi_s]^{\frac{1}{1-\gamma}}}{E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}}}, \\
  W l_s &= \frac{\gamma E[\pi_s]^{\frac{1}{1-\gamma}}}{(E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}})^\gamma}, \\
  V_s &= \frac{(1 - \gamma) E[\pi_s]^{\frac{1}{1-\gamma}}}{(E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}})^\gamma}, \\
  l_b &= \frac{E[\pi_b]^{\frac{1}{1-\gamma}}}{E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}}}; \\
  W l_b &= \frac{\gamma E[\pi_b]^{\frac{1}{1-\gamma}}}{(E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}})^\gamma}, \\
  V_b &= \frac{(1 - \gamma) E[\pi_b]^{\frac{1}{1-\gamma}}}{(E[\pi_s]^{\frac{1}{1-\gamma}} + \mu E[\pi_b]^{\frac{1}{1-\gamma}})^\gamma}.
\end{align*}
\]
C.2 Proofs for Main Analysis

Proof of Proposition 1

Use $\pi_s$ and $\pi_b$ as shorthand notations for $E[\pi_s]$ and $E[\pi_b]$.

\[
\frac{dl_s}{d\tau_d} = \frac{d(\pi_s^{1/\gamma})}{d\tau_d} \left( \pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma} \right)^{-1} - \pi_s^{1/\gamma} \left( \pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma} \right)^{-2} \left( \frac{d(\pi_s^{1/\gamma})}{d\tau_d} + \mu \frac{d(\pi_b^{1/\gamma})}{d\tau_d} \right) \tag{C.4}
\]

\[
\frac{dl_b}{d\tau_d} \propto \frac{d(\pi_b^{1/\gamma})}{d\tau_d} \left( \pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma} \right)^{-1} \left( 1 - \frac{\mu \pi_b^{1/\gamma}}{\pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma}} \right) \left( -\pi_s^{1/\gamma} \left( \pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma} \right)^{-2} \right) \frac{d(\pi_s^{1/\gamma})}{d\tau_d} \tag{C.7}
\]

For, $W_l$ and $V_l$, they are proportional to each other, it is sufficient to show the comparative statics for one group.

\[
\frac{dV_s}{d\tau_d} \propto \frac{d(\pi_s^{1/\gamma})}{d\tau_d} \left( \pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma} \right)^{-\gamma} - \gamma \pi_s^{1/\gamma} \left( \pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma} \right)^{-\gamma-1} \left( \frac{d(\pi_s^{1/\gamma})}{d\tau_d} + \mu \frac{d(\pi_b^{1/\gamma})}{d\tau_d} \right) \tag{C.9}
\]

\[
\frac{dV_b}{d\tau_d} \propto \frac{d(\pi_b^{1/\gamma})}{d\tau_d} \left( \pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma} \right)^{-\gamma} \left( 1 - \frac{\gamma \pi_b^{1/\gamma}}{\pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma}} \right) \left( -\gamma \pi_s^{1/\gamma} \left( \pi_s^{1/\gamma} + \mu \pi_b^{1/\gamma} \right)^{-\gamma-1} \right) \frac{d(\pi_b^{1/\gamma})}{d\tau_d} \tag{C.10}
\]
\[
\frac{dV_b}{d\tau_d} \propto \frac{d\left(\frac{1}{\tau_d^{\gamma}}\right)}{d\tau_d} \left(\frac{1}{\tau_d^{\gamma}} + \mu \frac{1}{\tau_d^{\gamma+\gamma}}\right)^{-1} \gamma \left(1 - \gamma \left(\frac{\mu \tau_d^{\gamma}}{\tau_d^{\gamma}} + \mu \frac{1}{\tau_d^{\gamma+\gamma}}\right)\right) \frac{d\left(\frac{1}{\tau_d^{\gamma}}\right)}{d\tau_d} < 0
\]

(C.11)

\[
\frac{d\left(\frac{1}{\tau_d^{\gamma}}\right)}{d\tau_d} > 0
\]

(C.12)

