Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Essays in Industrial Organization

Autor:

Ramiro de Elejalde

Director:

Jordi Jaumandreu

DEPARTAMENTO DE ECONOMIA

Getafe, Noviembre de 2015

Universidad Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Essays in Industrial Organization

Autor: ***Ramiro de Elejalde***

Director: **Jordi Jaumandreu**

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Secretario:

Calificación:

Getafe, de de

A Nieves, Teo, Marta Elena y Willie

Agradecimientos

Esta tesis no hubiese existido sin la ayuda de Jordi. Gracias por el tiempo compartido en reuniones sin límite de tiempo, por el apoyo cuando no veía claro hacia donde iba la tesis y por el apoyo moral en momentos de crisis. Gracias por dar el ejemplo con una impecable ética de trabajo.

A Nieves, gracias por todo. Por aguantar mis silencios cuando las cosas no salían, por aguantar los fines de semana trabajando en los papers, la distancia de los viajes a Boston, muchas gracias. Gracias por Teo, lo más hermoso del mundo.

A mis viejos, Marta Elena y Willie. Gracias por al apoyo incondicional y creer que su hijo estaba tomando la decisión correcta. Gracias por estar siempre ahí en la buenas y en las malas.

A mis hermanas, Vicky y Guchi. Gracias por poner el hombro cuando lo necesitaba aunque sea para escuchar un llamado por teléfono. Gracias por aguantar mi mal humor y rabietas, y por todo el amor fraternal.

A mis amigos del doctorado. A Zoe, Ainhoa, Fernando y Roberto. Porque fueron el sostén en el primer año, porque después de todo lo que pasamos y sobrevivimos ese año fue como si nos conociéramos de toda la vida. A Rodo, compañeros en varias batallas como colega, coautor y amigo. Gracias por ser tan generoso con tu tiempo y esfuerzo sin esperar nada a cambio. A Charly Pérez. Gracias por el apoyo, las conversaciones académicas y las no tan académicas. Gracias por las cervezas y mostrarme los mejores bares de Madrid. A Juancito, gracias por alojarme y acompañarme en tiempos complicados. A Gastón gracias por la generosidad intelectual para compartir ideas y estar dispuesto a explicar cuestiones que no eran obvias para mi. A Dani, Joaco, y Dolo gracias por el apoyo y los momentos compartidos. A la Lu, gracias por sentirme acompañado en la última batalla que parecía interminable.

A Eugenio gracias por el empuje, los retos para terminar los papers, y las discusiones a muerte por entender el problema en forma diferente. Gracias por enseñarme a ver el trabajo empírico con otros ojos. A Carlos gracias por el apoyo para terminar la tesis, las conversaciones en el patio de la facultad sobre la investigación y, más importante, sobre de la vida. Gracias por el aprecio que se percibe y es mutuo. Creo que me he convertido en mejor economista gracias a ustedes.

A Andrés, Julio, Matilde y Natalia gracias por el último empujón cuando me quedaban pocas fuerzas para llegar a la meta.

A todos. Muchas gracias.

Summary

In this thesis, I study two important issues in the field of industrial organization: competition between firms with different business models and its effects on market allocation, and the relation between innovative activity and employment at the firm level.

The first chapter entitled “Employment and innovation: Firm level evidence from Argentina” is a joint work with David Giuliodori and Rodolfo Stucchi. This chapter provides evidence about the effect of innovation on employment in Argentina in the period 1998-2001. In particular, we quantify the effect of process and product innovations on employment growth and the skill composition. Our results show that: (i) Product innovations have a positive effect on employment growth biased toward skill labor; (ii) Process innovations do not affect employment growth or composition; (iii) There are no heterogeneous effects in technology intensity and size, (iv) Most of the contraction in employment in this period was explained by noninnovators.

The second chapter entitled “Industry equilibrium with open-source and proprietary firms” is a joint work with Gastón Llanes. This chapter presents a model of industry equilibrium to study the coexistence of open-source and proprietary firms. Two novel aspects of the model are (i) participation in open source arises as the optimal decision of profit-maximizing firms, and (ii) open-source and proprietary firms may (or may not) coexist in equilibrium. Firms decide their type and investment in R&D, and sell packages composed of a primary good and a complementary private good. Open-source firms share their technological advances on the primary good, whereas

proprietary firms keep their innovations private. The main contribution of the chapter is to determine conditions under which open-source and proprietary firms coexist in equilibrium. Interestingly, this equilibrium is characterized by an asymmetric market structure, with few large proprietary firms and many small open-source firms.

The third chapter entitled "Competition between single-market and multimarket banks: Evidence from the US Banking Industry" estimates comparative advantages and disadvantages between single-market and multimarket banks. In the banking industry, due to the relevance of the branch's location, the relevant market is a local market. Based on this fact, the chapter defines two types of competitors: banks with branches in only one local market called *single-market banks*, and banks with branches in multiple local markets called *multimarket banks*.

The main results of the paper are: (i) incumbent single-market banks have a comparative advantage with multimarket banks in a sample of small urban and rural markets, and (ii) single-market banks face higher expansion costs into other local markets than mutlimarket banks. This chapter presents two main contributions. First, it contributes to the discussion about agency problems for hierarchical firms in markets with asymmetric information. Second, it suggests the possibility of unexpected consequences after a deregulation of the expansion restrictions that lesser the barrier to entry for multimarket banks.

Resumen

En esta tesis se estudian dos problemas relevantes en el campo de la organización industrial: competencia entre firmas con distintos modelos de negocios (firmas "open source" y propietarias en el capítulo 2, y firmas multimercado y unimercado en el capítulo 3) y sus efectos sobre el mercado como asignador de recursos, y la relación empírica entre actividad innovadora y empleo al nivel de firma.

El primer capítulo llamado "Empleo e innovación: Evidencia al nivel de firma para Argentina" es un trabajo conjunto con David Giuliodori y Rodolfo Stucchi. Este capítulo muestra evidencia del efecto de la innovación en el empleo para Argentina en el período 1998-2000. En particular, se mide el efecto de innovación en proceso y producto en la tasa de crecimiento del empleo y la composición de la cualificación del empleo. Los resultados del capítulo muestran que: (i) Las innovaciones de producto tienen un efecto positivo en el crecimiento del empleo sesgado hacia el trabajo calificado; (ii) las innovaciones de proceso no afectan el crecimiento del empleo o las composición de la cualificación del empleo; (iii) no hay efectos heterogéneos en función del sector tecnológico al que pertenece firma y el tamaño de la firma; (iv) La mayor parte de la contracción del empleo en el período se explica por la decisiones de contratación de la firmas no innovadoras.

El segundo capítulo llamado "Equilibrio en la industria con firmas "open source" y propietarias" es un trabajo conjunto con Gastón Llanes. En este capítulo se presenta un modelo de equilibrio en una industria para estudiar la coexistencia entre firmas "open source" y propietarias. El modelo presenta dos aspectos novedosos (i) la parti–

cipación en un proyecto open source surge endógenamente a partir de las decisiones óptimas de firmas maximizadoras de beneficios, y (ii) las firmas open source y propietarias pueden coexistir o no en equilibrio. Las firmas deciden su tipo y su inversión en I+D, y venden paquetes conformados por un bien primario y un bien complementario privado. Las firmas open source comparten sus avances tecnológicos en el bien primario mientras que las firmas propietarias mantienen sus innovaciones privadas. La principal contribución del capítulo es determinar las condiciones que permiten la coexistencia de ambas firmas en equilibrio. Un resultado importante es que el equilibrio se caracteriza por una estructura asimétrica de mercado con un número reducido de firmas propietarias de gran tamaño y numerosas firmas open source de menor tamaño.

El tercer capítulo llamado "Competencia entre bancos unimercado y multimercado: Evidencia para la industria bancaria de EEUU" estima las ventajas y desventajas comparativas entre bancos unimercado y multimercado. El mercado relevante en la industria bancaria de EEUU, debido a la importancia de la distancia geográfica en la industria, es esencialmente local. El trabajo considera dos tipos de competidores en función a la localización de las sucursales bancarias: bancos con presencia en un sólo mercado local llamados *bancos unimercado*, y bancos con presencia en más de un mercado local llamados *bancos multimercado*.

Los resultados principales del trabajo son: (i) los bancos unimercado incumbentes tienen ventajas comparativas frente a los bancos multimercados en mercados urbanos de menor tamaño y en mercados rurales, y (ii) los bancos unimercados enfrentan mayores costos de expansión hacia otros mercados locales que los bancos multimercado. Este capítulo presenta dos contribuciones principales. Primero, contribuye a la discusión sobre los problemas de agencia de firmas más jerárquicas en mercados con problemas de información asimétrica. Segundo, sugiere la posibilidad de consecuencias no esperadas de una política de desregulación a la expansión de las sucursales bancarias que disminuye la barreras a la entrada de los bancos multimercado.

Contents

List of Figures

List of Tables

Chapter 1

Employment and innovation: Firm level evidence from Argentina (with David Giuliodori and Rodolfo Stucchi)

1.1 Introduction

The effect of innovation on employment has been attracting the attention of economists and policy makers for a long time. This fact is not surprising. On the one hand, it has been argued that technical change could destroy jobs and, on the other hand, economic theory does not provide a clear answer about the employment effect of innovation. The relationship between innovation and employment is not straightforward. The literature has documented several compensation mechanisms that can counterbalance the initial effect of innovation and render the final effect undetermined (see Petit (1995); Pianta (2005); Piva and Vivarelli (2005); Vivarelli (1995, 2014)). Innovation can create or destroy jobs depending on the institutional setting, market structure, and the type of innovation the firm introduces.

The development –or the adoption– a new production process usually leads to greater efficiency in production, with savings in labor and/or capital and with a potential for price reduction. The first expected outcome is higher productivity with loss of employment. However, demand could grow after the innovation due to increased quality or lower price, and this increase in demand could lead to higher employment.

The introduction of a new or significantly improved product increases employment via an increase in demand. However, if after the innovation the innovator enjoys market power, it can set prices that maximize its profits but imply a reduction in output. Therefore, the net effect of a product innovation could be a contraction in employment. A new product can also destroy jobs if it is designed to reduce costs. It is also possible that product innovations do not change employment; this would be the case if new products replace old products without changes in demand.

Despite the fact that the theoretical effect of innovation on employment is ambiguous, several firm level studies have found that the fear that innovation could destroy jobs has little empirical support. In fact, the evidence shows a positive relationship between innovation and employment (Blanchflower and Burgess (1998); Entorf and Pohlmeier (1990); Giuliodori and Stucchi (2012); Piva and Vivarelli (2005); Smolny (1998); Van Reenen (1997)).

Firm-level evidence also suggests that while product innovation creates jobs, process innovation might in fact destroy jobs. To capture this idea, Harrison, Jaumandreu, Mairesse, and Peters (2014) pose a simple model to study the differential effect of product and process innovation on employment growth. They estimate their model for the manufacturing and service sectors in France, Germany, Spain, and the United Kingdom. They find that the increase in employment due to product innovations is large enough to compensate the negative effect of process innovations. The results are similar across countries, although there emerge some interesting differences. In fact, they find no evidence for a displacement effect of process innovation in Spanish manufacturing firms. They argue that this result can be explained by a greater pass-through of productivity improvements in lower prices. Hall, Lotti, and Mairesse (2008) estimate

Harrison, Jaumandreu, Mairesse, and Peters (2014)'s model using Italian data and find similar results.

Innovation affects not only the number of employees but also the composition of employment within each firm. The basic intuition is that innovations are skill biased because they replace tasks traditionally carried out by unskilled workers, with new jobs demanding qualified workers. Acemoglu (1998) argues that new technologies are not complementary to skills by nature, but innovators decide the direction of technological change. Acemoglu shows that an increase in the supply of skills can explain skill-biased technical change in the United States. More relevant to our article, Acemoglu (2003) extends the basic model of directed technical change to study the interaction between technology and international trade. In his model, technical change in developing countries is skill biased due to the transfer of technology from developed countries.

There is a vast empirical literature on the skill bias of technological change for developed countries. After the seminal work by Griliches (1969), the effect of innovation on the skills composition has been largely studied (Autor, Katz, and Krueger (1998); Bresnahan, Brynjolfsson, and Hitt (2002); Caroli and Van Reenen (2001); Doms, Dunne, and Troske (1997); Greenan (2003)). Giuliodori and Stucchi (2012) analyze a related question about the effect of innovation on the composition of employment in terms of labor contract. They present evidence for Spain, where the labor market is segmented in temporary and permanent contracts, and find that innovations can affect both types of contracts depending on the institutional environment.

Firm-level studies discussed above focus on the direct effect of innovation on employment –that is, the effect of innovation on the level of employment of the innovating firm. Innovation also has indirect effects –that is, on noninnovating firms. The indirect effects are intuitive for product innovations; it is not difficult to imagine a scenario in which a product innovation increases the demand of the innovating firm and its employment but reduces the demand of its competitors and their employment level. Process innovation also has indirect effects. The innovating firm can increase its productivity and by reducing price can gain market share, increase its demand for labor, and reduce the demand of labor of competitors. Pianta (2005) reviews several industry-level studies addressing these issues. The empirical evidence reviewed by Pianta (2005) shows that the effect of product innovation on employment is positive in industries characterized by high demand growth and an orientation toward product (or service) innovation, while process innovation leads to job losses. The evidence about overall effect is mixed; it depends on the country and period considered.

The evidence on the relationship between innovation and employment in Latin America is scarce, and because of the idiosyncratic nature of innovation in Latin America –mainly acquisition of technological knowledge from abroad– the evidence from developed countries cannot be simply extrapolated to this region. In addition, in Latin America there are important structural features that might lead to different outcomes of innovation on employment. First, the current production structure is strongly dominated by small- and medium-sized enterprises (SMEs). Second, Latin America's production structure is heavily dominated by the manufacturing of commodities and low technologically intensive goods. The available evidence comes from Benavente and Lauterbach (2008)), who estimate the Harrison, Jaumandreu, Mairesse, and Peters (2014) model for Chile. Contemporaneous to our article, Aboal, Garda, Lanzilotta, and Perera (2015), Alvarez, Benavente, Campusano, and Cuevas (2011), and Monge-Gonzalez, Rodriguez-Álvarez, Hewitt, Orozco, and Ruiz (2011) estimate Harrison, Jaumandreu, Mairesse, and Peters (2014) 's model for Chile, Costa Rica, and Uruguay, and Crespi and Tacsir (2013) present a comparative analysis for the several countries. They find that while product innovations increase employment, process innovations do not affect it. Additional evidence on the relation between innovation and employment comes from the evaluation of innovation public policies. Alvarez, Benavente, Campusano, and Cuevas (2012), who evaluate the effect of two innovation programs (FONTEC and FONDEF) in Chile, find that these programs increased employment and productivity. Castillo, Maffioli, Rojo, and Stucchi (2014) evaluate the effect of the Support Program for the Organizational Change in Argentina. They find that while support for both process and product innovation lead to increased employment, the support for product innovation has a higher effect on wages, survival rate, and exporting probability.

This article aims at providing evidence about the relationship between innovation and employment in the manufacturing sector in Argentina. More precisely, we aim at answering two important questions: (1) How do different types of innovation – product and process innovations– affect employment? (2) How do the different types of innovation affect the skill composition? In addition, given the Argentinean production structure, we are interested in knowing whether the results vary between low- and high-tech industries or small and large firms. To answer those questions, we use data from Innovation Surveys for Argentina for the period 1998-2001. This period coincides with one of the deepest recessions of Argentina's history. As a consequence, it provides an interesting opportunity to estimate the effects of innovation on employment growth in a highly recessive scenario. A few statistics help to describe the rough economic environment in Argentina in 1998-2001. The gross domestic product fell 8.4 percent between 1998 and 2001 (a negative average growth rate of 2.9 percent per year). Unemployment rose from 13.2 percent in May 1998 to 18.3 percent in October 2001. Investment plunged at an average annual rate of 12 percent. The manufacturing sector showed a significant contraction; the index of manufacturing activity fell 22 percent between 1998 and 2001. The innovation survey we use in the analysis shows a contraction in employment of 8 percent between 1998 and 2001. Interestingly, the reduction in employment was different between innovators and noninnovators. The reduction in employment was 7 percent in those firms that reported process or product innovations and 13 percent in firms that did not introduce any innovation.

We find that while product innovation creates jobs, process innovation does not affect the level of employment. Another important finding is that product innovation is skilled biased. In fact, we find that while product innovation creates both skilled and unskilled jobs, it creates a higher proportion of skilled jobs. In the case of process innovation, we find that there is no effect on skilled or unskilled jobs. These findings provide evidence not only against the fear that innovation could destroy jobs but also against the hypothesis that innovation could destroy unskilled jobs. From a policy perspective, our results shed light on two important issues. First, Argentina – like the rest of Latin American economies – faces a productivity problem that calls for attention. Our results point out that innovation programs – whose main objective is to increase productivity – could be attractive also from an employment point of view. Some warnings on SME policies that distort the size distribution of firms have been raised because they could affect the aggregate level of productivity (Guner, Ventura, and Xu (2008); Pagés (2010)). Interestingly, this is not the case of innovation policies because innovation is one of the main drivers of productivity growth. Second, the complementarity of innovation and skilled workers justify the need for training programs in addition to innovation programs.

The rest of the article is organized as follows. Section 1.2 presents the analytical framework. Section 1.3 presents the data and some descriptive statistics. Section 1.4 presents the main results of innovation on employment. Section 1.5 discusses the robustness of our results. Section 1.6 decomposes the contribution of product innovators, process innovators and noninnovators to employment growth. Finally, Section 1.7 concludes.

1.2 Analytical Framework

The analytical framework follows the model in Harrison, Jaumandreu, Mairesse, and Peters (2014). In this framework, a firm produces two types of products in period t: old or only marginally modified products ("old products" denoted by Y_{1t}) and new or significantly improved products ("new products" denoted by Y_{2t}). Assuming separability in the production of old and new products, the production function for product of type i in period t can be written as

$$Y_{it} = \theta_{it} F(L_{it}, K_{it}, M_{it}) e^{\eta + \omega_{it}},$$

where $F(.)$ is homogeneous of degree one in labor (L_{it}), capital (K_{it}), and intermediate goods (M_{it}); θ_{it} is a Hicks neutral technical change parameter (which can depend on process innovation); and $e^{\eta+\omega_{it}}$ is unobserved firm's productivity that can be decomposed in firm's attributes that are mainly time invariant (η) and productivity shocks (ω_{it}).

Under perfect competition in input markets, the cost function of a firm in period t is

$$C_t(w_t, Y_{1t}, Y_{2t}) = c(w_t)\left(\frac{Y_{1t}}{\theta_{1t} e^{\eta+\omega_{1t}}} + \frac{Y_{2t}}{\theta_{2t} e^{\eta+\omega_{2t}}}\right),$$

where w_t are input prices, and the conditional labor demand function is

$$L_{it} = c_{w_L}(w_t)\frac{Y_{it}}{\theta_{it} e^{\eta+\omega_{it}}},$$

where w_L is the price of labor, and $c_{w_L} = \partial c / \partial w_L$.

Using the labor demand function, we can approximate employment growth at the firm level as

$$\frac{\Delta L}{L} \approx \log\left(\frac{L_{12}}{L_{11}}\right) + \frac{L_{22}}{L_{11}} = -(\log\theta_{12} - \log\theta_{11}) + (\log Y_{12} - \log Y_{11}) + \frac{\theta_{11}}{\theta_{22}}\frac{Y_{22}}{Y_{11}} - (\omega_{12} - \omega_{11})$$

Employment growth is then decomposed into the part due to the increased efficiency in production of old products (which could be related to process innovations), the part due to sales of old products, and the part due to the introduction of new products. The estimating equation is given by

$$l = \alpha_0 + \alpha_1 d + g_1 + \beta g_2 + v, \qquad (1.2.1)$$

where l is total employment growth; g_1 is the nominal growth in sales of old products; g_2 is the nominal growth in sales of new products (product innovations); and d captures the introduction of process innovations in the production of old products.

The parameter β captures the relative efficiency in the production of old and new products: when $\beta < 1$ ($\beta > 1$) new products are produced more (less) efficiently than old products. The constant in Equation (1.2.1) represents (minus) the average efficiency growth in the production of old products for noninnovators.

We observe employment and total sales in 1998 and 2001, and firms report whether they introduced product or process innovations between those years. This is important because it provides us with information before and after the innovation. Moreover, in 2001, it is possible to know the percentage of sales corresponding to new products. This information is crucial to estimate Equation (1.2.1).

The effect of innovations on employment composition is estimated with a version of Equation (1.2.1) for employment growth of skilled and unskilled workers. For the two types of workers, skilled (s) and unskilled (u), we estimate

$$l^q = \alpha_0^q + \alpha_1^q d + g_1 + \beta^q g_2 + v^q \quad q = s, u, \qquad (1.2.2)$$

where l^q is the growth rate of employment of type q .