**Proof of Proposition 2**

It is sufficient to show comparative statics for \(\left(E[\pi_s]^{\frac{1}{\gamma}} + \mu E[\pi_b]^{\frac{1}{\gamma}}\right)^{-\gamma}\), which is proportional to \(V, WL\). Use \(\pi_s\) and \(\pi_b\) as shorthand notations for \(E[\pi_s]\) and \(E[\pi_b]\),

\[
\frac{d\left(\frac{1}{\tau_d^{\gamma}}\right)}{d\tau_d} = (1 - \gamma)\left(\frac{1}{\tau_d^{\gamma}} + \mu \frac{1}{\tau_d^{\gamma+\gamma}}\right)^{-\gamma} \gamma \left(\frac{1}{1 - \gamma} \frac{\tau_d^{\gamma}}{d\tau_d} d\pi_s + \mu \frac{1}{1 - \gamma} \frac{\tau_d^{\gamma}}{d\tau_d} d\pi_b\right),
\]

(C.13)

\[
\propto \pi_s^{\gamma} \frac{\tau_d^{\gamma}}{d\tau_d} + \mu \pi_b^{\gamma} \frac{\tau_d^{\gamma}}{d\tau_d},
\]

(C.14)

\[
\propto -\pi_s^{\gamma} + \pi_b^{\gamma}
\]

(C.15)

which is larger than 0 if and only if

\[
\left(1 - \alpha \sigma^2\right)\pi_s^c \frac{1}{\gamma} \pi_b^c \pi_b^c > 0.
\]

(C.16)

(C.16) is decreasing in \(\sigma^2\), so if \(\sigma^2\) is sufficiently low such that (C.16) holds at \(\sigma^2 = \sigma_o^2\), then \(\frac{dV}{d\tau_d} > 0, \frac{dWL}{d\tau_d} > 0\) hold for all \(\tau_d\). \(\phi^L\) is such that (C.16) just holds at \(\sigma_o^2\).

To obtain a closed-form expression of \(\phi^L\), I further focus on the case where competition prices are lower than the monopoly prices.\(^1\) Then the condition for \(\phi^L\) can be written in terms of model primitives by substituting in the expressions of \(\pi_b^c, \pi_s^m\) and \(\pi_s^c\).

\(^1\)For the other three possible cases: case (i), the large firm can charge a monopoly price because \(\phi^L > \theta \phi\), then substitute in \(\pi_b^c = \pi^m_b, \pi_s^m\) and \(\pi_s^c\); case (ii) the small firm can charge a monopoly price because \(\phi^H > \frac{\phi}{\bar{\phi}}\), then substitute in \(\pi_b^c\), and \(\pi_s^c = \pi^m_s\); case (iii) both firms can charge a monopoly price if they win the price competition, then substitute in \(\pi_b^c = \pi^m_b\), and \(\pi_s^c = \pi^m_s\).
\[
\phi^L < \phi - \left( \frac{(1 - \mu) (\frac{\phi}{\bar{g}} - \phi) \left( \frac{\sigma_o^2}{\hat{\sigma}_o^2} \right)^{1/v} + \mu \alpha \hat{\sigma}_o^2 (\phi^H - \phi) \left( \theta \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{1/v}}{1 - \alpha \hat{\sigma}_o^2} \right)^\gamma (\phi^H - \phi)^{(1-\gamma)} \left( \theta \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{\gamma \frac{1}{v-\gamma}} \left( \theta \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{\gamma \frac{1}{v-\gamma}} \equiv \phi^L \quad (C.17)
\]

On the other hand, at \( \hat{\sigma}_o^2 = 0 \), \((C.16)\) reads as

\[
(\pi^c_b)^{\frac{1}{1-\gamma}} - ((1 - \mu) \pi^m_s)^{\frac{\gamma}{1-\gamma}} \pi^c_s. \quad (C.19)
\]

\((C.19)\) is decreasing in \( \phi^L \), therefore if \( \phi^L \) is so high that it is only marginally lower than \( \phi \), then \((C.19)\) < 0, and \( \frac{dV}{d\tau_d} < 0, \frac{dWL}{d\tau_d} < 0 \) hold for all \( \tau_d \). \( \phi^L \) is such that \((C.16)\) = 0 holds.