A concern about the identification[1] of the coefficients in Equation (1.2.1) is the fact that innovation can be correlated with the error term, then ordinary least squares (OLS) can produce inconsistent estimates. The endogeneity of innovation comes from the fact that productivity is omitted from Equation (1.2.1), and it can be correlated

[1] The identification discussion focuses on Equation (1.2.1), but similar arguments apply to the identification of Equation (1.2.2).

with innovation. This is the case because innovations are the result of investment decisions, such as R&D, and those decisions depend on the firm's productivity. Then, if productivity is in the error term because it is an omitted variable, the error term will be correlated with innovation leading to an endogeneity problem.

In order to better understand the endogeneity problem, it is useful to decompose productivity in two unobserved components: firm's attributes that are mainly time invariant (such as managerial skills or organizational capital) and productivity shocks (which might lead the firm to reduce labor costs). Equation (1.2.1) is specified as a growth equation, and the influence of the time invariant part of productivity is removed from the error term.

The remaining source of correlation between innovation outputs and productivity are productivity shocks.

Part of the correlation between innovation and productivity shocks is the relationship between these variables and the business cycle. If both innovation and productivity are related to the business cycle as some literature has found –see, for example, Barlevy (2007) for innovation and Basu and Fernald (2001) for productivity– then endogeneity is a valid concern. To avoid this source of correlation, we include a set of industry dummies in the growth Equation (1.2.1). A set of industry dummies in Equation (1.2.1) is equivalent to the interaction between industry dummies and a 2001 dummy in a level equation. Therefore, these variables will capture the business cycle effect.

Once we control for time-invariant unobservables and industry-specific temporal shocks, there are good reasons to think that process innovation can be exogenous.[2] First, innovations expenditures are usually made well before they result in applicable innovations. Second, it seems realistic to assume that firms cannot predict future labor problems, supply chain disruptions, or demand shocks when deciding their innovations

[2]It should be noted that similar arguments can apply for product innovations. However, as we argue below, our measure of product innovation is measured with error. This additional source of endogeneity in the product innovation variable is not present in the process innovation variable.

expenditures. For these reasons, we treat process innovation as exogenous, but we run a robustness exercise instrumenting the process innovation variable.

Another possible source of endogeneity is the presence of measurement error in g_1 and g_2 . Ideally, we would use growth in real production, but we only observe growth in nominal sales. Then the growth in the price of old and new products is in the error term, and the correlation between the growth in prices and g_2 can create an attenuation bias in the estimation of β. To deal with this measurement error problem, we follow Harrison, Jaumandreu, Mairesse, and Peters (2014) . First, we use industry price indexes π as a proxy for the growth in prices of old products. Second, we use instrumental variables that are correlated with real growth in the production of new products but uncorrelated with its nominal growth.

The main advantage of Harrison, Jaumandreu, Mairesse, and Peters (2014)'s approach is the use of economic theory to model the possible mechanisms linking innovation and employment. In particular, under the assumption of perfect competition in input markets and general assumptions about the firm's technology, Harrison, Jaumandreu, Mairesse, and Peters (2014) derive an estimating equation linking process and product innovation with optimal hiring decisions. This, in turn, leads to parameters with a clear and meaningful economic interpretation. In comparison, a reduced form approach that regresses employment growth on innovative variables is difficult to interpret, and the mechanisms are difficult to disentangle. We illustrate this point by running a reduced form approach in Table 1.2 and remark how difficult it is to interpret the coefficients in that model.

Another advantage is that Harrison, Jaumandreu, Mairesse, and Peters (2014) estimate their model for several European countries that can serve as a benchmark for the effects of innovation on employment in developing countries. In that sense, we can not only interpret the evidence for Argentina, but also compare it with the evidence for developed countries.

An important point concerns the identification strategy and how it compares with alternative approaches. An alternative identification strategy used in the literature is a generalized method of moments (GMM) system estimator proposed in Blundell and Bond (1998). For a recent application, see Lachenmaier and Rottmann (2011). The basic approach is to specify a labor demand equation that depends on lags of the dependent variable, contemporaneous and lags of the innovation variables, industry controls, and a firm level fixed effect. Such a model is estimated using lagged level variables to instrument the difference equation and lagged differences to instrument the level equation. In this article, identification relies on contemporaneous level variables to instrument the difference equation controlling for industry and location dummies. Unfortunately, the data available consist of two periods, and this prevents us from including lagged variables as instruments. For that reason, we run several exercises to show that the results in the article are robust to departures from the basic assumptions such as instrument validity, alternative controls, and exogeneity of the process innovation.

1.3 Data and Descriptive Statistics

We use data from the Second National Innovation Survey (ENIT01).[3] ENIT01 was conducted in 2003 by the National Institute of Statistics and Censuses (INDEC) and collected retrospective information for each year between 1998 and 2001. The firms that were surveyed are the same firms surveyed in the Annual Industrial Survey – manufacturing firms with ten or more employees. The sample is representative of the manufacturing sector in the sense that the percentage of aggregate sales by industries in the sample is close to the percentage of sales by industry using the Annual Industrial Survey (Instituto Nacional de Estadísticas y Censos 2003).

The survey contains detailed information on firms' characteristics, innovative activity, and employment. Importantly, it also has detailed information on the compo-

[3]Segunda Encuesta Nacional de Innovación y Conducta Tecnológica de las Empresas Argentinas 1998-2001.

sition of sales, which allows us to compute the percentage of sales corresponding to new products. A firm in the survey reports the share of domestic sales (*pnd*) and the share of exports (*pxd*) that correspond to new or significantly improved products in 2001. Using that information, we construct sales of new products in 2001 as $Y_{22} = Domestic\,Sales\,2001\frac{pnd}{100} + Exports\,2001\frac{pxd}{100}$ and sales of old products in 2001 as $Y_{12} = Domestic\,Sales\,2001\frac{100-pnd}{100} + Exports\,2001\frac{100-pxd}{100}$. Using the above definitions, we can decompose the nominal growth in sales as $g = \left(\frac{Y_{12}+Y_{22}}{Y_{11}} - 1\right) 100 = \left[\left(\frac{Y_{12}}{Y_{11}} - 1\right) + \frac{Y_{22}}{Y_{11}}\right] 100 = g_1 + g_2$, where g_1 is the nominal growth in sales of old products, and g_2 is the nominal growth in sales of new products.

ENIT01 has also detailed information about the composition of employment by educational level, which allows us to study the effect of innovation on skill composition. We define skilled workers as employees with a university degree or tertiary education (one- to three-year degree related to technical professions), and unskilled workers are employees with primary or secondary education. As usual with firm-level data, prices are not reported at the firm level, and we use industry price indexes at the two-digit level to deflate nominal variables. Given that product prices can differ between firms or even within the firm for multiproduct firms, the use of price indexes introduces a measurement error problem in the estimation. In the empirical implementation, we use instrumental variables to correct this measurement error bias.

We classify firms in mutually exclusive categories according to their innovative activity: product innovators, process-only innovators, and noninnovators. Product innovators are firms that introduce product innovations; process-only innovators are firms that introduce process innovations or organizational change innovations, excluding product innovators; and noninnovators are firms not classified as product or process innovators.[4]

[4]Following Harrison, Jaumandreu, Mairesse, and Peters (2014), we classify firms that have introduced both product and process innovations as product innovators. The implicit assumption is that product and process innovators are more similar to product innovators than to process innovators. We will present some evidence supporting this assumption in the next section.

Table 1.1 shows the descriptive statistics for the share of innovative firms, employment growth, sales growth, and labor productivity, where labor productivity is defined as real sales per worker. A large share of firms (63 percent) introduced at least one innovation in 1998-2001. Most of the innovators are product innovators (48 percent) rather than process-only innovators (15 percent). Harrison, Jaumandreu, Mairesse, and Peters (2014) report a similar share of innovative firms for France, Germany, Spain, and the United Kingdom. The higher ratio of R&D expenditure to sales in developed countries suggests, however, that innovative activities undertaken by firms in Argentina are different from those in developed countries. Innovative firms in Argentina aim more at assimilating foreign technology or consist of incremental, marginal innovations, while innovative firms in developed countries invest primarily in research and development.

Interestingly, the reduction in employment was different between innovators and noninnovators. The annual reduction in employment was 2.5 percent for product innovators and 3.9 percent for process-only innovators while it was 6 percent for noninnovators. A similar pattern is observed in sales growth with a smaller reduction in sales for innovators than for noninnovators. The annual reduction in sales was 6.6 percent for product innovators, 8.1 percent for process-only innovators, and 12.5 percent for noninnovators.

For product innovators we decompose growth in sales in the part corresponding to old products (g_1) and the part corresponding to new products (g_2), as explained above. It is remarkable the rapid pace at which product innovators substituted old products with new products during the analyzed period in Argentina: sales of old products decreased 45 percent while sales of new products increased 40 percent. This pace, especially the decrease in sales of old products, is significantly faster than that for France, Germany, Spain, and the United Kingdom reported in Harrison, Jaumandreu, Mairesse, and Peters (2014). This difference might be explained by the recessive scenario in 1998-2001 in Argentina or by new products presenting only incremental, marginal innovations with respect to old products.

The decrease in labor productivity was 4.2 percent for product innovators, 4.3 for process-only innovators, and 6.5 percent for noninnovators. This evidence suggests that innovators might be able to compensate a negative aggregate shock through the introduction of new products or processes. Table 1.1 shows the descriptive statistics for skilled and unskilled labor. In our sample, employment contracted at an annual rate of 4 percent. However, skilled employment decreased 1.6 percent while unskilled employment decreased 5.3 percent. Differences in skilled-unskilled labor growth rates were greater for innovators than for noninnovators suggesting complementarity between innovation and skilled labor.

We study the presence of heterogeneous effects for sectors with different technological intensity. The nature of innovations can be very different for low-tech and high-tech, and this can be reflected in the employment effects of innovations. Sectors are classified as low-tech or high-tech sectors following Czarnitzki and Thorwarth (2012), who study the productivity effects of basic research in low-tech and high-tech industries in Belgium. The low-tech sectors are food, beverages, and tobacco (ISIC 15 and 16); textiles, wearing apparel, and leather products (ISIC 17, 18, and 19); wood, wood products, and furniture (ISIC 20 and 36); pulp, paper, and paper products (ISIC 21); publishing, printing, and reproduction of recorded media (ISIC 22); rubber and plastic products (ISIC 25); and basic metals, fabricated metals, and nonmetallic mineral products (ISIC 26, 27, and 28). The high-tech sectors are chemicals and chemical products, coke, refined petroleum, and nuclear fuel (ISIC 23 and 24); machinery, equipment, office machinery, computers, communication equipment, electrical machinery, and medical, precision, and optical instruments (ISIC 29, 30, 31, 32, and 33); and motor vehicles and transport equipment (ISIC 34 and 35). Low-tech sectors have a lower share of skilled labor but greater growth in employment than high-tech sectors. However, the difference in employment growth for innovators and noninnovators is similar for low-tech and high-tech sectors.

We also study the presence of heterogeneous effects for firms with different sizes. In developing countries, and Latin America in particular, the share of small firms in manufacturing is important. Thus, it is relevant to study whether the effects of innovation on employment vary by firm size. Small firms are firms with fewer than fifty employees, and large firms are firms with more than fifty employees. The share of innovators is 44 percent for small firms and 72 percent for large firms (Table 1.1). However, the difference in employment growth for innovators and noninnovators is greater for small firms than for large firms. This suggests that heterogeneous effects may exist between small and large firms.

1.4 Empirical Results

Exploratory Regressions

Table 1.2 shows OLS exploratory regressions of employment growth on innovation variables, real growth in sales, industry and location dummies, and a foreign ownership dummy. We run these regressions for two reasons: first, and more important, to illustrate how challenging it is to understand the mechanisms linking innovation and employment without imposing additional structure; second, to justify grouping product-only innovators and product and process innovators.

Columns (1), (2), and (3) differ in the allocation of those firms that introduce both product and process innovations. In column (1), product and process innovators are included in separate categories; in column (2), product and process innovators are included with product innovators; and in column (3), product and process innovators are included with process innovators.

The estimate of the constant is approximately -2.2 percent, and it captures the mean employment growth for noninnovators. Innovators are associated with a 2 percent higher employment growth than noninnovators. There are not statistically significant differences between product and process-only innovators, product-only innovators, and product and process innovators. It will be difficult to estimate separate effects on employment for process-only innovators, product-only innovators, and product and process innovators. Thus, we decide to group all product innovators (product and process innovators and product-only innovators). This decision is supported by the point estimates shown in the table.

The estimated coefficient on real growth in sales suggests that sales are associated with a less-than-proportional increase in employment: a 1 percent increase in sales growth of old products implies a 0.32 percent increase in employment growth. As a comparison, Harrison, Jaumandreu, Mairesse, and Peters (2014) found elasticities between 0.35 and 0.45 for European countries.

Innovation and Employment

Column (1) in Table 1.3 shows the estimates of the effect of innovations on employment using the Harrison, Jaumandreu, Mairesse, and Peters (2014) model. In all the specifications, we control for two-digit industry dummies, location dummies,[5] and foreign ownership. Panel A in Table 1.3 shows the OLS estimates. These results show that while product innovation has a positive and significant effect on employment, process innovation does not have a significant effect. The estimated coefficient on g_2 is close to one, which indicates no differences in efficiency in the production of old and new products. Panel B in Table 1.3 shows the instrumental variable (IV) estimates. As we discussed in the Analytical Framework section, there are two endogeneity problems that can bias the OLS estimation: an omitted variable problem because productivity shocks are included in the error term (with a negative sign), and a measurement error problem due to unobservability of prices at firm level. These endogeneity issues tend to generate a downward bias in the OLS estimate of the coefficient on g_2 .

The instrument used in the IV estimation is an indicator of the firm knowledge of

[5] "Location dummies" means a dummy variable for each province in Argentina. We consider that a firm is located in a province if its headquarters are located in that province. There are twenty-three provinces in Argentina, and around 64 percent of the firms are located in the city of Buenos Aires.

public support for innovation activities. The identification strategy relies on knowledge of public programs being exogenous once we control for industry, location, size, and time-invariant productivity. We believe this is a valid assumption for several reasons. First, if information acquisition is costly, only more productive and larger firms will be willing to make such an investment. Given that we control for invariant productivity and size, these effects are taken into account. In addition, it seems less likely that firms decide to invest in information acquisition because of productivity shocks that could be temporary. Second, public innovation policies can be targeted to specific regions, industries, or sizes. In those cases, the information cost would vary at that level, and we control for that. Third, in 1998-2001 in Argentina, there were policy changes that can provide some exogenous shocks that we exploit in the estimation. In particular, the main innovation program in Argentina is FONTAR. In 1998, this program introduced a new source of financing in the form of fiscal subsidies applied to income taxes (Binelli and Maffioli (2007)). Another important innovation program in this period was PRE, which was created at the end of 1997. These programs targeted SMEs, and conditional on size, there were no additional requirements to bias the provision of information about the public programs (Castillo, Maffioli, Rojo, and Stucchi (2014)).

A valid instrument must also satisfy a relevance condition that requires significant correlation between the instrument and the endogenous variable. This condition can be tested with a joint significance test on the excluded exogenous variables in the first-stage regression. Stock, Wright, and Yogo (2002) recommend an F statistic greater than ten to avoid weak instruments problem that can create small sample bias in IV estimates. The first column shows that this F statistic is approximately fourteen, showing no evidence of weak instruments problem. In addition, given that just-identified models are better behaved in small samples, we are confident that the instrument satisfies the relevance condition and the estimates have good small-sample properties.

Table 1.3 shows that the IV estimates of the coefficient on g_2 move upward, which is consistent with a downward bias in the OLS estimate. The estimate increases from 0.96 in the OLS estimation in panel A to 1.15 in the IV estimation in panel B. A coefficient greater than one offers evidence that new products are produced less efficiently than old products. However, this evidence is tenuous because the estimate is not statistically different from one. These results show that there is evidence that product innovations create employment (creation effect) due to demand enlargement.

Table 1.3 also shows that the IV estimate of the coefficient on process innovation is also corrected upward. The estimated coefficient is positive but not significant, suggesting that process innovations have no effect on employment. There are two plausible explanations for this result. First, process innovations may not generate important productivity gains; hence, there is no displacement effect on employment. Second, process innovations may generate productivity gains (displacement effect), which induce a demand enlargement through market competition (creation effect). In the end, the creation effect on employment compensates the displacement effect on employment.

We run a Davidson-MacKinnon test to assess the endogeneity of g_2 . We reject exogeneity of g_2 at 10 percent. Thus our preferred specification for the innovation-employment model is the IV estimation where g_2 is endogenous.

Skill-Biased Innovations

The effect of innovation activity on skilled and unskilled labor is central for the design of public policy. If innovation activities and skilled labor are complements, we expect that the introduction of innovations will be mainly reflected in a higher demand for skilled labor. This can justify the implementation of labor-training programs simultaneously with pro-innovation policies. Columns (2) and (3) in Table 1.3 show the OLS and IV results for skilled and unskilled labor. Consistent with the expected downward bias in the OLS estimation, the IV estimates of the coefficients in g_2 and d are greater than the OLS estimates.

Interestingly, the IV results suggest that product innovations are more skilled intensive. The p-value of the test $H_0 : \beta^u = \beta^s$ vs. $H_1 : \beta^u \neq \beta^s$ is equal to 0.106. If the alternative hypothesis is that innovation is skilled biased–that is, $H_1 : \beta^u < \beta^s$–it is possible to reject the null hypothesis at 10 percent (p-value 0.053). There is no evidence that process innovations affect the skill composition.

It should be noted that we cannot reject exogeneity of g_2 in the case of unskilled labor. Given the difference in point estimates between OLS and IV, the test fails to reject exogeneity of g_2 because of the lack of precision in IV estimates. For this reason, we follow the more conservative approach of treating g_2 as endogenous in all the specifications.

Heterogeneous Effects by Technology Intensity and Size

In Tables 1.4 and 1.5, we study heterogeneous effects by technology intensity (low-tech and high-tech sectors) and size (small and large firms).

Table 1.4 shows the results for low-tech and high-tech firms. There is no evidence of heterogeneous effects by technology intensity in the effect of innovation on employment. There is evidence of heterogeneous effects in the effect of innovation on employment composition: product innovations are skill biased for low-tech firms but not for high-tech firms. The evidence comes from a one-sided test against the alternative that product innovations are skill biased, and we reject the null at 10 percent. This result is even more surprising given that the power of the test is lower when we split the sample into low-tech and high-tech firms.

Table 1.5 shows the results for small and large firms. On the effect of innovation on employment and skill composition, there is no evidence of heterogeneous effects by firm size. We cannot reject the null hypothesis that product innovations are not skill biased, but this may be due to the small sample and the lack of power in the test.

1.5 Robustness checks

Innovation and Employment

In this section, we run some robustness checks to evaluate the sensitivity of the results about the effect of innovation on employment to alternative modeling assumptions. First, we include additional instruments to test for exogeneity of the instruments using a Sargan-Hansen overidentification test. The additional instrument is an indicator of positive R&D investment in each year (continuous R&D dummy). If continuous R&D is correlated with time-invariant firm attributes–something we control for–rather than productivity shocks, continuous R&D satisfies the exogeneity assumption. Given the definition of continuous R&D, exogeneity of continuous R&D seems like a sensible assumption. Column (1) in Table 1.6 shows the estimates for the overidentified model. The Sargan-Hansen test does not reject exogeneity of the instruments. These results provide additional evidence of the validity of the chosen instrument.

Second, we estimate the Harrison, Jaumandreu, Mairesse, and Peters (2014) model under the assumptions that both g_2 and process-only innovation are endogenous. Column (2) in Table 1.6 shows the results. The estimate on the coefficient on process-only innovation experiences an important loss in precision. However, the estimate of the coefficient on g_2 is similar to the estimate under exogeneity of the process innovation. Accordingly, the Davidson-MacKinnon test does not reject exogeneity of the process-only innovation variable.

Third, we evaluate whether product-only innovators are different from product and process innovators. To do that, we add an interaction between g_2 and a product and process innovator dummy as an additional covariate. This new variable is endogenous, so we use the interaction between knowledge of support for innovation activity and the product and process innovator dummy as an additional instrument. Column (3) in Table 1.6 shows the results. Although the estimated coefficient on g_2 increases, the interaction is not significant. We conclude that there is no compelling evidence to treat product and process innovators separately from product-only innovators.

Fourth, we control for industry-location shocks including the average employment growth at the industry-regional level as an additional regressor. In the basic specification, we control for industry specific shocks using two-digit industry dummies, and we control for location-specific shocks using location dummies. To control for industry-location shocks, we define five regions: Buenos Aires, Center, Cuyo, South, and North.[6] Then we construct the mean employment growth at the industry-regional level. We expect that this variable is able to capture industry-location-specific shocks. Column (4) in Table 1.6 shows that the variable is not significant and the results are similar to the basic model.

Fifth, given that part of the endogeneity comes from unobserved productivity, we include labor productivity as a proxy for unobserved productivity. The proxy for unobserved productivity is labor productivity in 1998 defined as real sales over workers. Column (5) in Table 1.6 shows that the variable is not significant and the results are similar to the basic model.

Sixth, measurement error in sales of new products can potentially bias our results. To ease concerns about the presence of measurement error in sales of new products, we use a more restricted definition for g_2 . We consider new products not already sold in local or international markets by other firms. Column (6) in Table 1.6 shows that the results are similar to the basic model.[7]

Skill-Biased Innovations

In this section, we run some robustness checks to evaluate the sensitivity of the results on the effect of innovation about employment composition. Table 1.7 shows the results

[6]Buenos Aires includes the city of Buenos Aires; Center includes the provinces of Buenos Aires, Córdoba, and Santa Fe; Cuyo includes Mendoza, San Luis, and San Juan; South includes Chubut, Neuquén, La Pampa, Santa Cruz, Rio Negro, and Tierra del Fuego; and North includes the rest of the provinces.

[7]We also extended the model to consider non-CRS. However, we could not reject the CRS hypothesis. These results are available from the authors on request.

of the different robustness exercises.

As in the previous case, we first include continuous R&D as an additional instrument. In this case, the effect of process innovation is again nonsignificant and therefore robust to different instruments. The effect of product innovation, however, is equal for skilled and unskilled labor. This contradicts the skill bias found using our preferred specification using only knowledge of public support. The fact that different instruments yield different results is shown in the fact that we reject overidentification for skilled labor. We interpret these results in two ways. First, if we have to choose between the two instruments, we are inclined to believe in the exogeneity of knowledge of public support. The arguments behind this statement are written in detail in the Empirical Results section. Second, if the effect of innovation is heterogeneous across firms, even if the two instruments are equally valid, the difference between two IV instruments is related to the fact that the IV estimate measures a local effect on compliers.

Next, we include several regressors used in the literature of skilled technical change for developing countries (see, for example, Meschi, Taymaz, and Vivarelli (2011)). We include log of exports; log of imports of physical capital, equipment, and inputs; and log of technology transfer. Exports capture skill-enhancing effects of exporting activity (learning-by-exporting); imports capture technological transfers embedded in physical capital; and technology transfers capture explicit transfer of technology through licenses and patents. The results in [2] in Table 1.7 show that these variables are not significant and the results do not change.[8]

[8]We also included in the regression the mean employment growth at the industry-regional level to capture industry-location-specific shocks. The coefficient of this variable is not significant, and the results do not change. The same occurs when we include labor productivity in 1998 as an additional regressor. These results are available from the authors on request.

1.6 Quantifying the Effect of Innovation on Employment and Productivity

The effect of each type of innovation on employment growth can be decomposed in a productivity trend, the contribution of noninnovators, the contribution of process-only innovators, and the contribution of product innovators. This decomposition is similar to the employment growth decomposition proposed in Harrison, Jaumandreu, Mairesse, and Peters (2014), but we modify the original decomposition to present separately the contribution of noninnovators. Firm's employment growth can be written as

$$l_i = \left(\sum_j \alpha_j \, industry_{ji} + \sum_k \alpha_k \, location_{ki}\right) + 1(g_{2i} = 0)(1 - d_i)(g_{1i} - \pi_i) + \quad (1.6.1)$$
$$d_i 1(g_{2i} = 0)(\alpha_1 + g_{1i} - \pi_i) +$$
$$1(g_{2i} > 0)(d_i \alpha_1 + g_{1i} - \pi_i + \beta g_{2i}) + v_i,$$

where $industry_{ji}$'s are industry dummy variables; $location_{ki}$'s are location (province) dummy variables; and 1(.) is an indicator function. Thus employment growth can be decomposed into four main components. The first component ($\sum_j \alpha_j \, industry_{ji} + \sum_k \alpha_k \, location_{ki}$) measures the contribution of the (industry-location specific) productivity trend; the second component ($1(g_{2i} = 0)(1 - d_i)(g_{1i} - \pi_i)$) measures the contribution of noninnovators; the third component ($d_i 1(g_{2i} = 0)(\alpha_1 + g_{1i} - \pi_i)$) measures the contribution of process-only innovators; and the fourth component ($1(g_{2i} > 0)(d_i \alpha_1 + g_{1i} - \pi_i + \beta g_{2i})$) measures the contribution of product innovators.

Column (1) in Table 1.8 shows the contribution of the different components to employment growth using the IV estimates. The contribution of the productivity trend is -0.6 percent, which shows a negligible increase in labor productivity in this period. This trend in productivity may be explained, at least in part, by the business cycle. Sales contracted at 9 percent per year, but firms did not translate the full extent of the adjustment to the labor force. This can be an optimal decision for the firms under the presence of labor adjustment costs or if firms have more optimistic expectations for the future (Basu and Fernald (2001)).

The contribution of noninnovators is -4.1 percent. This is the largest contribution and shows that the destruction of jobs during this period was concentrated in noninnovators. The contribution of process-only innovators is -0.6 percent. Two factors affect this contribution. First, there are few firms that introduce only process innovations (15 percent of the sample). Second, process innovations seem to have rather small effects on employment. The contribution of product innovators is 1.4 percent. These results show that product innovators substitute old products with new products at a rapid pace even in a recessive scenario. The result of the innovation-employment model that there are no efficiency gains in the production of new products might also suggest that product innovators are selling a similar product with small changes (incremental innovation).

Columns (2) and (3) in Table 1.8 show the decomposition for low- and high-tech sectors. Employment growth for low-tech firms is -3.5 percent, and employment growth for high-tech firms is -4.9 percent. The decomposition shows that the difference in employment growth can be fully explained by the contribution of product innovators. Given that the relative efficiency of new products is similar for low- and high-tech firms, the differential contribution of product innovators is associated with the larger real sales for product innovators in low-tech sectors.

Columns (4) and (5) in Table 1.8 show the decomposition for small and large firms. Employment growth for small firms is -3.5 percent, and employment growth for large firms is -4.2 percent. The decomposition shows that both innovators and noninnovators in small firms destroy more employment than innovators and noninnovators in large firms. Thus, the larger employment growth for small firms is explained by the lower productivity trend.

1.7 Conclusions

This article presents evidence about the relationship between innovation and employment in the manufacturing sector in Argentina. We aim at understanding whether different types of innovation create or destroy employment and the type of employment that is created or destroyed. To accomplish this, we estimated the model proposed in Harrison, Jaumandreu, Mairesse, and Peters (2014) using an IV approach with data from the Argentinean Innovation Surveys for the period 1998-2001.

The estimation of the effect of the different types of innovation on employment shows that product innovation generates employment, but process innovation has no effect on employment. In the case of product innovations, we find no evidence that new products are produced more efficiently than old products. Thus, the displacement effect of product innovation on employment has no empirical support in our data.

In the case of process innovation, there are two plausible explanations for its lack of effect on employment.

First, a process innovation may not generate important productivity gains; hence, there is no displacement effect on employment. Second, a process innovation may generate productivity gains (displacement effect), which induce a demand enlargement through market competition (creation effect). In the end, the creation effect on employment compensates the displacement effect on employment. Unfortunately, with the available data, we cannot distinguish one explanation from the other. Specification tests support use of an IV approach and the validity of the chosen instruments. These results are robust to using additional instruments, allowing different effects for product and process innovators, adding additional controls, endogeneity of process innovation, and using a different definition of new products.

Our results also show that product innovation is skilled biased. Although the innovation created both skilled and unskilled jobs, the proportion of skilled jobs was higher than the proportion of unskilled jobs. Therefore even if the innovation replaces tasks traditionally carried out by unskilled workers with new jobs demanding qualified workers, the increase in demand also leads to an increase in the demand of unskilled workers.

During the period we analyzed, there was an important contraction in employment due to the recession. We found that most of the contraction in employment was due to noninnovators. Process only innovators contributed –although marginally– to the reduction in employment while product innovators more than compensated for the effect of process innovators. These results were valid both for low- and high-tech industries and for small and large firms. Interestingly, low-tech firms destroyed fewer jobs than high-tech firms because sales decreased less for low-tech product innovators than for high-tech product innovators. Small firms destroyed fewer jobs than large firms because small firms had a lower productivity trend than large firms.

Appendix A: Figures and tables

Table 1.1: Descriptive Statistics

	All firms		Low Technology		High Technology		Small firms		Large firms	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Number of observations	1,415		953		462		417		998	
Distribution of firms										
Noninnovators	36	48	40	49	28	45	56	50	28	45
Process only innovators	15	36	16	37	13	34	12	32	17	37
Product innovators	48	50	43	50	59	49	32	47	55	50
Number of employees	233	556	242	608	215	427	28	12	319	642
Foreign Ownership	20	40	15	36	29	45	6	24	25	44
Located in Buenos Aires	65	48	62	49	70	46	65	48	65	48
Share of skilled labor										
All firms	34	28	30	26	44	30	28	28	37	28
Noninnovators	28	27	25	26	37	30	24	27	31	27
Process only innovators	34	27	29	24	48	29	41	33	32	25
Product innovators	39	28	34	27	46	29	29	27	41	28
Employment growth										
All firms	-4.0	12.3	-3.5	12.2	-4.9	12.4	-3.5	13.4	-4.2	11.7
Noninnovators	-6.0	12.8	-5.6	11.3	-7.2	16.4	-5.8	12.9	-6.2	12.7
Process only innovators	-3.9	12.3	-3.5	13.2	-4.8	9.7	1.5	11.6	-5.4	12.1
Product innovators	-2.5	11.6	-1.6	12.3	-3.8	10.5	-1.2	14.0	-2.8	11.0
Skilled labor growth										
All firms	-1.6	14.5	-1.0	14.4	-3.0	14.7	-2.7	17.6	-1.3	13.3
Noninnovators	-4.1	15.8	-3.7	14.6	-4.9	18.9	-4.8	18.6	-3.6	13.7
Process only innovators	-1.6	14.6	-0.9	15.4	-3.5	12.3	0.4	12.7	-2.2	15.0
Product innovators	-0.1	13.3	1.3	13.4	-2.0	12.9	-0.8	17.4	0.1	12.3
Unskilled labor growth										
All firms	-5.3	14.0	-5.0	13.7	-5.9	14.7	-4.5	16.0	-5.7	13.1
Noninnovators	-7.0	13.9	-6.6	12.7	-8.3	17.0	-6.9	15.2	-7.1	12.7
Process only innovators	-4.3	14.5	-4.4	13.7	-4.1	16.6	2.6	14.0	-6.2	14.1
Product innovators	-4.4	13.8	-3.8	14.4	-5.2	12.9	-2.8	17.2	-4.8	12.9

Descriptive Statistics (Continued)

	All firms		Low Technology		High Technology		Small firms		Large firms	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Sales growth (nominal)[a]										
All firms	-9.0	16.0	-8.5	16.1	-10.0	15.8	-9.7	17.4	-8.7	15.4
Noninnovators	-12.5	17.5	-12.2	17.0	-13.4	19.1	-12.8	17.8	-12.3	17.3
Process only innovators	-8.1	15.3	-7.4	15.4	-10.1	15.0	-3.0	16.5	-9.6	14.7
Product innovators	-6.6	14.5	-5.5	14.7	-8.4	14.0	-6.6	15.7	-6.7	14.2
Labor productivity growth[b]										
All firms	-5.0	15.2	-5.0	15.5	-5.1	14.6	-6.2	18.7	-4.5	13.5
Noninnovators	-6.5	17.4	-6.6	16.7	-6.2	19.5	-7.0	19.6	-6.0	15.4
Process only innovators	-4.3	13.4	-3.8	13.4	-5.4	13.5	-4.5	15.9	-4.2	12.6
Product innovators	-4.2	13.9	-3.9	15.0	-4.5	12.0	-5.4	18.1	-3.9	12.7
Prices growth [c]										
All firms	-2.0	2.3	-2.5	1.5	-1.2	3.2	-2.0	2.0	-2.0	2.4
Noninnovators	-2.3	2.2	-2.6	1.6	-1.5	3.2	-2.2	2.0	-2.4	2.3
Process only innovators	-1.9	2.0	-2.4	1.2	-0.5	2.6	-1.4	1.9	-2.0	2.0
Product innovators	-1.9	2.5	-2.3	1.5	-1.2	3.4	-1.9	1.9	-1.9	2.6

Notes: [i] s.d. = standard deviation. Growth rates are annual rates. [ii] Product innovators are firms that have introduced product innovations between 1998 and 2001. Process-only innovators are firms that have introduced process innovations or organizational change innovations excluding product innovators between 1998 and 2001. Noninnovators are firms not classified as product or process innovators. Skilled workers are employees with a university degree or tertiary education (one- to three-year degree related to technical professions). Unskilled workers are employees with primary or secondary education. Low-technology includes firms in the sectors of food, beverages, and tobacco; textiles, wearing apparel, and leather products; wood, wood products, and furniture; pulp, paper, and paper products; publishing, printing, and reproduction of recorded media; rubber and plastic products; basic metals, fabricated metals, and nonmetallic mineral products. [iii] High-technology includes firms in the sectors of chemicals and chemical products, coke, refined petroleum and nuclear fuel, machinery, equipment, office machinery, computers, communication equipment, electrical machinery, and medical, precision and optical instruments, motor vehicles and transport equipment. Small firms include firms with fewer than 50 employees. Large firms include firms with more than 50 employees.
[iv] Sample: Firms with information in all the relevant variables for the empirical analysis.
[a] Sales growth for each type of firm is the unweighted mean in growth rates across firms conditional on type.
[b] Labor productivity is real sales per worker. [c] Prices computed at the two-digit level of the ISIC and assigned to firms according to their activity.

Table 1.2: Exploratory regressions, OLS estimation

	All firms		Low Technology		High Technology	
Dependent Variable: l (Employment growth)	[1]	[2]	[1]	[2]	[1]	[2]
Constant	-2.219*	-2.214*	-2.122	-2.117	-2.462	-2.459
	(1.049)	(1.047)	(1.241)	(1.239)	(1.891)	(1.887)
Process only innovator (non-product innovator)	0.944	0.942	0.711	0.708	1.992	1.990
	(0.893)	(0.893)	(1.053)	(1.052)	(1.754)	(1.753)
Product only innovator (non-process innovator)	1.804		1.838		1.648	
	(1.125)		(1.519)		(1.819)	
Product and process innovator	2.035**		2.075*		1.780	
	(0.735)		(0.850)		(1.538)	
Product innovator		2.002**		2.040*		1.763
		(0.714)		(0.817)		(1.516)
Real sales growth ($g - \pi$)	0.325***	0.326***	0.311***	0.312***	0.357***	0.357***
	(0.031)	(0.031)	(0.033)	(0.033)	(0.070)	(0.070)
R-squared	0.24	0.24	0.23	0.23	0.27	0.27
Tests (p-values)						
Process&Product=Product only	0.825		0.875		0.918	
Process&Product=Process only	0.196		0.207		0.872	
Number of firms	1,415	1,415	953	953	462	462

Notes: [i] Robust standard errors. [ii] Significance level: *** 1%, ** 5%, and * 10%.
[iii] All regressions include as additional controls a dummy variable taking value one for those firms with more than 20 percent of foreign capital, dummy variables for the province where the firm's headquarters are located, and two-digit industry dummies. [iv] A product innovator is a firm that has introduced at least one product innovation. A process innovator is a firm that has introduced at least one process innovation or organizational change innovation.

Table 1.3: The Effect of innovation on employment and skill composition: OLS and IV Estimation of the Harrison et al. (2014) model

Dependent Variable: $l-(g_1-\pi)$	Labor	Skilled labor	Unskilled labor
A) OLS			
Process only innovator (d)	-0.560	-0.125	0.755
	(1.025)	(1.176)	(1.187)
Sales growth due to new products (g_2)	0.958***	0.963***	0.952***
	(0.013)	(0.015)	(0.015)
R-squared	0.84	0.81	0.80
B) IV			
Process only innovator (d)	1.252	2.998	2.265
	(1.612)	(2.094)	(1.818)
Sales growth due to new products (g_2)	1.151***	1.307***	1.102***
	(0.122)	(0.165)	(0.143)
R-squared	0.80	0.70	0.77
First stage (F-test)	13.94	12.13	
p-value	0.000	0.001	
Endogeneity test for g_2	2.79	6.51	1.17
p-value	0.095	0.011	0.279
Tests (p-values)			
$H_0: \beta = 1$	0.215	0.062	0.475
$H_0: \beta_{skilled} = \beta_{unskilled}$	0.106		
Number of firms	1,415	1,209	1,209

Notes: [i] Robust standard errors. [ii] Significance level: *** 1%, ** 5%, and * 10%. [iii] All regressions include as additional controls a dummy variable taking value one for those firms with more than 20 percent of foreign capital, dummy variables for the province where the firm's headquarters are located, and two-digit industry dummies. [iv] Skilled workers are employees with a university degree or tertiary education (one- to three-year degree related to technical professions). Unskilled workers are employees with primary or secondary education.
[v] Endogenous variables: g_2. Instruments: knowledge of public support for innovation activities.

Table 1.4: The effect of innovation on employment and skill composition: Heterogeneous effects by technology intensity

	Low Technology			High Technology		
Dependent Variable: $l - (g_1 - \pi)$	Labor	Skilled labor	Unskilled labor	Labor	Skilled labor	Unskilled labor
A) OLS						
Process only innovator (d)	-0.754	-0.181	0.261	-0.397	0.550	2.296
	(1.219)	(1.367)	(1.340)	(1.959)	(2.534)	(2.668)
Sales growth due to new products (g_2)	0.967***	0.961***	0.947***	0.935***	0.964***	0.957***
	(0.016)	(0.019)	(0.019)	(0.022)	(0.026)	(0.025)
R-squared	0.83	0.81	0.80	0.84	0.82	0.80
B) IV						
Process only innovator (d)	0.323	1.564	0.523	3.767	7.788	8.171
	(1.665)	(2.234)	(1.820)	(3.743)	(4.677)	(4.719)
Sales growth due to new products (g_2)	1.145***	1.266***	0.978***	1.143***	1.327***	1.246***
	(0.171)	(0.253)	(0.205)	(0.162)	(0.199)	(0.201)
R-squared	0.80	0.72	0.79	0.80	0.70	0.72
First stage (F-test)	6.56	4.77		8.25	8.47	
p-value	0.011	0.029		0.004	0.004	
Endogeneity test for g_2	1.17	2.09	0.02	1.78	4.62	2.73
p-value	0.280	0.148	0.888	0.183	0.032	0.099
Tests (p-values)						
$H_0 : \beta = 1$	0.397	0.293	0.916	0.376	0.100	0.222
$H_0 : \beta_{skilled} = \beta_{unskilled}$	0.179			0.590		
$H_0 : \beta_{low-tech} = \beta_{high-tech}$	0.994					
Number of firms	953	808	808	462	401	401

Notes: [i] Robust standard errors. [ii] Significance level: *** 1%, ** 5%, and * 10%.
[iii] All regressions include as additional controls a dummy variable taking value one for those firms with more than 20 percent of foreign capital, dummy variables for the province where the firm's headquarters are located, and two-digit industry dummies. [iv] Skilled workers are employees with a university degree or tertiary education (one- to three-year degree related to technical professions). Unskilled workers are employees with primary or secondary education.
[v] Low-technology includes firms in the sectors of food, beverages, and tobacco; textiles, wearing apparel, and leather products; wood, wood products, and furniture; pulp, paper, and paper products; publishing, printing, and reproduction of recorded media; rubber and plastic products; basic metals, fabricated metals, and nonmetallic mineral products. High-technology includes firms in the sectors of chemicals and chemical products, coke, refined petroleum and nuclear fuel, machinery, equipment, office machinery, computers, communication equipment, electrical machinery, and medical, precision and optical instruments, motor vehicles and transport equipment.
[vi] Endogenous variables: g_2. Instruments: knowledge of public support for innovation activities.

Table 1.5: The effect of innovation on employment and skill composition: Heterogeneous effects by size

	Small firms			Large firms		
Dependent Variable: $l-(g_1-\pi)$	Labor	Skilled labor	Unskilled labor	Labor	Skilled labor	Unskilled labor
A) OLS						
Process only innovator (d)	-3.300	-4.172	-0.598	-0.706	0.185	0.488
	(2.522)	(3.236)	(3.112)	(1.118)	(1.243)	(1.268)
Sales growth due to new products (g_2)	0.962***	0.985***	0.959***	0.952***	0.950***	0.941***
	(0.032)	(0.043)	(0.039)	(0.013)	(0.015)	(0.016)
R-squared	0.78	0.73	0.75	0.87	0.85	0.83
B) IV						
Process only innovator (d)	-3.045	-4.183	-0.210	2.158	3.714	2.378
	(2.656)	(3.816)	(3.021)	(2.481)	(2.705)	(2.638)
Sales growth due to new products (g_2)	1.150***	1.357***	1.075***	1.170***	1.239***	1.085***
	(0.208)	(0.388)	(0.308)	(0.174)	(0.194)	(0.184)
R-squared	0.73	0.61	0.71	0.82	0.77	0.80
First stage (F-test)	6.43	3.29		6.20	6.43	
p-value	0.012	0.071		0.013	0.011	
Endogeneity test for g_2	0.96	1.40	0.14	1.90	3.08	0.64
p-value	0.329	0.237	0.706	0.169	0.080	0.424
Tests (p-values)						
$H_0 : \beta = 1$	0.471	0.358	0.806	0.329	0.218	0.643
$H_0 : \beta_{skilled} = \beta_{unskilled}$	0.312			0.328		
$H_0 : \beta_{small} = \beta_{large}$	0.994					
Number of firms	414	304	304	997	902	902

Notes: [i] Robust standard errors. [ii] Significance level: *** 1%, ** 5%, and * 10%.
[iii] All regressions include as additional controls a dummy variable taking value one for those firms with more than 20 percent of foreign capital, dummy variables for the province where the firm's headquarters are located, and two-digit industry dummies. [iv] Skilled workers are employees with a university degree or tertiary education (one- to three-year degree related to technical professions). Unskilled workers are employees with primary or secondary education.
[v] Small firms include firms with fewer than 50 employees. Large firms include firms with more than 50 employees.
[vi] Endogenous variables: g_2. Instruments: knowledge of public support for innovation activities.