Again, focus on the case where competition prices are lower than the monopoly prices, the expression of \( \phi^L \) can be obtained by substitute in \( \pi^c_b, \pi^m_s \) and \( \pi^c_s \),

\[
\phi^L > \phi - (1 - \mu)^\gamma \left( \frac{\phi}{\hat{\sigma}_o^2} \right)^{\frac{1}{v-\gamma}} (\phi^H - \phi)^{(1-\gamma)} \left( \theta \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{\gamma \frac{1}{v-\gamma}} \left( \theta \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{\gamma \frac{1}{v-\gamma}} \equiv \phi^L. \quad (C.20)
\]

Moreover, to compare \( \phi^L \) and \( \phi^L \), we need to know the sign of the following expression,

\[
\left( \frac{(1 - \mu) (\frac{\phi}{\bar{g}} - \phi) \left( \frac{\sigma_o^2}{\hat{\sigma}_o^2} \right)^{1/v} + \mu \alpha \hat{\sigma}_o^2 (\phi^H - \phi) \left( \theta \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{1/v}}{1 - \alpha \hat{\sigma}_o^2} \right) - (1 - \mu) (\frac{\phi}{\theta} - \phi) \left( \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{\gamma \frac{1}{v-\gamma}} \left( \theta \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{\gamma \frac{1}{v-\gamma}} \equiv \phi^L \quad (C.21)
\]

\[
= (1 - \mu) (\frac{\phi}{\theta} - \phi) \left( \frac{\sigma^2}{\hat{\sigma}_o^2} \right)^{\gamma \frac{1}{v-\gamma}} \left( \frac{1}{1 - \alpha \hat{\sigma}_o^2} \right) - 1 \] + positive constants \quad (C.22)

> 0. \quad (C.23)

Therefore, \( \phi^L \) is always greater than \( \phi^L \).

If \( \phi^L \) is such that \((C.16)\) is violated at \( \hat{\sigma}_o^2 = \hat{\sigma}_o^2 \), but \((C.16)\) > 0 for some \( \tau_d > \tau_d^* \), then \( V \) and \( WL \) first decrease for \( \tau_d \in [0, \tau_d^*] \) then increase for \( \tau_d \in [\tau_d^*, +\infty] \).
Proof of Lemma 1

First note that the price of product $j$ that maximizes consumption of product $j$ equals its marginal cost $\phi_j$. To see this, $p_j$ should solve the following maximization problem,

$$\max_{p_j} \quad x_j(p_j, A) - \phi_j x_j(p_j, A), \quad (C.24)$$

$$\iff \max_{p_j} \quad \left( \frac{\theta}{p_j} \right)^{\frac{\theta}{1-\theta}} - \phi_j \left( \frac{\theta}{p_j} \right)^{\frac{1}{1-\theta}} \quad (C.25)$$

The first order condition is

$$\frac{-\theta}{1-\theta} \theta^{\frac{\theta}{1-\theta}} p_j^{-\frac{\theta}{1-\theta}-1} + \theta^{\frac{1}{1-\theta}} \frac{1}{1-\theta} \phi_j p_j^{-\frac{1}{1-\theta}-1} = 0 \quad (C.26)$$

$$\Rightarrow p_j = \phi_j. \quad (C.27)$$

$(C.25)$ is monotonically decreasing for $p_j \geq \phi_j$, which means given a level of marginal cost, firms generate more consumption if they charge a lower price, as long as this price is above the marginal cost. Given a level of marginal cost, the monopoly price is weakly higher than the Bertrand competition price, i.e., $p^m_i \geq p^c_i$. Therefore $c^m_i \leq c^c_i, \forall i \in \{s, b\}$.

Consider $c^m_s$ and $c^c_s$, consumption generated by small firms under monopoly and Bertrand competition. For small firms, the marginal cost always equals $\phi$, and the above reasoning immediately applies here: $p^m_s = \phi_s \geq \phi^H = p^c_s$, thus $c^m_s \leq c^c_s$.

Consider $c^c_s$ and $c^c_b$, consumption generated by small and large firms respectively under competition. (i) If $\phi^H$ and $\phi^L$ are such that $\phi_s \leq \phi^H$, and $\phi^L \leq \phi$, then both firms can effectively charge monopoly price, and $c^c_s = c^m_s, c^c_b = c^L_m$. Then $c^c_s = c^m_s < c^L_m = c^c_b$, by Lemma 3.