Table 1.6: Robustness exercises on the effect of innovation on employment

Dependent Variable: $l-(g_1-\pi)$	[1]	[2]	[3]	[4]	[5]	[6]
Process only innovator (d)	0.000	23.066	-3.590	1.218	1.215	1.182
	(1.117)	(20.216)	(3.407)	(1.667)	(1.593)	(1.611)
Sales growth due to new products (g_2)	1.018***	1.099***	1.632**	1.148***	1.137***	1.282***
	(0.045)	(0.091)	(0.557)	(0.130)	(0.122)	(0.237)
$g_2\times$ product and process innovator ($g_2 \times prod\&proc$)			-0.467			
			(0.403)			
Mean employment growth				0.040		
				(0.116)		
Labor productivity in 1998					0.002	
					0.002	
R-squared	0.83	0.76	0.68	0.80	0.80	0.59
First stage for g_2	38.18	40.19	25.37	12.40	13.34	8.81
p-value	0.000	0.000	0.000	0.000	0.000	0.003
First stage for d		8.56				
p-value		0.000				
First stage for $g_2 \times prod\&proc$			149.12			
p-value			0.000			
Overidentification test	1.88					
p-value	0.170					
Endogeneity test for g_2	1.52	2.37	2.51	2.53	2.32	2.45
p-value	0.217	0.123	0.113	0.112	0.128	0.118
Endogeneity test for d		1.47				
p-value		0.225				
Endogeneity test for $g_2 \times prod\&proc$			1.64			
p-value			0.200			
Tests (p-values)						
$H_0: \beta = 1$	0.691	0.277	0.257	0.255	0.264	0.233
$H_0: \beta_{prod\&proc} = \beta_{product\,only}$			0.247			
Number of firms	1,415	1,415	1,415	1,415	1,415	1,415

Notes: [i] Robust standard errors. [ii] Significance level: *** 1%, ** 5%, and * 10%.
[iii] All regressions include as additional controls a dummy variable taking value one for those firms with more than 20 percent of foreign capital, dummy variables for the province where the firm's headquarters are located, and two-digit industry dummies. [iv] [1] Endogenous variables: g_2. Instruments: knowledge of public support for innovation activities and a continuous R&D dummy. [2] Endogenous variables: g_2 and d. Instruments: knowledge of public support for innovation activities and a continuous R&D dummy. [3] Endogenous variables: g_2 and g_2 × product and process innovator. Instruments: knowledge of public support for innovation activities and knowledge of public support for innovation activities × product and process innovator. [4] Endogenous variables: g_2. Instruments: knowledge of public support for innovation activities. Additional control: Mean employment growth at the industry and regional level. [5] Endogenous variables: g_2. Instruments: knowledge of public support for innovation activities. Additional control: Firm's labor productivity in 1998. [6] Endogenous variables: g_2, sales of new products not already sold in the market by other firms. Instruments: knowledge of public support for innovation activities.

Table 1.7: Robustness exercises on the effects of innovation on skill composition

	[1]		[2]	
Dependent Variable: $l-(g_1-\pi)$	Skilled labor	Unskilled labor	Skilled labor	Unskilled labor
Process only innovator (d)	0.774	1.730	3.479	2.335
	(1.285)	(1.291)	(2.585)	(2.077)
Sales growth due to new products (g_2)	1.050***	1.040***	1.394***	1.118***
	(0.052)	(0.053)	(0.242)	(0.193)
Exports in 1998 (in logs)			0.031	0.061
			(0.108)	(0.094)
Imports in 1998 (in logs)			-0.294	-0.102
			(0.201)	(0.158)
Technology transfer in 1998 (in logs)			-0.592	-0.126
			(0.382)	(0.307)
R-squared	0.80	0.79	0.65	0.77
First stage (F-test)	35.87		6.71	
p-value	0.000		0.010	
Overidentification test	4.82	0.26		
p-value	0.028	0.612		
Endogeneity test for g_2	2.51	2.46	5.63	0.81
p-value	0.113	0.117	0.018	0.368
Tests (p-values)				
$H_0 : \beta = 1$	0.335	0.449	0.103	0.542
$H_0 : \beta_{skilled} = \beta_{unskilled}$	0.813		0.126	
Number of firms	1,209	1,209	1,209	1,209

Notes: [i] Robust standard errors. [ii] Significance level: *** 1%, ** 5%, and * 10%.
[iii] All regressions include as additional controls a dummy variable taking value one for those firms with more than 20 percent of foreign capital, dummy variables for the province where the firm's headquarters are located, and two-digit industry dummies. [iv] Skilled workers are employees with a university degree or tertiary education (one- to three-year degree related to technical professions). Unskilled workers are employees with primary or secondary education.
[v] [1] Endogenous variables: g_2. Instruments: knowledge of public support for innovation activities and a continuous R&D dummy. [2] Endogenous variables: g_2. Instruments: knowledge of public support for innovation activities. Additional controls: Exports in 1998 (in logs), Imports in 1998 (in logs), and Technology transfer in 1998 (in logs).

Table 1.8: Contributions of innovation to employment growth (annual rates of growth 1998-2001 in percentage)

	All firms	Low-technology	High-technology	Small firms	Large firms
Firms' employment growth	-4.0	-3.5	-4.9	-3.5	-4.2
Productivity trend	-0.6	-0.5	-0.3	1.9	-2.0
Contribution non-innovators	-4.1	-4.2	-3.9	-5.9	-3.4
Contribution process only innovators	-0.6	-0.6	-0.6	-0.4	-0.7
Contribution product innovators	1.4	1.8	0.0	0.9	2.0

[i] Low-technology includes firms in the sectors of food, beverages, and tobacco; textiles, wearing apparel, and leather products; wood, wood products, and furniture; pulp, paper, and paper products; publishing, printing, and reproduction of recorded media; rubber and plastic products; basic metals, fabricated metals, and nonmetallic mineral products. [ii] High-technology includes firms in the sectors of chemicals and chemical products, coke, refined petroleum and nuclear fuel, machinery, equipment, office machinery, computers, communication equipment, electrical machinery, and medical, precision and optical instruments, motor vehicles and transport equipment. [iii] Small firms include firms with fewer than 50 employees. [iv] Large firms include firms with more than 50 employees.

Chapter 2

Competition between single-market and multimarket banks: Evidence from the U.S. Banking Industry

2.1 Introduction

After a long history of branching restrictions, the market structure in the U.S. banking industry presents direct competition between different organizational forms in the same geographic market: *Single-market* and *Multimarket banks*. For example, in the city of Boston there is competition between Bank of America and Meetinghouse Bank. Bank of America owns 4,896 offices located in 35 different states and has 1,600 billions of $ in assets; Meetinghouse Bank owns 2 offices located in Boston and has 126 millions of $ in assets.

How can such different firms coexist in the same market? A plausible explanation is that single-market banks have a niche market where they enjoy a comparative advantage over multimarket banks. Who are the clients in this niche market? Previous work suggest that borrowers with projects whose profitability depends crucially on the borrower's reputation.[1] Importantly this reputation can only be learned by

[1]See, for example, Petersen and Rajan (1994) and Stein (2002).

the lender through personal interaction and cannot be truthfully transmitted within a hierarchical organization. These loans are usually called "soft information" loans. On the other hand, multimarket banks have a comparative advantage if the borrower reputation can be learned using "hard information" like financial statements, credit history, credit scoring, etc.

In this paper I look for evidence consistent with the relationship lending explanation for the comparative advantage of single-market banks. To do that, I estimate differences in profitability between single-market and multimarket banks in small geographic markets. I focus on small geographic markets because in these markets firms are smaller and less sophisticated thus more likely to be good candidates for a relationship loan. Moreover, in smaller markets is more likely that the branch manager has information about the community that is difficult to transmit within a hierarchical organization. The results of the paper show that single-market banks have a comparative advantage in small markets controlling for market structure, market size variables and time-invariant market unobservables.

A better understanding of the differences between single-market and multimarket is a relevant economic question for the U.S. banking industry due to recent regulatory changes in the industry. Historically the banking industry in the United States was subject to strict regulation from the state and federal governments. There were particularly stringent restrictions to opening branches in different geographic markets in both the same state and across different states. In the period 1970-1990, some states relaxed some of these regulations, and, finally, in 1994 the Riegle-Neal Interstate Banking and Branching Efficiency Act almost eliminated any branching restrictions between states. The welfare effects of the Riegle-Neal may be important. On the one hand, some economists stress that the act lowers barriers to entry that can increase competition at the local market level, and it can allow banks to exploit economies of scale and risk diversification strategies. On the other hand, a change in the distribution of banks may imply a change in the access to or the cost of credit for agents traditionally served for single-market banks. If there is compelling evidence that single-market banks have a comparative advantage with some agents, we may learn that a pro-competition policy may have some unexpected distributional consequences that policy makers might want to consider when designing new policies.

After the Riegle-Neal Act the U.S. banking industry shows a slow but continuous process of structural change in the market structure of the industry. In the sample of small markets, the average market structure changed from 3 single-market banks and 2 multimarket banks in 1994 to 2 single-market banks and 4 multimarket banks in 2007. What is a reasonable explanation behind this change in the market structure? A potential suspect is the expansion cost of single-market banks. Although this cost is unobservable for the researcher, Pakes, Ostrovsky, and Berry (2007)(POB) propose to infer the entry costs of firms using data on entry and exit decisions, the evolution of the market structure and the firm's profit function. I apply an extension of POB approach to estimate the entry costs for single-market and multimarket firms. The results show that single-market banks face higher entry costs than multimarket banks. These results can explain the findings in Aguirregabiria, Clark, and Wang (2015) that the Riegle-Neal Act opened the door for potential gains due to risk diversification for banks, but these gains remained largely unexploited for small banks. This paper offers a plausible explanation: the expansion costs for small banks are too large so it is not profitable to follow a risk diversification strategy.

This paper relates with several empirical papers that study agency problems in soft information loans in the banking industry. In a seminal paper Petersen and Rajan (1994) find that a solid relationship with a bank (repeated interaction or number of products) has a positive effect on the availability of funds but it has no effect on the interest rate paid by the creditor. In this paper I seek to understand if the bank type affects whether or not a solid relationship with a bank alleviate financial constraints. Brickley, Linck, and Smith Jr. (2003) using data from Texas in 1998, estimate the probability that a large bank owns a branch conditional on the type of local market: large urban market, small urban market, or rural market. They find that the probability that a large bank owns a branch decreases significantly in a rural market. Their findings are consistent with the results is this paper, but the identification strategy is different. This paper exploits the panel data structure of the data while Brickley, Linck, and Smith Jr. (2003) exploit the cross-section variation in the data. Finally, Berger, Miller, Petersen, Rajan, and Stein (2005) using bank-firm matched data find evidence consistent with small banks specialization in soft information loans. In particular, they find that large banks lend to less opaque borrowers, at a greater distance between firm and bank headquarters, with less personal interaction with the borrower, and they form shorter and less exclusive relationships. A difference with Berger, Miller, Petersen, Rajan, and Stein (2005) is that in this paper the relevant difference between banks is not size but geographic scope that is more relevant for the soft information story.

A related literature studies the effect of the market structure on small firms. Sapienza (2002) measures the effects of bank mergers on loan interest rates and credit supply for Italy. Sapienza finds evidence that, after a large bank acquires a small bank, small borrowers of the target bank are less likely to borrow from the consolidated bank. This evidence is consistent with small banks having a comparative advantage in small business loans. In a recent paper, Canales and Nanda (2012) evaluate whether the presence of decentralized banks is beneficial for small firms in Mexico. Interestingly, they find that the result is ambiguous and it depends on the level of competition: only in a more competitive environment a greater presence of decentralized banks implies lower loan interest rates for small firms.

Finally, there are some papers that analyze the effects of the Riegle-Neal Act. In a sample of large urban markets (MSA) Dick (2007) finds that, after the Riegle-Neal Act, banking markets became less concentrated and banks became more efficient. Rice and Strahan (2010) find that the Riegle-Neal Act increased the probability of borrowing from banks for small firms and it decreased the loan interest rates paid by small firms.

These results do not contradict the findings in this paper because Rice and Strahan did not distinguish the effects in markets of different size. Finally, Aguirregabiria, Clark, and Wang (2015) seek to measure efficiency effects due to the Riegle-Neal Act. They find that the act expanded substantially the potential gains in risk diversification due to the possibility of expanding to other states. However, few small banks took advantage of these risk diversification possibilities.

The rest of the paper is structured as follows. The next section reviews the history of the U.S. banking stressing recent changes in market structure. The data used to estimate the model are described in section 2.3. Section 2.4 presents the entry/exit model used in the structural estimation. Section 2.5 describes the two-stage procedure used in the estimation. Section 2.6 reports and discusses the main results of the estimation. Final remarks and conclusions are presented in section 2.7.

2.2 Industry background

The U.S. banking industry has a remarkable feature compared with the banking industry in other countries: 75 percent of U.S. banking institutions have presence in only one local market, and these banks make around 40 percent of all banking loans to small businesses and farms.

This feature of the U.S. banking industry is not random but the consequence of restrictions on the geographic expansion of banks that have a long history in the United States. Because the U.S. Constitution prevented states from issuing fiat money and from taxing interstate trade, the states used their power to grant bank charters to generate a substantial part of state revenues. A state received no charter fees from banks incorporated in other states, so states prohibited out-of-state banks from operating in their territories. These were called *interstate branching restrictions*. States would grant a charter for a specific location or limit bank branches to that city or county. By adopting branching restrictions, states created a series of local monopolies from which they could extract part of the rents. These were called *intrastate branching restrictions*.

In the period 1970-1994, states started to deregulate these geographic restrictions. There were different stages of deregulation. First, states relaxed intrastate branching restrictions. Second, states signed bilateral agreements allowing banks chartered in one state to open branches in the other state and vice versa. Though the deregulation phenomenon was quite extended, different states showed different timing and intensity in their deregulation. See Kroszner and Strahan (1999) for a political economy explanation of the different speed in the deregulation across states.

In 1994, the Congress passed the Riegle-Neal Interstate Banking and Branching Efficiency Act that effectively permitted banks and holding companies to enter any state. Until 1997 the states had the option to opt-in or opt-out some of the provisions in the act, but most of the states decided to opt-in in 1994.

In part as a consequence of such regulatory changes, the industry became more concentrated. Figure 2.1 shows that the number of banks halved between 1979 and 2007 and decreased 25 percent between 1994 and 2007.

Aggregate industry trends may differ with the trends at the local market level. Table 2.1 shows the evolution of concentration and market structure at the aggregate level and at the local market level. Local markets are classified as large urban markets, small urban markets and rural markets. A large urban market is defined as a Metropolitan Statistical Area, an urban area with a core of more than 50,000 inhabitants and contiguous counties. A small urban market is defined as a Micropolitan Statistical Area, an urban area with a core of more than 10,000 inhabitants but less than 50,000 inhabitants and contiguous counties. A rural market is a county not classified as a Metropolitan or Micropolitan Statistical Area.[2] Interestingly, concentration indexes at the local market level remained steady even when aggregate concentration

[2]Metropolitan Statistical Areas and Micropolitan Statistical Areas are geographic areas defined by the U.S. Office of Management and Budget and used by the Census Bureau and other federal government agencies for statistical purposes.

increased.

The market structure changed after the deregulation. Figure 2.1 shows that the number of multimarket banks increased a 27 percent in the period 1994-2007 and the number of single-market banks decreased more than 40 percent in the same period. Looking at the numbers in terms of branches is even more striking as shown in figure 2.2. Table 2.1 reports the evolution of the market structure for local markets. The mean market structure in a large urban market changed from 19 single-market banks (SM banks) and 9 multimarket banks (MM banks) in 1994 to 14 SM banks and 15 MM banks in 2007; the mean market structure in a small urban and rural market changed from 3 SM banks and 2 MM banks in 1994 to 2 SM banks and 4 MM banks in 2007.

Tables 2.2 and 2.3 show the sources of variation in the number of single-market and multimarket branches. The dynamics in the number of branches is decomposed in the contribution of branch openings, branch closings, mergers at the bank level, and switching in the bank type (a SM bank becoming a MM bank). Branch opening and closing explain an important part of the variation in the number of branches. Bank mergers are important in the sample period. However most of the mergers during this period were out-of-market mergers and this type of merger do not affect the market structure at the local market level. In the rest of the paper, I consider that merger decisions are exogenous from the point of view of the local market. This assumption is a reasonable approximation to the problem at hand given the difficulty to model merger decisions taken at the bank level using a model at the level of a local market.

2.3 Data

2.3.1 Data sources

The data set consists of yearly data for nearly all commercial banks and thrifts in United States and socioeconomic data at the market level for the period 1994-2007. The bank data were obtained from four sources: the Summary of Deposits (SOD) and the Institution Directory from the Federal Deposit Insurance Corporation (FDIC), the Reports on Condition and Income (Call Reports) from the Federal Reserve Bank of Chicago (FED), and the Thrift Financial Reports from the Office of Thrift Supervision (OTS).

The Summary of Deposits is a yearly survey conducted by the FDIC, a government corporation that insures deposits, examines and supervises financial institutions, and manages receiverships. The FDIC requests all FDIC-insured banks to submit a form with the amount of deposits in each branch at June, 30th each year. The Summary of Deposits includes deposits, branch's location, and branch's ownership information. The Summary of Deposits data are used to construct the number of SM incumbents, the number of MM incumbents, and entry and exit variables.

To identify mergers I complemented the data in the SOD with the FDIC's Institution Directory. The Institution Directory lists structural changes in bank institutions: the date a bank began operations, the date a bank finished operations, and the reasons for finishing operations. In case a bank merges with another bank, the Institution Directory includes the bank code of the acquirer.

Call Reports are available for all commercial banks regulated by the FED, FDIC, and the Comptroller of the Currency. The Call Reports contain balance sheet information collected on a quarterly basis. I use the Call Reports data to construct deposit interest rates, loan interest rates, default rates, wage expenditures, and other costs. I use the second quarter information to make it comparable with the SOD information. Thrift Financial Reports provide the same information for thrifts.

Demographic data at the local market level come from the U.S. Census Bureau, the Bureau of Economic Analysis (BEA), and the Bureau of Labor Statistics (BLS). Population, income per capita, number of employees, and average wage come from the BEA's Local Area Personal Income, number of establishments comes from the U.S. Census Bureau' County Business Patterns, and consumer price indexes come from the BLS's CPI.

2.3.2 Variable definitions

Local market definition

A local market definition should balance two contradictory objectives. On the one hand, it should be large enough so that households and firms in the local market do not use banking services from banks outside the local market. On the other hand, it should be small enough so that there are no distinct or overlapping submarkets within the local market.

To satisfy those requirements I selected: (i) Small urban markets and rural markets, (ii) with less than 100,000 inhabitants, and (iii) with less than 8 single-market and 8 multimarket incumbent banks in the period 1994-2007.

As mentioned above, a small urban market is a Micropolitan Statistical Areas, a set of counties with at least one urban cluster of at least 10,000 but less than 50,000 population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties. A rural market is a county not classified as Micropolitan or Metropolitan Statistical Areas.[3]

I dropped large urban areas and markets with more than 100,000 inhabitants to avoid the presence of distinct and overlapping submarkets. I dropped markets with more than 8 single-market and 8 multimarket incumbent banks for computational reasons. The computational time of the estimation increases exponentially in the number of incumbents so there are practical limits to number of incumbents that the estimation can handle. I relaxed the selection rule slightly to check that the results are robust to small changes in the maximum number of incumbents.

The final sample consists of 1,691 local markets. The sample includes small urban and rural markets and it covers 12 percent of the total population in the United States. Figure 2.3 shows a map of the United States with large urban markets in grey, small urban markets in black, and rural counties in white. The map shows that the sample

[3]A Metropolitan Statistical Area is a set of counties with at least one urbanized area of 50,000 or more population.

is more representative of the Midwest and South regions rather than the West and Northeast regions. For example, the sample represents 30 percent of the population in the East South Central Division (Alabama, Kentucky, Mississippi, Tennessee) and West North Central Division (Iowa, Nebraska, Kansas, North Dakota, Minnesota, South Dakota, Missouri), but around 5 percent of the Middle Atlantic Division (New Jersey, New York, Pennsylvania) and Pacific Division (Alaska, California, Hawaii, Oregon, Washington).

The selection of the sample reflects the fact that it is difficult to properly identify a relevant market in large urban areas for retail activities. To deal with this problem, empirical researchers select markets where a relevant market is clearly defined. For example, Bresnahan and Reiss (1991)'s seminal paper selected 202 isolated markets at least 20 miles from the nearest town of 1,000 people or more, Mazzeo (2002) selected 492 markets excluding MSA and counties with more than 15 firms, and Seim (2006) selected 151 markets which consisted in cities with population between 40,000 and 150,000.

Table 2.4 shows descriptive statistics for the sample of local markets. The median market has a population of 15,000 inhabitants, an income per capita of $21,000, and 343 business establishments. This shows that local markets in the sample are relatively small. However there is enough variability in the sample in population, employees, establishments, and population density that can be exploited in the estimation. As usual the distribution of population, business employees, and business establishments per market is asymmetric with a right tail.

Single-market and multimarket banks

The empirical definition of a single-market is a bank that holds more than 80 percent of its total deposits in a single local market. Otherwise the bank is classified as a multimarket bank.

Table 2.5 shows differences in observable characteristics between single-market and multimarket banks. I compute statistics at the bank-market level and at the bank level. For the variables at the bank-market level I calculate a simple mean, for the variables at the bank level I calculate a mean weighted by the number of branches.

Single-market and multimarket banks differ in geographic scope, size, ownership structure, and lending practices. The average SM bank was active in 2 local markets and 1 state in 2007 while the average MM bank was active in 73 local markets and 5 states. Also the average SM bank owned 6 branches in 2007 while the average MM bank owned 570 branches in 2007. These differences in geographic scope and size have increased since 1994 as a result of deregulation and consolidation in the banking industry. Regarding ownership structure, 25 percent of SM banks were owned by a multibank holding company in 2007 while 58 percent of MM banks were owned by a multibank holding company in 2007.

More relevant are differences in lending practices. In 2007, the average SM bank lent 86 percent of its business loans to small businesses while the average MM bank lent 59 percent of its business loans to small businesses. Also, the average SM bank lent 83 percent of its farm loans to small farms while the average MM bank lent 65 percent of its farm loans to small farms.[4] Such ratios show some evidence of SM banks having a comparative advantage in the provision of loans to small businesses and small farms.

There are differences between SM and MM banks in average interest rates paid on deposits and average interest rates charged on loans. The average SM bank paid a higher interest rate on deposits and charged a higher interest rate on loans than the average MM bank. The higher interest rate paid on deposits by a SM bank may be evidence that a SM bank has to offer a higher return to attract depositors. The higher interest rate paid on loans by a SM bank may be evidence of SM banks exploiting an

[4]Loans to businesses includes loans secured by nonfarm nonresidential property and commercial and industrial loans, and loans to farms includes loans secured by farmland and loans to finance agricultural production and other loans to farmers. Loans to small businesses are loans to businesses with amounts smaller than $ 1,000,000, and loans to small farms are loans to farms with amounts smaller than $ 500,000.

informational advantage when lending to more opaque lenders or may be evidence of SM lending to riskier lenders. It can also be the case that MM banks charge lower loan interest rates because they can exploit some economies of scale.

There are also differences in the ratio of equity to assets and non-performing loans to loans. The average SM bank had a higher equity-assets ratio and a higher non-performing loans-loans ratio. These might highlight that a MM bank can decrease the risk in their portfolio through diversification in different geographic markets.

Finally, there does not exist observable differences in the number of branches per market of an incumbent bank of the SM or MM type. Both type of incumbents owned approximately 2 branches. This statistic is useful to rule out a possible interpretation for the differences in entry costs between the 2 types: they are not driven by differences in the number of branches.

To sum up, the average SM bank has less geographic scope, is smaller, and has a simpler ownership structure. Hence a SM bank cannot enjoy economies of scale, economies of scope, or geographic risk diversification but it may be able to exploit informational advantages in relationship lending.

There is a caveat to the descriptive statistics reported in table 2.5. Given that most of the statistics are computed at the bank level, a mean at the bank level for a MM bank includes the lending practices in other local markets where those banks are incumbents while a mean at the bank level for a SM bank includes the lending practices in the local markets in the sample. Thus part of the observed differences for SM and MM banks might be explained by different lending practices across local markets that affect the means reported by MM banks. The descriptive statistics should be interpreted with such limitation in mind.

Table 2.6 shows the distribution of market structures in the final sample. Each cell reports the percentage of times that such market structure is observed in the data. The most observed market structure is a local market with 2 SM banks and 1 MM bank. The local markets in the final sample are concentrated markets: 75 percent of the observed markets had less than 4 incumbents of each type. It does not exist a clear pattern between banks types and market concentration that can drive the results such as SM banks being incumbents in more concentrated markets or vice versa. For example, a SM monopolist is observed 2.7 percent of the time and a MM monopolist is observed 3.8 percent of the time.

Potential entrants

I use 2 different definitions to capture potential entrants. For the first definition I assume one potential entrant of each type, and I drop a few markets that experience multiple entries. For the second definition I compute the maximum number of SM incumbents in each market across time, and I define the number of SM potential entrants in period t as the difference between that maximum and the number of SM incumbents in a period t. I apply the same procedure the number of MM potential entrants. The rationale behind this definition is that in each geographic market we observe all potential entrants being active at some point in time. Dunne, Klimek, Roberts, and Xu (2013) use a similar definition but for a model with homogenous firms. To check the sensitivity of the estimates to the definition of potential entrants, I estimate the model using the 2 different pools of entrants.

Exit and entry definition

I matched branches over time using 7 variables: branch's FDIC code, bank's FDIC code, bank holding company's FED code, address, city reported, ZIP code, state, and county.[5] I use exact merge and fuzzy merge with different subsets of these variables. The exact merge matched 95 percent of the branch-year observations and the fuzzy merge matched an additional 1.5 percent-2 percent of the unmatched branch-year observations.

[5] I use multiple variables for the matching because the branch code in the SOD has 10 percent of missing values and it does not change in a consistent manner.

The longitudinal data set at the branch level allows me to identify opening and closing of branches: a new branch in the data set is an opening, and a branch that drops from the data set is an closing. I use the branch code and branch's address to differentiate true opening and closing of branches from changes of ownership, i.e. a bank acquired by another bank or a bank selling some branches to another bank.

Finally, I identified that a bank enters in a local market when a bank is not an incumbent in $t-1$ but is an incumbent in t and all its branches are opened in t. I identified that a bank exits from a local market when a bank is an incumbent in $t-1$ but is not an incumbent in t and all its branches are closed in $t-1$. I do not consider an entry or an exit those situations where there is a change in ownership: when a bank acquired another bank or when a bank enters in a local market by buying branches from another bank.

I clean some cases from the data to avoid measurement error issues. Specifically, I do not consider as an entry or exit: a bank that enters and exit more than once in the same local market, and an entry/exit using the bank code but not the branch code. I also drop banks without deposits, and branches in Hawaii, Alaska, Puerto Rico or American Islands.

In table 2.7 I computed some entry and exit statistics for the different number of incumbents in the local market. I observe that both entries and exits increased in absolute terms with the number of incumbents but the number of entries increases less than proportionally with the number of incumbent and the number of exits increases proportionally with the number of incumbents. Dunne, Klimek, Roberts, and Xu (2013) find a similar evidence for a sample of doctors and chiropractors. Without a structural model it is difficult to interpret whether this observation is driven by differences in the number of potential entrants, entry costs, or profitability. I also observe that the entry proportion is larger than the exit rate. This is evidence of the geographic expansion of banks after the deregulation and a quite profitable period for the banking industry. Also, the saving and loans crisis of the 80's and early 90's caused the more inefficient banks to exit the industry, and this contributed to the lower exit rate afterwards.

More relevant for this paper are differences between SM and MM banks. I observe that the entry proportion is larger for SM than for MM banks, and the exit rate is slightly smaller for SM than MM banks. This evidence shows that the decrease in the number of SM branches was due either to SM banks downsizing, SM expanding to new local markets and becoming MM banks, or a MM banks acquiring a SM banks. But it is also true that SM banks were active players in the industry during this period. It is not clear whether SM banks continue to open branches because they face low entry costs than compensate demand or production cost disadvantages with respect to MM banks, or SM banks face higher entry costs that were compensated by informational advantages with respect to MM banks.

The structural model introduced in the next section tries to identified such differences in entry costs, demand, and production costs between SM and MM banks.

2.4 Model

The model is an oligopolistic model of entry/exit with imperfect information similar to POB. There are incumbent and potential entrant banks competing in geographic markets. Each period a bank observes a private information shock. An incumbent bank decides whether to continue or exit, and a potential entrant bank decides whether to enter or not. Banks choose optimal actions based on their beliefs about their competitors' behavior. In equilibrium, those beliefs are correct. The model departs from POB framework by allowing heterogeneity between banks based on their geographic scope: single-market and multimarket banks.

Both single-market and multimarket banks take entry/exit decisions separately in each market based on the market profitability. Such an assumption is reasonable for identification of heterogeneity between single-market and multimarket banks in several dimensions: preferences, technology, sell-off value, and entry costs.

There are $m \in \{1, 2, \ldots, M\}$ geographic markets and infinite periods $t \in \{1, 2, \ldots, \infty\}$. Banks are indexed by $i \in \mathcal{N}$, where the set of banks $\mathcal{N}$ can be partitioned in the set of single-market banks and the set of multimarket banks.

Bank's profitability depends on common knowledge and private information state variables. Common knowledge state variables are number of incumbents of type $\tau \in \{1, 2\}$, n^{τ}_{mt}, and market state variables, z_{mt}. The private information state variable for the incumbent is its sell-off value ϕ^{τ}_{imt}, and for the potential entrant is its entry cost κ^{τ}_{imt}.

The timing of the game is as follows. First, a bank observes the number of incumbents, the market state, and its private information realization. Second, banks simultaneously choose actions. Third, entrants pay the entry cost and incumbents earn current profits. Finally, at the end of the period, banks that exit earn the sell-off value, entrants become incumbents, and the market evolves to a new state. The timing is summarized in figure 2.4.

The incumbent's current payoff is

$$\Pi^{\tau}_{imt} = \begin{cases} \pi^{\tau}(n^1_{mt}, n^2_{mt}, z_{mt}; \theta^{\tau}_P) & \text{if } a^{\tau}_{imt} = 1, \\ \pi^{\tau}(n^1_{mt}, n^2_{mt}, z_{mt}; \theta^{\tau}_P) + \beta\, \phi^{\tau}_{imt} & \text{if } a^{\tau}_{imt} = 0, \end{cases}$$

where $\pi^{\tau}(.; \theta^{\tau}_P)$ is the bank profit function parameterized by θ^{τ}_P, β is the intertemporal discount factor, and a^{τ}_{imt} is the action continue/exit for the type τ incumbent or enter/not enter for the type τ potential entrant. The current payoff is consistent with the timing of the game: incumbents earn current profits, and, at the end of the period, exiters receive the sell-off value.

The potential entrant's current payoff is

$$\Pi^{\tau}_{imt} = \begin{cases} -\kappa^{\tau}_{imt} & \text{if } a^{\tau}_{imt} = 1, \\ 0 & \text{if } a^{\tau}_{imt} = 0. \end{cases}$$

I assume that potential entrants are short-lived to avoid timing of entry issues. Entrants pay the entry cost this period but become incumbents next period, not entrants receive a zero payoff.

Private information shocks are IID over banks, markets, and time with CDFs,

$$\phi^{\tau}_{imt} \sim F(.; \theta^{\tau}_{X}),$$
$$\kappa^{\tau}_{imt} \sim G(.; \theta^{\tau}_{E}).$$

θ^{τ}_{X} and θ^{τ}_{E} are the parameters of the sell-off value CDF and entry cost CDF. Although the entry costs and sell-off values are private information, their CDF is common knowledge.

To simplify notation the set of parameters is denoted by $\theta = (\theta^{\tau}_{P}, \theta^{\tau}_{X}, \theta^{\tau}_{E})^{2}_{\tau=1}$, the set of common knowledge variables is denoted by $s = (n^1, n^2, z)$, and the private information shock is denoted generically by ν_i.

Each bank maximizes the discounted expected value of future payoffs,

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \pi_i(a_{imt}, s_{mt}, \nu_{imt}) | s_{m0}, \nu_{im0}\right],$$

where the expectation is taken over beliefs about its competitors' actions and the evolution of the market state variables.

I focus on Markov Perfect Equilibria (MPE) of the game. A MPE is a Subgame Perfect equilibrium in payoff relevant strategies or Markov strategies. Formally, a Markov strategy is a mapping $\sigma_i(s, \nu_i) \mapsto \{0, 1\}$ which assigns an action to each possible realization of the state variables. A Markov strategy profile $\sigma = (\sigma_1, \ldots, \sigma_N)$

assigns an action to each player. Then a Markov strategy is a MPE if and only if for all s, ν_i, i, and alternative strategies σ_i',

$$V_i(s, \nu_i | \sigma_i, \sigma_{-i}) \geq V_i(s, \nu_i | \sigma_i', \sigma_{-i}),$$

where $V_i()$ is the value function of bank i associated to the corresponding strategy profile. I focus on symmetric MPE. Therefore a Markov strategy for an incumbent and a potential entrant of each type completely characterize the equilibrium.

The integrated value function is the value function with the private information shock integrated out. Under an IID assumption for the private information shock there is no loss of generality in working with the integrated value function.

The integrated value function for the incumbent can be written as the solution of a functional equation:

$$V_{in}^{\tau}(s;\theta) = \pi^{\tau}(s;\theta_P^{\tau}) + \beta \int [\max\{\phi_i^{\tau}, VC^{\tau}(s;\theta)\}] \, dF(\phi_i^{\tau};\theta_X^{\tau}), \tag{2.4.1}$$

$$VC^{\tau}(s;\theta) = \sum_{s'} V_{in}^{\tau}(s';\theta) P_{in}^{\tau}(s'|s, a_i^{\tau} = 1). \tag{2.4.2}$$

$VC^{\tau}()$ is the continuation value, i.e. the expected value function in $t+1$ conditional on the state in t and the bank continuing. $P_{in}^{\tau}()$ is the transition probability for the state variables conditional on continuing and it embodies the beliefs about its competitors' strategies. Equations (2.4.1) and (2.4.2) can be solved for the integrated value function or the continuation value, but it is straightforward to write optimal policies in terms of the continuation value.

The optimal choice for a type τ incumbent is to exit if $\phi_i^{\tau} > VC^{\tau}()$, otherwise to continue. The exit probability of a type τ bank i is the expected behavior of the bank before the realization of the private information shock,

$$Pr(\tau\, exits|s;\theta) = Pr(\phi_i^\tau > VC^\tau(s;\theta)),$$
$$= 1 - F(VC^\tau(s;\theta);\theta_X^\tau). \tag{2.4.3}$$

Given arbitrary beliefs on rivals strategies equation (2.4.3) is the expected best response of type τ bank, but it is the expected behavior of bank i if other firms are playing equilibrium strategies.

The integrated value function for the potential entrant can be obtained as the solution of the following equation

$$V_{en}^\tau(s;\theta) = \max\{0, -\kappa_i^\tau + \beta VE^\tau(s;\theta)\}, \tag{2.4.4}$$
$$VE^\tau(s;\theta) = \sum_{s'} V_{in}^\tau(s';\theta) P_{en}^\tau(s'|s, a_i^\tau = 1). \tag{2.4.5}$$

$VE^\tau()$ is the entry value or expected value of a potential bank next period conditional on entry, and $P_{en}^\tau()$ is the transition probability of the state variables conditional on entry. The entry value is a function of the continuation value through the incumbent value function.

The optimal choice for a type τ potential entrant is to enter if $\kappa_i^\tau < \beta VE^\tau()$, otherwise not entering is optimal. The entry probability of a type τ bank is

$$Pr(\tau\, enters|s;\theta) = Pr(\kappa_i^\tau \leq \beta VE^\tau(s;\theta)),$$
$$= G(\beta VE^\tau(s;\theta);\theta_E^\tau). \tag{2.4.6}$$

The econometric implementation rests on the theoretical exit probability in equation (2.4.3) and the theoretical entry probability in equation (2.4.6). The estimation strategy is based on finding parameter values that minimize the distance between theoretical and observed exit and entry probabilities. Under similar observed entry probabilities for SM and MM banks can lay heterogeneity in different economic primitives: profitability, sell-off values, or entry costs. I impose the economic structure, and use the data to identify the parameters that affects the entry/exit behavior of SM and MM banks.

2.5 Empirical implementation

I estimate the parameters of the model using a 2-stage procedure. In the first stage I obtain an estimation of VC and VE based on profit, exit probability, and transition probability estimates. In the second stage I use the estimated continuation value $\widehat{VC}$ and entry value $\widehat{VE}$ to compute theoretical exit and entry probabilities that depend on θ. Then the parameter estimates are those values that minimize a distance between theoretical and observed probabilities.

I assume that sell-off values follow an exponential distribution, and entry costs follow a logistic distribution. The exponential assumptions allows me to obtain an explicit expression for $\widehat{VC}$, but it can be replaced by another parametric distribution at the cost of complicating the computation of $\widehat{VC}$. The exponential probability has also the nice property of restricting sell-off values to be positive.

In theory the CDF for entry costs is non-parametrically identified, in practice data constraints require to assume a parametric distribution. I choose the logistic distribution because is similar to the normal distribution and its CDF has an analytical expression which decreases computational time in the estimation. Notice that the independent of irrelevant alternatives critique of the logistic does not apply here because there are only 2 choices: continue/exit for incumbents or entry/no entry for potential entrants.

Assumption 1. *Distribution of entry costs and sell-off values.*

1. *The sell-off values follow an exponential distribution with mean and variance θ_X^τ,*

$$F(\phi;\theta_X^\tau) = 1 - \exp\left(\frac{\phi}{\theta_X^\tau}\right) \text{ with } \phi \in (0,\infty). \tag{2.5.1}$$

2. *The entry costs follow a logistic distribution with mean θ_E^τ and variance $\pi^2/3$,*

$$G(\kappa;\theta_E^\tau) = \frac{\exp(\kappa-\theta_E^\tau)}{1+\exp(\kappa-\theta_E^\tau)} \text{ with } \kappa \in \mathbb{R}. \tag{2.5.2}$$

I fixed the variance of the entry costs to be $\pi^2/3$ [6] and estimate the mean of the distribution θ_E^τ. θ_E^τ might depend on market size so the mean entry cost changes with market size. Moreover, the effect of market size on entry costs might be different for SM and MM banks. In the estimation I explore these different alternatives.

I assume that z follows an exogenous Markov process. This assumption helps to alleviate the curse of dimensionality when estimating transition matrices non-parametrically.

Assumption 2. *Exogenous Markov process for z. z follows an exogenous first order Markov process, $Pr(z'|n^1,n^2,z,a) = Pr(z'|z)$.*

For an exponential distributed random variable, $\mathbb{E}[\phi_i^\tau|\phi_i^\tau > VC^\tau] = \theta_X^\tau + VC^\tau$ holds.[7] Then the VC can be written as,

$$VC^\tau(s;\theta) = \sum_{s'} \{\pi^\tau(s';\theta_P^\tau) + \beta\, P_{exit}^\tau(s')\,\theta_X^\tau + \beta\, VC^\tau(s';\theta)\}\, P_{in}^\tau(s'|s,a_i^\tau=1),$$

where $P_{exit}^\tau(s)$ is the reduced form exit probability in state s. Such a functional equation can be solved for the continuation value in matrix form as

$$VC^\tau(\theta) = [I - \beta M_{in}^\tau]^{-1} M_{in}^\tau[\pi^\tau(\theta_P^\tau) + \beta P_{exit}^\tau\,\theta_X^\tau], \tag{2.5.3}$$

[6] $\pi^2/3$ is the variance of a standard logistic random variable

[7] Due to this property of the exponential distribution the continuation value has an explicit expression, otherwise the continuation value is a fixed point in a functional equation. Then the exponential assumption avoids solving a functional equation for each parameter value tried in the estimation.

where VC^{τ},π^{τ},P^{τ}_{exit} are vectors that stack the continuation value, profit function, and exit probability in each state, M^{τ}_{in} is a matrix with the transition probability between states conditional on the incumbent continuing, and I is the identity matrix. Analogously the entry value can be written in matrix form as

$$VE^{\tau}(\theta) = M^{\tau}_{en}[I + \beta(I - \beta M^{\tau}_{in})^{-1}M^{\tau}_{in}][\pi^{\tau}(\theta^{\tau}_{P}) + \beta P^{\tau}_{exit}\,\theta^{\tau}_{X}], \tag{2.5.4}$$

where M^{τ}_{en} is a matrix with the transition probability between states conditional on entering in the market.