(ii) If $\phi_s \leq \phi^H$, and $\phi^L > \phi$, then only the small firm charges monopoly price. In this case, $c^c_s = c^m_s < c^L_m < c^c_b$. (iii) If $\phi_s > \phi^H$, and $\phi^L \leq \phi$, then only the large firm charges monopoly price. However, the large firms’ monopoly price is still lower than the small firms’ Bertrand competition price: $p^c_b = p^m_b = \phi^L \leq \phi < \phi^H = p^c_s$. Even if the small firm would have marginal cost $\phi^L$, the
small firm would generate less consumption, because it charges a higher price. As $\phi > \phi^L$, the small firm also produces with a higher cost, thus $c^s < c^b$ a fortiori. (iv) If $\frac{\phi}{H} > \frac{\phi}{L}$, and $\frac{\phi^L}{\theta} > \phi$, then both firms charge Bertrand competition price. Given $p^c_b = \phi < \phi^H = p^c_s$, and $\phi^L < \phi$, the argument in case (iii) also applies here, and $c^s < c^b$.

**Proof of Proposition 3**

To prove the proposition, it is useful to decompose the effect of increasing small firms’ disclosure quality into two parts. First, it shifts innovation from small to big firms (ls decreases and lb increases). Second, given the levels of innovation, it shifts production (thus also consumption contribution) from small to big firms. This is reflected in the derivative of welfare with respect to disclosure $\tau_d$:

$$\frac{dC}{d\tau_d} = \gamma E[c_s]I_s^{\gamma-1}\frac{dl_s}{d\tau_d} + \gamma \mu E[c_b]I_b^{\gamma-1}\frac{dl_b}{d\tau_d} + \mu(c^c_b - c^c_s)\frac{d(1 - \alpha \hat{\sigma}^2)}{d\tau_d}.$$  

(A: innovation reallocation)  

(B: production reallocation)  

(C.28)

The term labeled “innovation reallocation” shows the innovation incentive change as a result of large firms’ learning and competition with small firms (as summarized in Proposition 1). The term labeled “production reallocation” shows product market competition change, where large firms are more likely to obtain low marginal cost and drive out small firms. The innovation reallocation and production reallocation generally have countervailing effects, because high profit margins drive firms to innovate but negatively affect the consumption intensities of the innovations.

I prove the proposition by first proving the following two lemmas.

**Lemma 4** If $\phi^L \to \phi$, or if $\phi^H \to \phi$ the production effect dominates.

Proof: As $\phi^L \to \phi$, large firms’ ex-post product market profit $\pi^c_b \to 0$, and large firms’ expected profit $E[\pi_b]$ does not respond to change in $\tau_d$, i.e.,

$$\left.\frac{dE[\pi_b]}{d\tau_d}\right|_{\phi^L \to \phi} = \left.\frac{d}{d\tau_d}\left(1 - \alpha \hat{\sigma}^2\right)\pi^c_b\right|_{\phi^L \to \phi} \to 0.$$  

(C.29)
As a result, large firms’ innovation decision does not respond to change in \( \tau_d \); by equation (C.7),
\[
\frac{dl_b}{d\tau_d}\bigg|_{\phi^L \rightarrow \phi} \rightarrow 0. \tag{C.30}
\]

So is small firms’ innovation decision:
\[
\frac{dl_s}{d\tau_d}\bigg|_{\phi^L \rightarrow \phi} \rightarrow 0. \tag{C.31}
\]

Similarly, if \( \phi^H \rightarrow \phi \), small firms’ expected profit does not respond to change in \( \tau_d \), i.e.,
\[
\frac{d E[\pi_s]}{d\tau_d}\bigg|_{\phi^H \rightarrow \phi} = \frac{d \left((1 - \mu)\pi_s^m + \mu\alpha \hat{\sigma}^2\right)}{d\tau_d}\bigg|_{\phi^H \rightarrow \phi} \rightarrow 0.
\]

Hence, \( \frac{dl_s}{d\tau_d}\bigg|_{\phi^H \rightarrow \phi} \rightarrow 0 \), and \( \frac{dl_b}{d\tau_d}\bigg|_{\phi^H \rightarrow \phi} \rightarrow 0 \). Therefore, term \( A \rightarrow 0 \).