The first stage estimates the profit function, the transition probability for market state, exit and entry probabilities, and transition probabilities. With these estimates, continuation and entry values are estimated.

2.5.1 First stage estimation

Profit function

The parameters of the profit function can be estimated with profit and covariates data without imposing the dynamic model.

Unfortunately I cannot observe bank's profits at the market level, but I observe the deposits held in a bank in a market and I can compute average interest rates, average tax rates, wage expenditure, and other expenses at the bank level. Using the available data I compute a measure of bank's current profits using

$$\pi_{imt} = (1 - tax_{it})\left(q_{imt}\,(r^{L}_{it} - r^{D}_{it}) - wages_{imt} - other_{imt}\right),$$

where q_{imt} are deposits held by bank i in market m, r^{D}_{it} is the bank's average deposit interest rate, r^{L}_{it} is the bank's average loan interest rate (adjusted by expected loan losses), tax_{it} is the average tax rate paid by the bank, $wages_{imt}$ is the wage expenditure in the market, and $other_{imt}$ includes other costs incurred by the bank in the market.

I computed average deposit interest rate as the ratio of interest expense on deposits net of service charges on deposit accounts to total deposits. Average loan interest rate are interest and fee income on loans and income from lease financing receivables to total loans and lease financing receivables. To adjust the loan interest rate for different risks on the bank's loans, I subtract from the average loan interest rate the ratio of allowances for loans and lease losses to total loans and lease financing receivables. Average tax rate paid by the bank is the ratio of applicable income taxes to income before income taxes. Wage expenditure in the market is expenses on salaries and employees benefits allocated to the market using market deposits. Other costs include expenses of premises and fixed assets allocated to the market using market deposits.

The measure of profits is reasonable because loans and deposits are the main retail activities carried out by banks. Measurement error in average interest rates measured at the bank level rather than the bank-market level is a minor concern since empirical evidence supports MM banks charging uniform prices at the state level.[8] The estimating equation for the profit function is,

$$\pi^{\tau}_{imt} = \sum_{\tau=1}^{2} \mathbf{1}(i \in \tau)\, g^{\tau}(n^1_{mt}, n^2_{mt}; \theta^{\tau}_N) + \theta'_Z\, z_{mt} + \eta_m + u_{imt}, \tag{2.5.5}$$

where

$$\begin{aligned} g^{\tau}(n^1_{mt}, n^2_{mt}; \theta^{\tau}_N) = &\theta^{\tau}_{N,0} + \theta^{\tau}_{N,1} \times \text{presence of first type } \tau \text{ competitor} \\ &+ \theta^{\tau}_{N,2} \times \text{presence of second type } \tau \text{ competitor} \\ &+ \theta^{\tau}_{N,3} \times \text{additional type } \tau \text{ competitors} \\ &+ \theta^{\tau}_{N,4} \times \text{presence of first type } \tau' \text{ competitor} \\ &+ \theta^{\tau}_{N,5} \times \text{presence of second type } \tau' \text{ competitor} \\ &+ \theta^{\tau}_{N,6} \times \text{additional type } \tau' \text{ competitors,} \end{aligned}$$

[8]Some papers showing empirical evidence of uniform pricing are Biehl (2002) and Heitfield and Prager (2004).

where η_m is a market fixed effect, and z_{mt} includes population, income per capita, number of business establishments, and number of business employees in the market. Controlling for unobserved market profitability is important for two reasons. First, an unobserved variable positively correlated with both own profitability and rival's presence creates a positive bias in the estimates. Second, the estimated market fixed effects are used as unobserved correlated state variables in the dynamic game.[9] I estimate equation (2.5.5) by OLS with variance-covariance matrix robust to heteroscedasticity, and time series and within market correlation.

Transition probability for market state variables

To apply a non-parametric estimator I discretize the market state variables. First, I construct $\hat{z}_{mt} = \hat{\theta}'_Z z_{mt} + \eta_m$ to reduce the dimensionality of z. Then, I choose 10 group specific bins for $\hat{z}_{mt}$ such that they contain the same number of observations and I assign the mean value of $\hat{z}_{mt}$ to each bin. Finally, I estimate separate transition probabilities for each market type $\widehat{M}_z$ using a non-parametric estimator:

$$\widehat{M}_z(i,j) = \frac{\sum_{(m,t)\in T(z_i)} \mathbf{1}(z_{m,t+1} = z_j)}{\#T(z_i)}.$$

$M_z(i,j)$ is the estimated probability of being in state z_j tomorrow give the market is in state z_i, $T(z)$ is the set of observations with market state z, and $\#T(z)$ is the number of observations in $T(z)$.

Exit and entry probability

The exit probability of a type τ bank is estimated as the mean of observed exit probabilities. Let $T(n^1, n^2, z) = \{(m,t) : (n^1_{mt}, n^2_{mt}, z_{mt}) = (n^1, n^2, z)\}$ be the set of observations satisfying a given state configuration. Then the estimated exit probability

[9] Dunne, Klimek, Roberts, and Xu (2013) follow a similar procedure in their paper about dentists and chiropractors.

is

$$\widehat{P}^{\tau}_{exit}(s) = \frac{1}{\#T(n^1, n^2, z)} \sum_{(m,t)\in T(n^1,n^2,z)} \frac{x^{\tau}_{mt}}{n^{\tau}},$$

where x^{τ}_{mt} is the number of type τ banks that exit the market.

The entry probability of a type τ bank is estimated as the mean of observed entry probabilities:

$$\widehat{P}^{\tau}_{entry}(s) = \frac{1}{\#T(n^1, n^2, z)} \sum_{(m,t)\in T(n^1,n^2,z)} \frac{e^{\tau}_{mt}}{E^{\tau}_{mt}},$$

where e^{τ}_{mt} is the number of type τ entrants, and $E^{\tau}_{mt}$is the number of type τ potential entrants. In general is difficult to identify potential entrants so I follow different approaches. I estimate the maximum number of type τ banks in a market as $N^{\tau}_m = \max_t(n^{\tau}_{mt})$, so the number of potential entrants is $E^{\tau}_{mt} = N^{\tau}_m - n^{\tau}_{mt}$. Another approach is to assume one potential entrant of each type per market. Though both approaches are imperfect, the estimation results are robust to the chosen potential entrant definition.

Incumbent and potential entrant transition probability

Transition probability estimates are weighted non-parametric estimators with weights given by the number of incumbents that continue or the number of entrants. The probability of n^1_j, n^2_j conditional on being in n^1_i, n^2_i, z_i and the type τ incumbent continuing is estimated by

$$\widehat{M}^{\tau}_{in,n}(i,j) = \frac{\sum_{(m,t)\in T(n^1_i,n^2_i,z_i)} (n^{\tau}_i - x^{\tau}_{mt}) \mathbf{1}((n^1_{m,t+1}, n^2_{m,t+1}) = (n^1_j, n^2_j))}{\sum_{(m,t)\in T(n^1_i,n^2_i,z_i)} (n^{\tau}_i - x^{\tau}_{mt})}.$$

The probability of n_j^1, n_j^2 conditional on being in n_i^1, n_i^2, z_i and the type τ bank entering is estimated by

$$\widehat{M}^{\tau}_{en,n}(i,j) = \frac{\sum_{(m,t)\in T(n_i^1,n_i^2,z_i)} e^{\tau}_{mt}\mathbf{1}((n^1_{m,t+1},n^2_{m,t+1})=(n^1_j,n^2_j))}{\sum_{(m,t)\in T(n_i^1,n_i^2,z_i)} e^{\tau}_{mt}}.$$

Estimation of the type τ incumbent transition probability matrix $\widehat{M}^{\tau}_{in}$ derives directly from $\widehat{M}^{\tau}_{in,n}$ and $\widehat{M}_z$, and similarly for the type τ entrant transition probability matrix $\widehat{M}^{\tau}_{en}$.

Estimation of continuation and entry values

Plugging $\widehat{\pi}^{\tau}$, $\widehat{P}^{\tau}_{exit}$, and $\widehat{M}^{\tau}_{in}$ in equation (2.5.3) we obtain an estimation of the continuation value

$$\widehat{VC}^{\tau}(\theta) = \widehat{W}^{\tau}_{in,0}\,\widehat{\pi}^{\tau} + \widehat{W}^{\tau}_{in,1}\,\theta^{\tau}_X, \tag{2.5.6}$$

where $\widehat{W}^{\tau}_{in,0} = [I-\beta\widehat{M}^{\tau}_{in}]^{-1}\widehat{M}^{\tau}_{in}$, and $\widehat{W}^{\tau}_{in,1} = \widehat{W}^{\tau}_{in,0}\beta\widehat{P}^{\tau}_{exit}$. Similarly, plugging $\widehat{\pi}^{\tau}$, $\widehat{P}^{\tau}_{exit}$, $\widehat{M}^{\tau}_{in}$, and $\widehat{M}^{\tau}_{en}$ in equation (2.5.4) we obtain an estimation of the entry value

$$\widehat{VE}^{\tau}(\theta) = \widehat{W}^{\tau}_{en,0}\,\widehat{\pi}^{\tau} + \widehat{W}^{\tau}_{en,1}\,\theta^{\tau}_X, \tag{2.5.7}$$

where $\widehat{W}^{\tau}_{en,0} = \widehat{M}^{\tau}_{en}[I+\beta(I-\beta\widehat{M}^{\tau}_{in})^{-1}\widehat{M}^{\tau}_{in}]$, and $\widehat{W}^{\tau}_{en,1} = \widehat{W}^{\tau}_{en,0}\beta\widehat{P}^{\tau}_{exit}$. Estimated continuation and entry value are linear functions of the parameters of interest.

2.5.2 Second stage estimation

In the second stage, the estimated continuation and entry value are used to construct theoretical probabilities that depend on the parameters. The estimates are those parameter values than minimize a distance between theoretical and observed probabilities.

Plugging the estimated continuation value in equation (2.5.3) in the exit probability in equation (2.4.3) and using the distributional assumption, the theoretical exit probability is

$$Pr(\tau\, exit|s;\theta,\widehat{P}) = \exp\left\{-\frac{1}{\theta_X^\tau}[\widehat{W}_{in,0}^\tau(s)\,\widehat{\pi}^\tau(s) + \widehat{W}_{in,1}^\tau(s)\,\theta_X^\tau]\right\}. \tag{2.5.8}$$

$\widehat{P}$ denotes the exit and transition probabilities used to estimate the continuation value. Plugging the estimated entry value in equation (2.5.4) in the entry probability in equation (2.4.6) and using the distributional assumption, the theoretical entry probability is

$$Pr(\tau\, entry|s;\theta,\widehat{P}) = \frac{\exp\left\{\beta[\widehat{W}_{en,0}^\tau(s)\,\widehat{\pi}^\tau(s) + \widehat{W}_{en,1}^\tau(s)\,\theta_X^\tau] - \theta_E^\tau\right\}}{1+\exp\left\{\beta[\widehat{W}_{en,0}^\tau(s)\,\widehat{\pi}^\tau(s) + \widehat{W}_{en,1}^\tau(s)\,\theta_X^\tau] - \theta_E^\tau\right\}}.$$

Finally, I apply a minimum distance estimator that minimize a metric in the difference between theoretical and observed probabilities,

$$\widehat{\theta} = \arg\max_{\theta\in\Theta}(\widehat{\pi} - \widehat{h}(\theta))' A_T(\widehat{\pi} - \widehat{h}(\theta)),$$

where $\widehat{\pi} = (\widehat{P}_X^{1\prime}, \widehat{P}_X^{2\prime}, \widehat{P}_E^{1\prime}, \widehat{P}_E^{2\prime})'$ is the vector that stacks the reduce form probabilities for each state, $\widehat{h}(\theta) = (Pr^1(exit;\theta,\widehat{P})', Pr^2(exit;\theta,\widehat{P})', Pr^1(entry;\theta,\widehat{P})', Pr^2(entry;\theta,\widehat{P})')'$ is the vector that stacks the theoretical probabilities for each state, and A_T is matrix that weights the different equalities.

The weighting matrix A_T is block diagonal with blocks

$$A_T(i,i) = \begin{pmatrix} \frac{\#T(s_1)^2}{T^2} & \frac{2\,\#T(s_1)\#T(s_2)}{T^2} & \cdots & \frac{2\,\#T(s_1)\#T(s_S)}{T^2} \\ \frac{2\,\#T(s_1)\#T(s_2)}{T^2} & \frac{\#T(s_2)^2}{T^2} & \cdots & \frac{2\,\#T(s_2)\#T(s_S)}{T^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{2\,\#T(s_1)\#T(s_S)}{T^2} & \frac{2\,\#T(s_2)\#T(s_S)}{T^2} & \cdots & \frac{\#T(s_S)^2}{T^2} \end{pmatrix},$$

where $\#T(s)$ is the number of observation in state s, and T is the total number of observations. A_T is not the asymptotic optimal matrix but reduces the finite bias, and it is equivalent to the method of moments estimator proposed by POB. Pesendorfer and Schmidt-Dengler (2008) call this class of estimators asymptotic least square estimators and prove consistence and asymptotic normality. Usual standard errors are not valid due to the estimation error in the first stage thus standard errors are computed using a non-parametric bootstrap.

2.6 Empirical results

In this section I comment the results of the estimation. The main results are that a single-market bank has a profit advantage over a multimarket bank, but it pays a higher entry cost.

2.6.1 Profit function

The estimates of the profit function shown in table 2.8 have the expected sign, and are statistically significant. A single-market bank has an advantage in profits over a multimarket bank, increasing competition decreases profits, and a larger market size increases profits.

The single-market dummy is positive and significant at 1 percent: in mean the profits of a single-market monopolist bank is 0.2 million $ higher than the profits of a multimarket monopolist bank. For the market configuration more common in the data with 1 SM and 2 MM banks the model predicts an average current profit for a SM bank of 1.1 million $ and for a MM bank of 0.96 million $. This result is robust to different specifications of the profit function and alternative profit definitions. A related result is obtained by Adams, Brevoort, and Kiser (2007) who estimate a deposit demand using a generalized extreme value model, and find that a SM bank faces a less elastic demand than a MM bank in both rural markets and MSAs.

A plausible explanation for this SM advantage is the soft vs hard-information story.

The main idea is that there are different types of loans: transaction loans and relationship loans. Transaction loans are based on hard information like financial statements, collateral, covenants, credit scoring, etc. Relationship loans are based on soft information collected through repeated lender-borrower interactions. Soft information cannot flow easily within a formal organizational structure, and this creates an advantage for less hierarchical organizations like SM banks. It is natural to think on small businesses and farmers relying more in relationship lending.

Complementary evidence supporting this idea is provided by Berger, Bonime, Goldberg, and White (2004). Berger, Bonime, Goldberg, and White find an increase in entry of small banks after a large out-of-market bank acquires a small incumbent bank. The authors interpret their findings as small banks entering to supply credit to some relationship-dependent small businesses.

Another explanation is that the profit approximation includes the traditional retail banking activities: loans and deposits. A MM bank obtains a higher proportion of its profits from sources like brokerage fees, securitization, etc. which are not directly included in the profit approximation use in the empirical application. Then estimation results highlight that SM might have advantages in retail banking activities while MM may have advantages in non-traditional banking activities where economies of scale are more important.

The effect of increasing competition is negative, and almost all of the competition effects are statistically significant. When there exist 1 SM incumbent and 2 MM incumbents, the expected effect of an additional SM competitor is to decrease the profit of the SM incumbent in .086 million $ and to decrease the profit of a MM incumbent in .153 million $. While the expected effect of an additional MM competitor is to decrease the profit of the SM incumbent in .101 million $ and to decrease the profit of a MM incumbent in .087 million $.

The effect of increasing competition on profit is quantitatively similar for competitors of different types, and for the first, second, or additional competitors. Cohen and Mazzeo (2007) use data for banks and thrifts in 2001 and 2003, and exploit the cross-section variation in the number of competitors and market size to estimate a profit function for banks. They conclude that competition among banks of the same type is greater than competition among different types, and find decreasing effects of the number of competitors on profits. Adams, Brevoort, and Kiser (2007) also find higher cross price elasticities within types than between types.

The effect of average wage, number of business establishments, and number of employees in the geographic market have the expected positive sign while the effect of population and income per capita have an negative sign. These results are not surprising given that the market size regressors are highly collinear but I choose to keep all of them in the regression to capture more variability of profits. I tried alternatives functional forms: quadratic, logarithmic, interacted with income per capita. The estimates were robust to the different specifications I choose a more simple model with a linear functional form.

A possible concern for the second stage is the fit of the model to the data: the model explains 1.7 percent of the within profit variation. Although this is expected given such simple econometric model, it may signal the need of a richer model of firm heterogeneity to capture the variability observed in the profit data.

2.6.2 Market state variables

The methodology I apply for the estimation requires a discrete state space. Number of incumbents of each type is a discrete variable, but the market variables must be discretized. Following Dunne, Klimek, Roberts, and Xu (2013) to reduce the dimensionality of the market state variables I use the estimated coefficients to construct a new artificial variable that capture the effects of population and income per capita. I work with the market state variable $\hat{z}_{mt} = \hat{\theta}_Z z_{mt} + \eta_m$ where z_{mt} are market state variables, η_m is unobserved market profitability, and $\hat{\theta}_Z$ are the estimated coefficients of the market state variables in the first stage. Then, I choose 10 group specific bins for $\hat{z}_{mt}$ such that they contain the same number of observations and I assign the mean value of $\hat{z}_{mt}$ to each bin.

Table 2.9 shows descriptive statistics for each market state. More profitable markets tend to be larger and to have larger income per capita. This is the case despite the estimated coefficient of population and income per capita are negative in the profit estimation. The fixed effect captures differences that do not change over time between markets so it is capturing most of the variation in population and income per capita between markets. The average population is 2,496 inhabitants for the 1st group, and increases up to 55,000 for the 10th group. The increasing difference in average population between contiguous groups is due to the skewed population distribution.

As expected the number of banks, single-market banks, and multimarket banks are increasing in the market state. But the number of MM banks increases at a faster rate than the number of SM banks. It seems that MM banks presence is relatively greater in more profitable markets, while the SM presence is relatively greater in less profitable markets.

Population per bank and profit per bank are increasing in the market state. If we associate market state with population, the results are in line with Bresnahan and Reiss (1991)'s seminal paper. Increasing competition decreases markups, and a firm needs a larger demand to cover its fixed costs. Note that profit increases at the lower rate than population that seems to confirm the competition story.

2.6.3 Sell-off values and entry costs

I estimate the sell-off values and entry costs using a minimum distance estimator. I minimize the objective function using a Compass Search algorithm. The standard errors are computed using a non-parametric bootstrap.

The results of the sell-off value and entry cost estimation are in table 2.10. The sell-off estimates are basically zero that is not surprising given the low exit probability in the data. It is reasonable that banks do not close many branches in a period characterized by expansion in the number of branches. Likely many of the non-profitable branches were closed during the saving a loan crisis of the '80s. The reluctance of banks to close branches can be explained by brand concerns when closing branches.

The main result in the entry cost estimation is that SM banks face a higher entry costs than MM banks, and the cost of a bank which decides to enter in a new local market is around 10 million $ for a SM bank and 7 million $ for a MM bank. This paper is one of the first papers that attempt at measuring entry costs in the banking industry, and the results and interpretations should be consider as a first approach to the issue.[10]

The estimated differences in entry costs are driven by differences in profit: SM banks have a profit advantage, so they should face a higher entry cost if they enter in same proportion as MM banks. The result is robust to the pool of potential entrants used.

A SM bank should pay a entry costs which is around 3 million $ higher than a MM bank. In relative terms it is a 30 percent more expensive for a SM bank to enter in a new market than for a MM bank. There are some plausible explanations for the cost differential. In general, a SM bank that enters in a new market is a denovo bank and a denovo bank must pay start-up costs than a bank already operating avoid. Though, a MM bank that enter in a new state could face some red tape costs, it is reasonable to assume that are less important. Advertisement can be in part fixed at the bank, at in part fixed at the bank-market level. A MM bank has economies of scale advantages over the bank level or institutional advertisement. Another factor is hiring costs for management positions. A multimarket bank has many branches in different local markets, and could find it less costly to look for a manager for a new branch: directors can promote an employee to a manager position, or reallocate a manager from another branch. A single-market bank has to search for a manager in the job

[10] Up to my knowledge, Aguirregabiria, Clark, and Wang (2015) is the another recent paper that also seeks to measure to the entry costs in the U.S. banking industry.

market that is more costly.

Finally, as expected entry costs are higher the higher the market size. But though this effect is higher for MM banks than SM banks does not close the gap between both types of banks in larger markets.

2.7 Conclusions

Historic restrictions to the geographic expansion of banks have greatly affected the market structure of the U.S. banking industry. Recent deregulation, the most important the Riegle-Neal Act in 1994, foster a structural change in the market structure in the industry. These two facts motivate trying to understand the differences between two types of business models that coexist in the banking industry: single-market and multimarket banks.

An important result of the paper is that single-market incumbents are more profitable than multimarket incumbents in small local markets. This profit advantage is consistent with single-market banks having a comparative advantage in loans to small business and farmers.

Another feature that can differentiate single-market and multimarket banks is expansion costs, the cost to enter in a new local market. Measuring expansion costs is particularly relevant after recent interstate branching deregulation (Riegle-Neal Act) because it can help explain why single-market banks did not exploit new efficiency opportunities after the deregulation. It can also help in understanding the dynamics of the industry and some unexpected welfare consequences of the deregulation.

The second important results of the paper is that a single-market bank paid an entry costs which is 30 percent higher than the entry cost for a multimarket bank. This higher entry costs can be linked to start up costs, or higher advertisement and recruitment costs faced by single-market banks.

Appendix A: Construction of the data set

This appendix defines the relevant market for the U.S. banking industry, describes the data sources, and the construction of the entry/exit variables.

2.7.1 Definition of the relevant market

In this paper the relevant market for the U.S. banking industry is a geographic market: a Metropolitan Statistical Areas (MSA), a Micropolitan Statistical Area (MicroSA)[11], or a county outside a MSA or MicroSA. It is natural to relate MSAs to large urban markets, MicroSAs to small urban markets, and counties outside a MSA or MicroSA to rural markets.

Given the size of some geographic markets, distinct or overlapping submarkets might exist within a geographic market. In turn, this makes difficult to distinguish which banks are effectively competing with whom within the geographic market. For this reason I restrict the analysis to those markets with less than 100,000 inhabitants. This condition follows the selection criteria applied in Cohen and Mazzeo (2007).[12]

An example of a urban market in the sample is Bogalusa Micropolitan Statistical Area in Louisiana that had a population of 43,926 in 2000. Figure 2.5 shows the bank's branches located in Bogalusa. Almost all branches are located in the two biggest cities of the local market: Bogalusa and Franklinton.

The definition of a relevant market as a local geographic market seems appropriate for the U.S. banking industry. Indeed several papers present empirical evidence suggesting the relevant market for financial services is local.

A first strand of papers use survey data to document that households and small

[11]Micropolitan Statistical Areas are defined by the U.S. Census Bureau and it consist on a urban center plus adjoining counties. A Micropolitan Statistical Areas (MicroSA) has a smaller population than a MSA.

[12]Cohen and Mazzeo (2007) use Labor market areas (LMAs) defined by Bureau of Labor Statistics. LMAs are MSA, MicroSA, or small labor market areas. In New England, small labor market areas are based on a geographic level not available in my data. Hence, outside MSA and MicroSA I decided to work with individual counties rather than excluding New England states.

businesses use financial services from local institutions. Amel and Starr-McCluer (2002) using the 1998 Survey of Consumer Finances (SCF) find the median distance between a household and its depository institution is 3 miles, and for 75 percent of the households the distance is smaller than 10 miles. Moreover, 90 percent of checking accounts, savings accounts and certificates of deposits are acquired within 30 miles of home or workplace. Kwast, Starr-McCluer, and Wolken (1997) find similar results for small businesses. Using the 1993 Survey of Small Business Finances (SSBF) they find that more than 75 percent of small businesses did their banking business within 15 miles of their offices.

A second strand of papers estimate the demand of financial services by households, and the sensitivity of demand to changes in distance. Most of these studies use data on bank's branch locations (and branch characteristics), household locations (and household characteristics), market shares and prices to estimate a demand function. Ho and Ishii (2011) use data for MSA and counties in California, Oregon and Washington, and find the cross-price elasticities are very low for banks more than one mile away. Ishii (2005) use data for MSA in Massachusetts. Ishii finds that households are indifferent between a decrease in 1.1 miles in the distance to the nearest branch or an one standard deviation increase in the deposit interest rate. Finally, Wang (2009) use data on 132 isolated, middle-sized, U.S. geographic markets, and finds that a branch that is one mile closer is equivalent to a branch with one standard deviation higher deposit interest rate.

2.7.2 Data sources

The main data source is the Summary of Deposits (SOD) for the period 1994-2007 collected by the Federal Deposit Insurance Corporation (FDIC). The SOD contains data for all branches and offices owned by FDIC-insured institutions. The FDIC collects deposit balances for commercial and savings banks as of June 30 of each year, and the Office of Thrift Supervision (OTS) collects the same data for savings institutions.

Data are collected annually. The SOD contains information about location, ownership, and amount of deposits at the branch level.

I complemented the SOD data with balance sheet data at the bank level. I used data for the period 1994-2007 from the Call Reports of the Federal Reserve Bank of Chicago (commercial banks) and the Thrift Financial Reports of the Office of Thrift Supervision (saving institutions). I used the balance sheet data to compute average deposit interest rates, average loan interest rates, and other financial ratios.

2.7.3 Construction of entry/exit variables

The SOD reports information for almost all depository institutions in the United States. But given that it is administrative data it required a lengthy cleaning process before the data is suitable for research purposes. In particular the construction of entry/exit variable is sensitive to measurement error problems. For example, missing information in a year can be interpreted as exit and subsequent entry, or consolidation of bank charters by a Multibanking Holding Company can be interpreted as bank exit and bank entry in several markets.

In the SOD there is a bank code (variable `cert`) and a branch code (variable `uninumbr`). The variable `uninumbr` can be used to build a panel data of branches, but such a variable presents some problems. First, if there is an ownership change of the branch, `uninumbr` usually changes for the same branch. Second, there are missing values for `uninumbr`. Missing values in `uninumbr` are more pervasive for savings institutions whose information is collected by the OTS.[13]

Given the above-mentioned problems with the branch code variable, I relied also on the information provided by other branch variables. I matched observations using the branch's code, address, city, ZIP code, state, county, bank code, and bank holding company code. I used a *fuzzy merge*[14] that links observations based on a matching

[13]`uninumbr` is missing for 8.65 percent of the commercial banks and for 57.18 percent of the saving banks.

[14]I applied the fuzzy merge using the command `reclink` in Stata.

algorithm. If the matching score is higher than a user given threshold, two observations are linked. I applied the fuzzy merge recursively to observations in t and $t+1$.

Once the panel was constructed at the branch level, I defined the relevant variables to study the market dynamics: entry, exit, and ownership changes. A branch entry refers to the opening of a new branch, and a branch exit refers to closing an existing branch. Hence, entry is defined as

$$entry_{mt} = active_{mt} \times (1 - active_{m,t-1}),$$

and exit is defined as,

$$exit_{mt} = (1 - active_{mt}) \times active_{m,t-1},$$

where $active_{mt}$ is an indicator function which equals 1 if the branch is active in market m at time t.

Part of the dynamics in the industry is due to changes in ownership. I considered a change in ownership when an active branch experiences a change in its bank charter (variable `cert`). There are several economic situations behind a change of ownership: a merger or acquisition of a bank, a bank holding company that consolidates the bank charters of its subsidiaries, or a bank which sells individual branches. I distinguished a change in ownership by merger or acquisition (M&A) from other reasons using data from the Federal Reserve Bank of Chicago about mergers. As expected most of the changes of ownership matched mergers at the bank level.

I dropped some observations from the final data set. First, I deleted a branch that satisfy at least one of the following conditions: (1) it does not have deposits in more than 75 percent of the observations, (2) it has only 1 observation with positive deposits, (3) it is active for only 1 period, or (4) it has less than 100,000 $ of deposits in average. Second, I consolidated branches with the same address and bank code. Third, I made linear interpolations for deposits in case there is a gap of one period in the observations of a branch.

Appendix B: Figures and tables

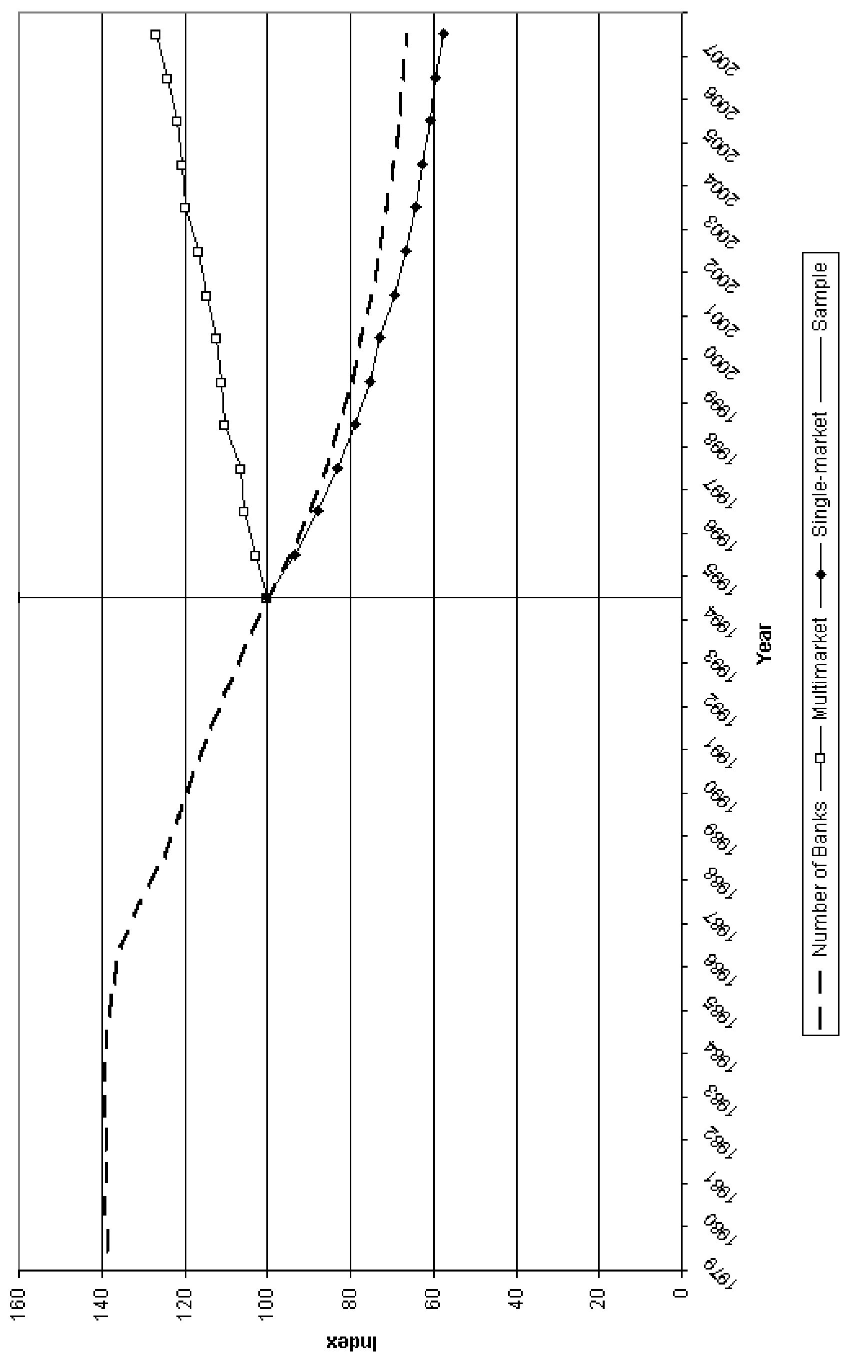


Figure 2.1: Number of banks. U.S. commercial and saving banks. 1979-2007. Index 1994=100.

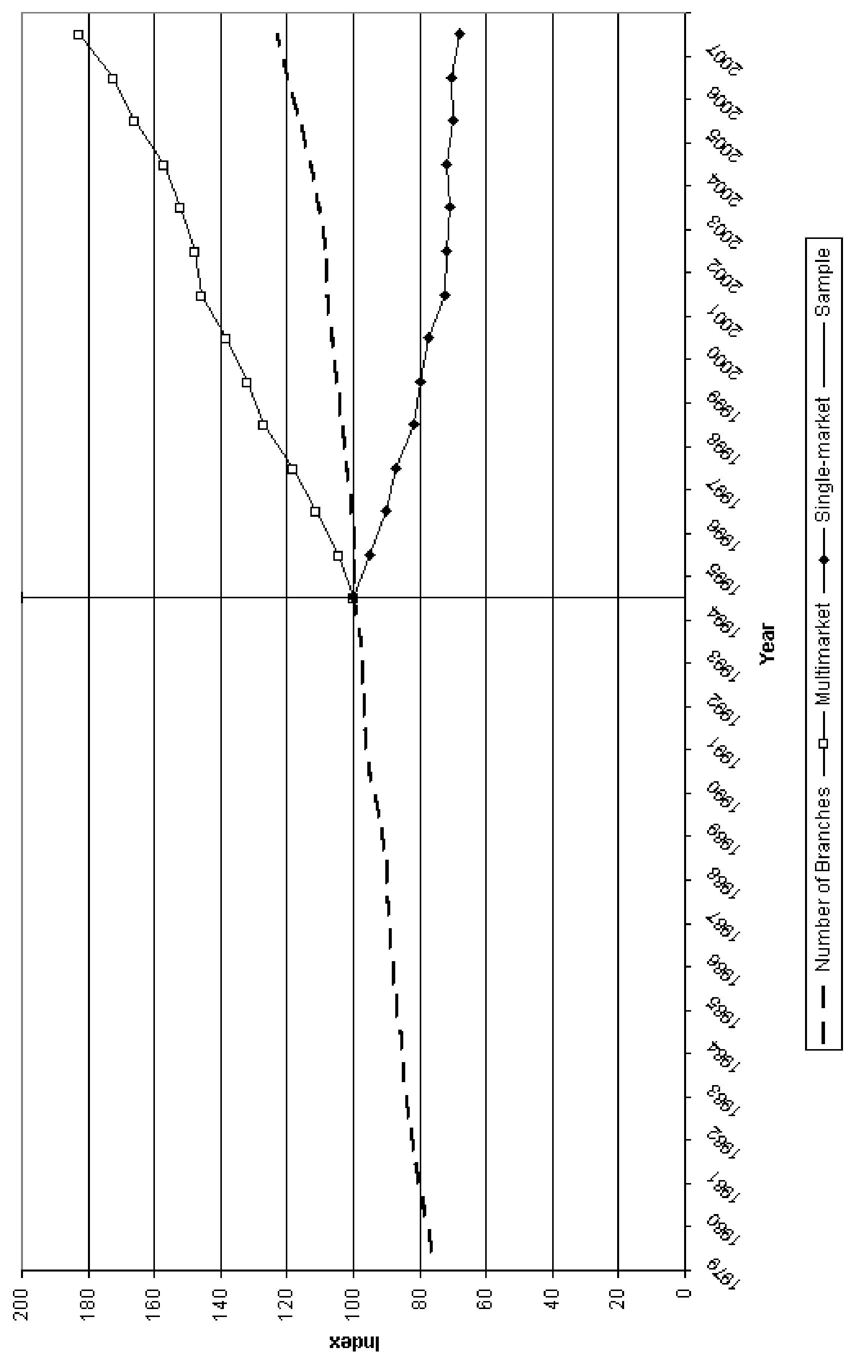


Figure 2.2: Number of branches. U.S. commercial and saving banks. 1979-2007. Index 1994=100.

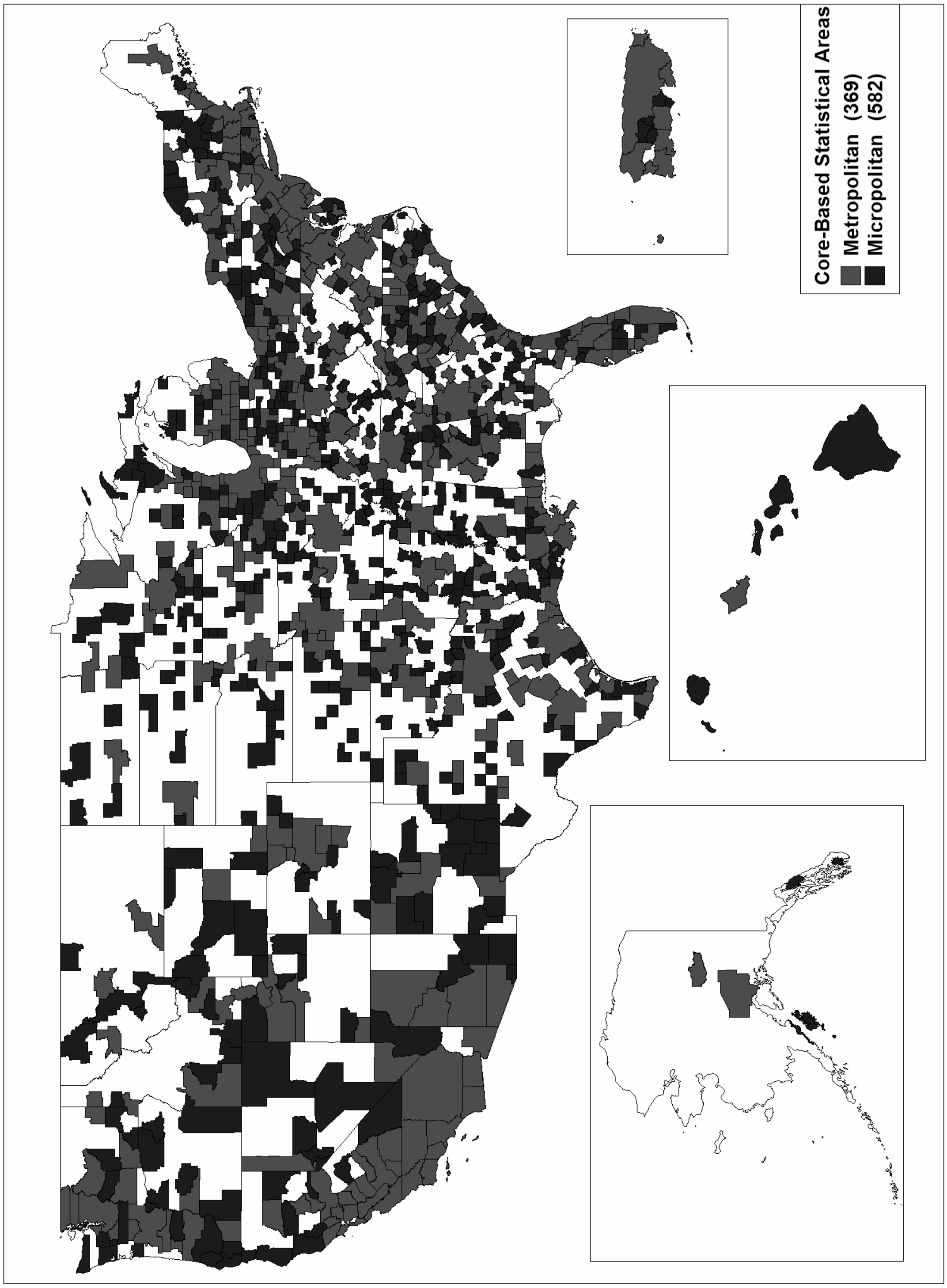


Figure 2.3: Metropolitan and Micropolitan Statistical Areas.