**Lemma 5** The production reallocation effect is positive if and only if the large firms’ expected gross profit from innovation is sufficiently high. That is,
\[
B > 0 \iff E[\pi_b] > E[\pi_s] \left(\frac{c^e_s}{c^e_b}\right)^{\frac{1-\gamma}{\gamma}}. \tag{C.32}
\]

Proof:
\[
(B) = \mu(c^e_b I^Y_b - c^e_s I^Y_s) \frac{d (1 - \alpha \hat{\sigma}^2)}{d\tau_d}
\]
\[
\propto c^e_b I^Y_b - c^e_s I^Y_s \tag{C.33}
\]

(C.34) > 0 if and only if \( \frac{c^e_b}{c^e_s} > \left(\frac{L_b}{L_s}\right)^{\gamma} \), which is equivalent to \( E[\pi_b] > E[\pi_s] \left(\frac{c^e_s}{c^e_b}\right)^{\frac{1-\gamma}{\gamma}} \).

When \( \phi^L \rightarrow \phi \), \( E[\pi_b] = (1 - \alpha \hat{\sigma}^2) \pi^c_b \rightarrow 0 \), by Lemma 5, \( B < 0 \), i.e., the production effect is negative. Moreover, when \( \phi^L \rightarrow \phi \), by Lemma 4, the production effect dominates. Therefore, \( \frac{dc}{d\tau_d} < 0 \).

When \( \phi^H \rightarrow \phi \), the production effect again dominates by Lemma 4. Then \( \frac{dc}{d\tau_d} > 0 \) as long as \( B > 0 \). Because \( c^e_b \geq c^e_s \) by Lemma 1, a sufficient condition for \( B > 0 \) is thus \( L_b > L_s \).
which is equivalent to $E[\pi_b] > E[\pi_s]$. Therefore, if $E[\pi_b \mid \hat{\sigma}^2 = \hat{\sigma}^2_\sigma] > E[\pi_s \hat{\sigma}^2 = \hat{\sigma}^2_\sigma]$, then $E[\pi_b] > E[\pi_s]$ holds for all $\tau_d$, and $\frac{dC}{d\tau_d} > 0$ holds for all $\tau_d$. $\phi^L$ is such that $(1 - \alpha \hat{\sigma}^2_\sigma) \pi^c_b \geq ((1 - \mu)\pi^m_s + \mu \alpha \hat{\sigma}^2_\sigma)$ just holds.
C.3 Proofs for Alternative Model Assumption Analysis

Proof of Proposition 4.

Use $\pi_s$ and $\pi_b$ as shorthand notations for $E[\pi_s]$ and $E[\pi_b]$.

$$
\frac{dl_s}{d\tau_d} = -\pi_s^{1-\gamma} (\pi_s^{1-\gamma} + \mu \pi_b^{1-\gamma})^{-2} \mu \frac{d(\pi_b^{1-\gamma})}{d\tau_d} < 0
$$  \hspace{1cm} (C.35)

$$
\frac{dl_b}{d\tau_d} \propto \frac{d(\pi_b^{1-\gamma})}{d\tau_d} (\pi_s^{1-\gamma} + \mu \pi_b^{1-\gamma})^{-1} \left( 1 - \frac{\mu \pi_b^{1-\gamma}}{\pi_s^{1-\gamma} + \mu \pi_b^{1-\gamma}} \right) > 0
$$  \hspace{1cm} (C.36)

Since $W_l$ and $V_i$ are proportional to each other, for the rest two groups of quantities, it is sufficient to show the comparative statics for one group.

$$
\frac{dV_s}{d\tau_d} \propto -\gamma \pi_s^{1-\gamma} (\pi_s^{1-\gamma} + \mu \pi_b^{1-\gamma})^{-\gamma-1} \mu \frac{d(\pi_b^{1-\gamma})}{d\tau_d} < 0
$$  \hspace{1cm} (C.37)

$$
\frac{dV_b}{d\tau_d} \propto \frac{d(\pi_b^{1-\gamma})}{d\tau_d} (\pi_s^{1-\gamma} + \mu \pi_b^{1-\gamma})^{-\gamma} \left( 1 - \gamma \frac{\mu \pi_b^{1-\gamma}}{\pi_s^{1-\gamma} + \mu \pi_b^{1-\gamma}} \right) > 0
$$  \hspace{1cm} (C.38)

Proof of Proposition 5.

$$
W = \gamma \left( E[\pi_s]^{1-\gamma} + \mu E[\pi_b]^{1-\gamma} \right)^{1-\gamma} \text{ and } V = (1 - \gamma) \left( E[\pi_s]^{1-\gamma} + \mu E[\pi_b]^{1-\gamma} \right)^{1-\gamma},
$$

therefore, it is sufficient to show that $\left( E[\pi_s]^{1-\gamma} + \mu E[\pi_b]^{1-\gamma} \right)^{1-\gamma}$ is increasing in $\tau_d$. This is indeed the case since $E[\pi_s] = \pi''$ is constant and $E[\pi_b]$ is increasing in $\tau_d$. 

Proof of Lemma 3.

Firms’ monopoly prices are a constant markup over their marginal cost: \( p^m_i = \frac{\phi_i}{\theta} \), and the corresponding monopoly quantity is \( x^m_i = \left( \frac{\phi_i}{\theta} \right)^{1/\gamma} \), which generates consumption

\[
c_i^m = (x^m_i) - \phi_i x^m_i = \left( \frac{\theta^2}{\phi_i} \right)^{1/\gamma} - \phi_i \left( \frac{\theta^2}{\phi_i} \right)^{1/\gamma} = \left( \frac{\theta^2}{\phi_i} \right)^{1/\gamma} \left( 1 - \theta^2 \right) \tag{C.39}
\]

(C.39) is decreasing in marginal cost \( \phi_i \), therefore the lower the marginal cost \( \phi_i \), the more consumption each product generates. As \( \phi^L < \phi < \phi^H \), \( c^H_m < c^m_s < c^L_m \).

Proof of Proposition 6.

Recall that \( C = c_s^m l_s^\gamma + \mu (1 - \alpha \hat{\sigma}^2) c_b^L l_b^\gamma + \mu \alpha \hat{\sigma}^2 c_b^H l_b^\gamma \).

\[
\frac{dC}{d\tau_d} = \gamma c_s^m l_s^\gamma - \frac{dl_s}{d\tau_d} + \gamma \mu (1 - \alpha \hat{\sigma}^2) c_b^L l_b^\gamma - \frac{dl_b}{d\tau_d} + \gamma \mu \alpha \hat{\sigma}^2 c_b^H l_b^\gamma - \frac{dl_b}{d\tau_d}
\]

\[
\underbrace{(A')}_{(B') > 0} + \mu c_b^L l_b^\gamma - \frac{dl_b}{d\tau_d} - \mu c_b^H l_b^\gamma - \frac{dl_b}{d\tau_d}
\]

\[
\text{The last two terms (marked as (B)) reflect the redistribution of innovation from high cost to low cost large firms, and its always positive since } c_b^L > c_b^H. \text{ Consider the first three terms, since } l_s = 1 - \mu l_b, \text{ these three terms can be rewritten as follows}
\]

\[
(A') = \gamma c_s^m l_s^\gamma - \frac{dl_s}{d\tau_d} + \gamma \mu (1 - \alpha \hat{\sigma}^2) c_b^L l_b^\gamma - \frac{dl_b}{d\tau_d} + \gamma \mu \alpha \hat{\sigma}^2 c_b^H l_b^\gamma - \frac{dl_b}{d\tau_d}
\]

\[
\propto -c_s^m l_s^\gamma - \frac{dl_s}{d\tau_d} + (1 - \alpha \hat{\sigma}^2) c_b^L l_b^\gamma - \frac{dl_b}{d\tau_d} + \alpha \hat{\sigma}^2 c_b^H l_b^\gamma - \frac{dl_b}{d\tau_d}
\]

\[
\propto -c_s^m l_s^\gamma + E[c_b^m l_b^\gamma - \frac{dl_b}{d\tau_d}]
\]
(C.43) > 0 if and only if \( \frac{E[c_\theta^n]}{c_\theta^n} > \left( \frac{l_b}{\Gamma} \right)^{1-\gamma} = \frac{E[\pi_b]}{\pi_s} \). But \( \frac{E[c_\theta^n]}{c_\theta^n} = (1-\alpha \hat{\sigma}^2) \phi_L \frac{\pi_b}{\pi_L} + \alpha \hat{\sigma}^2 \phi_H \frac{\pi_b}{\pi_H} = \frac{E[\pi_b]}{\pi_s} \). Therefore, (A) = 0 and \( \frac{dC}{dt_d} > 0 \).

**Proof of Proposition 7**

Inspection of equation (C.5) and (C.7) shows that given certain values of expected profits \( E[\pi_s] \) and \( E[\pi_b] \), the proof of Proposition 1 critically relies on \( \frac{dE[\pi_s]}{dt_d} < 0 \) and \( \frac{dE[\pi_b]}{dt_d} > 0 \). If these two inequalities hold, Proposition 1 continues to hold, as the relations between \( E[\pi_s] \), \( E[\pi_b] \) and \( l_s, l_b \) are unaffected by the specific values of \( E[\pi_s] \) and \( E[\pi_b] \).

\[
\frac{dE[\pi_s]}{dt_d} \propto \frac{d\hat{\sigma}^2}{dt_d} (\pi_s^{\phi_H} - \pi_s^{\phi_L}) \tag{C.44}
\]

\[
\frac{dE[\pi_b]}{dt_d} \propto \frac{d\hat{\sigma}^2}{dt_d} (\pi_b^{\phi_H} - \pi_b^{\phi_L}) \tag{C.45}
\]

Therefore, to show that \( \frac{dE[\pi_s]}{dt_d} < 0 \) and \( \frac{dE[\pi_b]}{dt_d} > 0 \), I need to show \( \pi_s^{\phi_H} > \pi_s^{\phi_L} \) and \( \pi_b^{\phi_H} < \pi_b^{\phi_L} \), a sufficient condition for which is \( \frac{d\pi_s}{d\phi_b} > 0 \) and \( \frac{d\pi_b}{d\phi_b} < 0 \). Taking derivative of the equilibrium profits with respect to \( \phi_b \),

\[
\frac{d\pi_b}{d\phi_b} = \frac{\partial \pi_b}{\partial (x_s + x_b)} \frac{d(x_s + x_b)}{d\phi_b} + \frac{\partial \pi_b}{\partial x_b} \frac{dx_b}{d\phi_b} < 0. \tag{C.46}
\]

where \( \frac{\partial \pi_b}{\partial (x_s + x_b)} > 0 \) and \( \frac{\partial \pi_b}{\partial x_b} > 0 \) are obtained from equation (B.14); \( \frac{d(x_s+x_b)}{d\phi_b} < 0 \) is by equation (B.17); \( \frac{dx_b}{d\phi_b} < 0 \) can be seen from equation (B.19), where the numerator is decreasing in \( \phi_b \) and the denominator is increasing in \( \phi_b \).
\[
\frac{d\pi_s}{d\phi_b} \propto 2(\phi_b - \phi\theta)(\phi + \phi_b)\frac{2 - \theta}{\theta - 1} + (\phi_b - \phi\theta)^2 \frac{2 - \theta}{\theta - 1} (\phi + \phi_b)^{\frac{2 - 2\theta}{\theta - 1}}, \\
\propto 2 + (\phi_b - \phi\theta)\frac{2 - \theta}{\theta - 1} (\phi + \phi_b)^{-1}.
\]

(C.47) (C.48)

For (C.48) to be greater than 0, I effectively need to show

\[\frac{2}{\phi_b - \phi\theta} > \frac{2 - \theta}{1 - \theta \phi + \phi_b} \cdot\]

(C.49)

If \(\phi_b > \phi\),

\[
LHS > \frac{2}{\phi - \phi\theta} = \frac{2}{\phi(1 - \theta)} > \frac{2 - \theta}{1 - \theta \phi + \phi_b} = RHS;
\]

(C.50)

if \(\phi_b < \phi\),

\[
LHS > \frac{2}{\phi - \phi\theta} = \frac{2}{\phi(1 - \theta)} > \frac{2 - \theta}{1 - \theta \phi + \phi_b} = RHS.
\]

(C.51)

Therefore \(\frac{d\pi_s}{d\phi_b} > 0\).

In summary, \(\frac{d\pi_s}{d\phi_b} > 0\), \(\frac{d\pi_b}{d\phi_b} < 0\) imply \(\frac{dE[\pi_s]}{dr^d} < 0\), \(\frac{dE[\pi_s]}{dr^d} > 0\), and subsequently Proposition 1 holds.

Proof of Proposition 8

Equation (C.13 - C.14) in the proof of Proposition 2 are based on expected product market profits and thus still hold. The only difference with the Cournot competition case is that equation (B.22) and (B.23) are substituted in as the expressions of expected profits \(E[\pi_s]\) and \(E[\pi_b]\). More specifically, aggregate innovation expenditure is increasing (decreasing) in disclosure if the
following expression is greater (less) than 0,

\[-E[\pi_s]^{\gamma} (\pi_s^{\phi_H} - \pi_s^{\phi_L}) - E[\pi_b]^{\gamma} (\pi_b^{\phi_H} - \pi_b^{\phi_L})\]

where, by Proposition 7, \(\pi_s^{\phi_H} - \pi_s^{\phi_L} > 0\), and \(\pi_b^{\phi_H} - \pi_b^{\phi_L} < 0\). Moreover, because \(\frac{dE[\pi_s]}{d\tau_d} < 0\) and \(\frac{dE[\pi_b]}{d\tau_d} > 0\), expression (C.52) is increasing in \(\tau_d\) (or equivalently, decreasing in \(\hat{\sigma}^2\)).

The same logic of proof for Proposition 2 also applies here. Namely, if \(\hat{\sigma}^2_{o}\) is sufficiently low such that \(\text{(C.52)} > 0\) holds at \(\hat{\sigma}^2 = \hat{\sigma}^2_{o}\), then \(\frac{dV}{d\tau_d} > 0\), \(\frac{dWL}{d\tau_d} > 0\) hold for all \(\tau_d\).

As is shown in the proof of Proposition 7 that \(\frac{d\pi_s}{d\phi_b} > 0\) and \(\frac{d\pi_b}{d\phi_b} < 0\), it is straightforward that \(\frac{dE[\pi_s]}{d\phi_b} > 0\) and \(\frac{dE[\pi_b]}{d\phi_b} < 0\), because \(E[\pi_i]\) are linear combinations of \(\pi_i\)'s. Therefore, expression (C.52) is decreasing in \(\phi_b\). Expressing this condition in terms of \(\phi_L\), \(\frac{dV}{d\tau_d} < 0\), \(\frac{dWL}{d\tau_d} < 0\), \(\forall \tau_d\) if \(\phi_L < \phi^*L\), where \(\phi^*L\) is such that (C.52) just holds.

At the other extreme, if expression (C.52) < 0 at \(\hat{\sigma}^2 = 0\), then \(\frac{dV}{d\tau_d} < 0\), \(\frac{dWL}{d\tau_d} < 0\) hold for all \(\tau_d\). Namely, if

\[-\left((1 - \mu)\pi_s^m + \mu \pi_s^{\phi_L}\right)^{\gamma} (\pi_s^{\phi_H} - \pi_s^{\phi_L}) - \left(\pi_b^{\phi_H} - \pi_b^{\phi_L}\right) < 0\]  \hspace{1cm} (C.53)

\(\frac{dV}{d\tau_d} < 0\), \(\frac{dWL}{d\tau_d} < 0\). Again, because (C.52) is decreasing in \(\phi_L\), \(\frac{dV}{d\tau_d} < 0\), \(\frac{dWL}{d\tau_d} < 0\), \(\forall \tau_d\) if \(\phi_L > \phi^{*L}\), where \(\phi^{*L}\) is such that (C.53) just holds.

Finally, if \(\phi_L \in [\phi^*L, \phi^{*L}]\), then (C.52) < 0 at \(\hat{\sigma}^2 = \hat{\sigma}^2_{o}\), but (C.52) > 0 for some \(\tau_d > \tau^*_d\), then \(V\) and \(WL\) first decrease for \(\tau_d \in [0, \tau^*_d]\) then increase for \(\tau_d \in [\tau^*_d, +\infty]\).