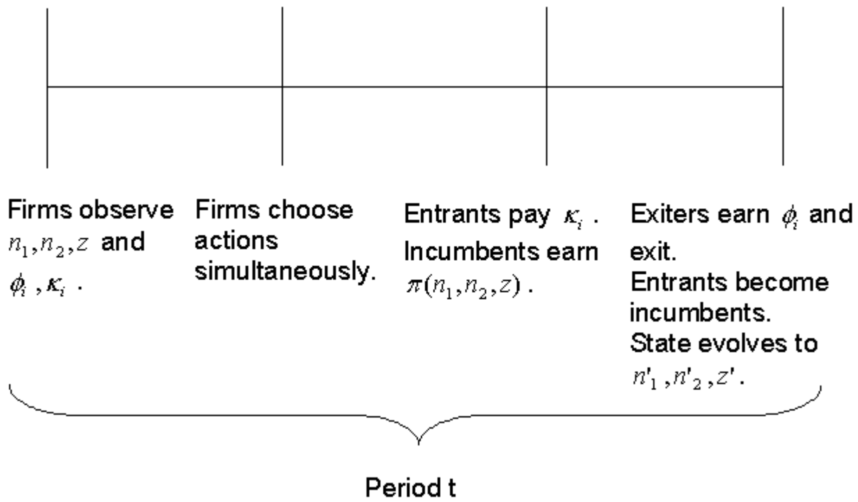
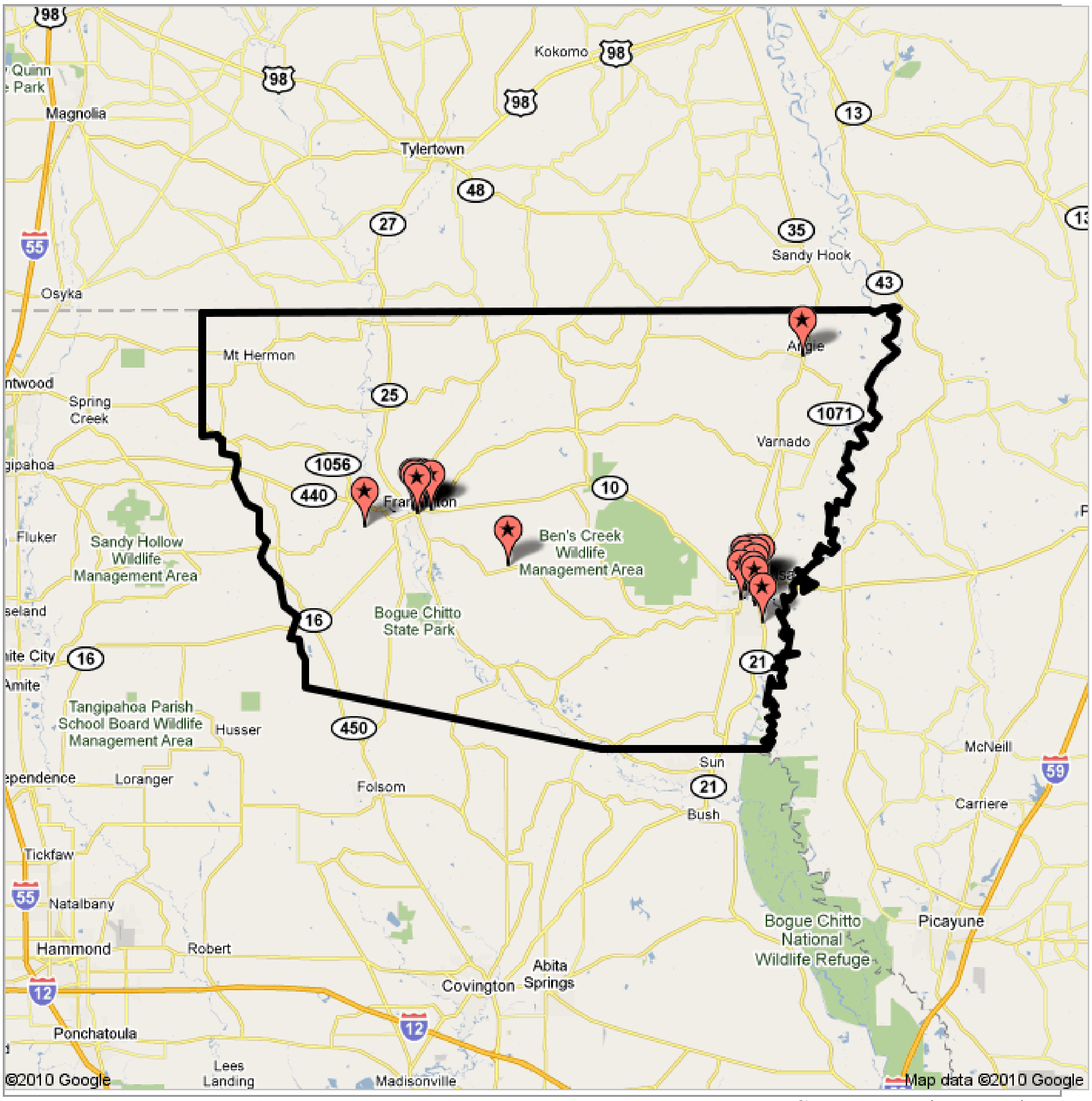


Figure 2.4: Timing of the game in period t.

9/21/2010 Google Maps



Untitled

Figure 2.5: Bank's branch locations in Bogalusa Micropolitan Statistical Area, LA.

Credit Union Service Centers

Credit Union Service Centers

balloon_text	$[name]Ste 100, 5500 Veterans Memorial BlvdMetairie, LA 70003-1746(504) 455-7764 $[geDirections]
contents	Credit Union Service Centers
city	Metairie
region	US-LA

Zellco Federal Credit Union

Table 2.1: Evolution of concentration and market structure in the banking industry, U.S., 1994-2007

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Aggregate level														
C4 (percentage)	8	7	8	10	13	14	15	16	17	18	21	26	25	26
HHI	33	32	36	53	71	78	107	114	126	137	157	221	210	215
Number of banks	12,928	12,221	11,630	11,125	10,677	10,288	10,060	9,703	9,433	9,219	9,034	8,825	8,733	8,567
Single-market	11,325	10,564	9,929	9,413	8,908	8,507	8,246	7,865	7,559	7,296	7,101	6,873	6,749	6,535
Multimarket	1,602	1,656	1,700	1,711	1,768	1,780	1,813	1,837	1,873	1,922	1,932	1,951	1,983	2,031
Number of branches	75,468	75,692	76,412	77,651	78,856	79,946	81,121	81,716	82,411	83,556	85,542	87,855	90,435	92,824
Single-market	39,155	37,389	35,541	34,258	32,252	31,498	30,327	28,460	28,215	27,847	28,230	27,350	27,652	26,455
Multimarket	36,300	38,290	40,858	43,379	46,589	48,434	50,780	53,242	54,181	55,694	57,297	60,489	62,766	66,353
Large urban markets[a]														
C1 (percentage)	25	25	26	26	26	26	26	26	26	26	26	27	26	26
HHI	1,440	1,461	1,491	1,498	1,519	1,512	1,507	1,505	1,505	1,503	1,518	1,538	1,514	1,505
Number of banks	29	28	27	27	27	27	27	27	27	27	27	27	28	29
Single-market	19	18	17	17	16	15	15	15	14	14	14	14	14	14
Multimarket	9	10	10	10	11	11	12	13	13	13	13	14	14	15
Small urban and rural markets[a]														
C1 (percentage)	45	45	45	45	45	45	45	45	45	45	45	45	45	45
HHI	3,585	3,567	3,555	3,549	3,537	3,532	3,517	3,500	3,494	3,495	3,488	3,471	3,438	3,437
Number of banks	5	5	5	5	6	6	6	6	6	6	6	6	6	6
Single-market	3	3	3	3	2	2	2	2	2	2	2	2	2	2
Multimarket	2	3	3	3	3	3	3	3	4	4	4	4	4	4

Notes: [i] Source: FDIC Summary of Deposits. [ii] C1 is the deposit market share of the largest bank in the market; C4 is the deposit market share of the four largest banks in the market; and HHI is the Herfindahl-Hirschman index computed with deposit market shares. [iii] Single-market banks are banks that hold more than 80 percent of its deposits in a single local market. Multimarket banks are banks not classified as single-market banks. [iv] Large urban markets are Metropolitan Statistical Areas defined by the U.S. Census Bureau. Small urban and rural markets are Micropolitan Statistical Areas defined by the U.S. Census Bureau and rural counties. [a] Means over local markets.

Table 2.2: Sources of variation in the number of bank branches, Large urban markets, U.S., 1994-2006.

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
All banks													
Branch openings	3.5	4.0	5.5	5.0	4.7	4.2	3.9	3.6	3.9	4.8	5.0	5.0	4.9
Branch closings	-3.4	-3.1	-3.9	-3.5	-3.3	-2.6	-3.2	-2.6	-2.1	-1.8	-1.7	-1.5	-1.8
Net change	*0.2*	*0.9*	*1.6*	*1.6*	*1.4*	*1.6*	*0.7*	*1.0*	*1.7*	*3.0*	*3.3*	*3.5*	*3.1*
Single-market banks													
Branch openings	2.1	2.3	2.7	2.5	2.5	2.3	2.0	1.8	1.7	1.6	1.9	2.0	1.8
Branch closings	-1.6	-1.3	-1.5	-1.0	-0.9	-0.8	-0.7	-0.7	-0.6	-0.5	-0.5	-0.4	-0.4
Net bank change by M&A	-1.5	-2.1	-1.9	-1.9	-1.7	-2.2	-1.7	-0.8	-0.5	-0.9	-0.7	-0.6	-1.7
Net switch SM to MM	-1.3	-1.5	-0.9	-2.3	-0.5	-0.8	-1.9	-0.3	-0.8	0.6	-1.8	-0.4	-1.1
Net change	*-2.3*	*-2.7*	*-1.6*	*-2.7*	*-0.6*	*-1.5*	*-2.3*	*-0.1*	*-0.2*	*0.8*	*-1.1*	*0.6*	*-1.3*
Multimarket banks													
Branch openings	1.5	1.8	2.8	2.5	2.2	1.9	1.9	1.8	2.2	3.2	3.1	3.0	3.0
Branch closings	-1.8	-1.8	-2.4	-2.4	-2.4	-1.8	-2.4	-1.8	-1.5	-1.3	-1.2	-1.1	-1.4
Net bank change by M&A	1.5	2.1	1.9	1.9	1.7	2.2	1.7	0.8	0.5	0.9	0.7	0.6	1.7
Net switch SM to SM	1.3	1.5	0.9	2.3	0.5	0.8	1.9	0.3	0.8	-0.6	1.8	0.4	1.1
Net change	*2.4*	*3.6*	*3.2*	*4.3*	*2.0*	*3.1*	*3.1*	*1.1*	*2.0*	*2.1*	*4.4*	*2.8*	*4.4*

Notes: [i] Source: FDIC Summary of Deposits. [ii] Single-market banks are banks that hold more than 80 percent of its deposits in a single local market. Multimarket banks are banks not classified as single-market banks. [iii] Large urban markets are Metropolitan Statistical Areas defined by the U.S. Census Bureau. Small urban and rural markets are Micropolitan Statistical Areas defined by the U.S. Census Bureau and rural counties. [iv] A branch opening means a denovo branch, a branch closing means a branch closed by a bank, a net bank change by M&A means branches owned by a different bank due to a bank merger, and a switch SM to MM means a single-market bank becoming a multimarket bank. [v] Percentage over the number of incumbents.

Table 2.3: Sources of variation in the number of bank branches, Small urban and rural markets, U.S., 1994-2006.

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
All banks													
Branch openings	2.9	3.0	3.3	3.9	3.4	3.1	2.8	2.3	2.2	1.9	2.1	2.3	2.3
Branch closings	-2.1	-1.9	-1.6	-2.4	-2.1	-2.0	-2.0	-2.0	-1.8	-1.4	-1.2	-1.2	-1.2
Net change	*0.7*	*1.0*	*1.7*	*1.5*	*1.3*	*1.1*	*0.8*	*0.3*	*0.3*	*0.5*	*0.9*	*1.1*	*1.2*
Single-market banks													
Branch openings	1.6	1.8	1.9	1.9	1.7	1.6	1.3	1.1	1.1	0.9	0.9	0.9	1.1
Branch closings	-0.9	-0.8	-0.7	-0.9	-0.6	-0.6	-0.5	-0.5	-0.5	-0.3	-0.2	-0.3	-0.3
Net bank change by M&A	-1.7	-1.3	-1.4	-1.3	-1.8	-1.1	-1.4	-0.6	-0.5	-0.2	-0.5	-0.5	-0.8
Net switch SM to MM	-1.6	-1.3	-1.6	-1.9	-1.4	-1.2	-1.6	-1.0	-1.1	-1.0	-1.0	-0.8	-1.3
Net change	*-2.6*	*-1.6*	*-1.9*	*-2.2*	*-2.2*	*-1.3*	*-2.2*	*-1.0*	*-1.1*	*-0.7*	*-0.7*	*-0.7*	*-1.3*
Multimarket banks													
Branch openings	1.3	1.2	1.5	2.0	1.7	1.4	1.5	1.2	1.1	1.0	1.2	1.3	1.3
Branch closings	-1.3	-1.2	-0.8	-1.6	-1.4	-1.4	-1.5	-1.5	-1.3	-1.1	-1.0	-0.9	-0.9
Net bank change by M&A	1.7	1.3	1.4	1.3	1.8	1.1	1.4	0.6	0.5	0.2	0.5	0.5	0.8
Net switch SM to MM	1.6	1.3	1.6	1.9	1.4	1.2	1.6	1.0	1.1	1.0	1.0	0.8	1.3
Net change	*3.3*	*2.7*	*3.6*	*3.7*	*3.5*	*2.4*	*3.0*	*1.4*	*1.5*	*1.2*	*1.6*	*1.8*	*2.5*

Notes: [i] Source: FDIC Summary of Deposits. [ii] Single-market banks are banks that hold more than 80 percent of its deposits in a single local market. Multimarket banks are banks not classified as single-market banks. [iii] Large urban markets are Metropolitan Statistical Areas defined by the U.S. Census Bureau. Small urban and rural markets are Micropolitan Statistical Areas defined by the U.S. Census Bureau and rural counties. [iv] A branch opening means a denovo branch, a branch closing means a branch closed by a bank, a net bank change by M&A means branches owned by a different bank due to a bank merger, and a switch SM to MM means a single-market bank becoming a multimarket bank. [v] Percentage over the number of incumbents.

Table 2.4: Descriptive statistics for local markets, Selected local markets, U.S., 1994-2007.

Variable	Mean	s.d.	Median	Percentiles	
				1%	99%
Population	20,097	16,916	15,044	844	78,184
Population density (per squared mile)	37	90	23	1	169
Income per capita (2007 $)	21,545	4,381	21,164	13,234	34,387
Average wage (2007 $)	23,295	3,928	22,853	16,438	35,383
Number of business employees	7,633	7,221	5,172	290	33,834
Number of business establishments	470	418	343	17	1,951

Notes: [i] Source: Local Area Personal Income of the Bureau of Economic Analysis (population, income per capita, average wage, number of business employees), County of Business Patterns of the U.S. Census Bureau (number of business establishments), and Summary File 3 of the U.S. Census 2000 (land area). [ii] Sample: Small urban and rural markets with less than 100,000 inhabitants, less than 8 SM incumbent banks, and less than 8 MM incumbent banks.

Table 2.5: Descriptive statistics for single-market and multimarket banks, U.S., 1994 and 2007.

	1994		2007	
	SM	MM	SM	MM
Bank level				
Number of markets	1	14	2	73
Number of states	1	2	1	5
Number of branches	4	60	6	570
Employees/branches	12	13	12	18
Multibank holding company	31	50	25	58
Loans & leases/assets	53	61	62	69
Real estate loans/loans & leases	50	58	62	65
Agricultural loans/loans & leases	20	11	13	8
Commercial & industrial loans/loans & leases	13	13	14	15
Loans to individuals/loans & leases	17	18	10	9
Loans to small businesses/loans & leases	23	22	26	20
Loans to small farms/loans & leases	29	16	21	12
Non-performing Loans/loans & leases	1.0	0.9	1.1	1.0
Equity/assets	10.0	8.1	11.4	10.3
Return on equity (ROE)	12.1	11.9	10.7	11.1
Return on assets (ROA)	1.2	0.9	1.1	1.1
Deposit interest rate	2.6	2.5	2.5	2.3
Loan interest rate	6.9	6.8	6.6	6.3
Bank-market level				
Number of branches	1.7	1.7	1.7	1.8
Deposits per branch	27,460	23,474	34,934	33,729

Notes: [i] Source: Reports of Condition and Income from Federal Reserve Bank of Chicago and Thrift Financial Reports from the Office of Thrift Supervision.
[ii] Sample: Banks with branches in the selected local markets. [iii] Single-market banks are banks that hold more than 80 percent of its deposits in a single local market. Multimarket banks are banks not classified as single-market banks. [iv] Variables computed at the bank level are means weighted by the number of branches; variables computed at the bank-market level are simple means.

Table 2.6: Descriptive statistics for the market structure, Selected local markets, U.S, 1994-2007

		Number of multimarket banks									
		0	1	2	3	4	5	6	7	8	Total
Number of single-market banks	0	0.9	3.8	4.6	4.2	3.4	2.1	1.7	0.8	0.2	21.7
	1	2.7	4.7	5.8	4.4	3.2	2.6	1.4	0.8	0.2	25.8
	2	2.7	4.9	3.7	3.6	2.7	2.0	1.2	0.7	0.2	21.7
	3	1.7	2.4	2.8	3.0	1.9	1.2	0.9	0.3	0.1	14.4
	4	0.9	1.1	1.8	1.4	1.2	0.9	0.5	0.2	0.1	8.2
	5	0.4	0.7	0.8	0.8	0.7	0.5	0.3	0.1	0.1	4.3
	6	0.1	0.5	0.5	0.5	0.4	0.3	0.1	0.1	0.0	2.5
	7	0.1	0.2	0.2	0.2	0.3	0.2	0.0	0.0	0.0	1.1
	8	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.4
	Total	9.5	18.3	20.3	18.1	13.9	9.8	6.2	3.0	0.9	100.0

Notes: [i] Source: Summary of Deposits (FDIC).
[ii] Sample: Small urban and rural markets with less than 100,000 inhabitants, less than 8 SM incumbent banks, and less than 8 MM incumbent banks. [iii] Single-market banks are banks that hold more than 80 percent of its deposits in a single local market. Multimarket banks are banks not classified as single-market banks. [iv] Percentage of total market-year observations.

Table 2.7: Entry and exit statistics, Selected local markets, U.S, 1994-2007

Number of Incumbents	Entry proportion (%)			Exit rate (%)		
	All	SM	MM	All	SM	MM
1	3.52	2.21	1.31	0.69	0.34	0.34
2	2.19	1.04	1.15	0.43	0.11	0.32
3	1.90	1.05	0.85	0.51	0.19	0.32
4	1.78	0.81	0.97	0.67	0.20	0.48
5	1.89	1.03	0.87	0.70	0.21	0.48
6	1.44	0.74	0.69	0.53	0.19	0.34
7	1.53	0.77	0.75	0.55	0.17	0.38
8	1.53	0.80	0.73	0.63	0.20	0.43
9	1.12	0.62	0.50	0.57	0.16	0.41
10 or +	0.01	0.01	0.00	0.01	0.00	0.00
Average	1.62	0.85	0.77	0.59	0.20	0.40

Notes: [i] Source: Summary of Deposits (FDIC). [ii] Sample: Small urban and rural markets with less than 100,000 inhabitants, less than 8 SM incumbent banks, and less than 8 MM incumbent banks. [iii] Single-market banks are banks that hold more than 80 percent of its deposits in a single local market. Multimarket banks are banks not classified as single-market banks. [iv] Entry proportion is the ratio of entrants to incumbents, and exit rate is the ratio of exits to incumbents.

Table 2.8: Estimation of the profit function

Dependent variable: bank-market profits (in million of 2007 $)	
Single-market bank dummy	0.205***
	(0.039)
Effects of competition on single-market banks	
First single-market competitor	-0.086**
	(0.036)
Second single-market competitor	-0.091
	(0.076)
Additional single-market competitors	-0.137***
	(0.014)
First multimarket competitor	-0.040
	(0.025)
Second multimarket competitor	-0.141***
	(0.025)
Additional multimarket competitors	-0.101**
	(0.042)
Effects of competition on multimarket banks	
First single-market competitor	-0.117***
	(0.034)
Second single-market competitor	-0.153***
	(0.016)
Additional single-market competitors	-0.081***
	(0.015)
First multimarket competitor	-0.072***
	(0.026)
Second multimarket competitor	-0.087**
	(0.036)
Additional multimarket competitors	-0.141***
	(0.016)
Market size controls	
Log(population)	-0.320***
	(0.103)
Log(income per capita)	-0.184**
	(0.072)
Log(wage)	0.042
	(0.072)
Log(n of establishments)	0.178**
	(0.073)
Log(n of employees)	0.369***
	(0.070)
Market fixed effects	Yes
Number of observations	111,484
Number of markets	1,678
Mean of dependent variable	0.703

Notes: [i] Sample: Small urban and rural markets with less than 100,000 inhabitants, less than 8 SM incumbent banks, and less than 8 MM incumbent banks.
[ii] Asymptotic standard errors robust to heteroscedasticity, and time series and within market correlation. [iii] Significance level: *** 1%, ** 5%, and * 10%.

Table 2.9: Descriptive statistics for the market state variable

Market state	N. of obs.	Population	Income per capita	N of banks	N of SM banks	N of MM banks	Pop. per bank	Profits per bank
1	2,367	2,496	20,479	1.55	0.73	0.82	1,795	0.258
2	2,367	5,189	21,254	2.73	1.24	1.49	2,404	0.363
3	2,368	8,096	20,609	3.32	1.56	1.76	2,998	0.439
4	2,367	10,675	20,676	3.85	1.71	2.14	3,390	0.503
5	2,367	13,307	21,230	4.65	2.04	2.61	3,474	0.522
6	2,368	17,154	21,435	5.06	2.12	2.94	3,927	0.612
7	2,367	21,933	21,540	5.35	2.21	3.15	4,572	0.723
8	2,368	28,423	21,694	6.14	2.40	3.74	5,166	0.767
9	2,367	38,165	22,172	7.04	2.58	4.47	5,974	0.818
10	2,368	55,516	24,364	7.88	2.68	5.20	7,355	1.117

Notes: [i] Sample: Small urban and rural markets with less than 100,000 inhabitants, less than 8 SM incumbent banks, and less than 8 MM incumbent banks.
[ii] Single-market banks are banks that hold more than 80 percent of its deposits in a single local market. Multimarket banks are banks not classified as single-market banks. [iii] The statistics for population, income per capita, number of banks, SM banks, MM banks, and population per bank are means over market-year observations, and the statistics for profits per banks are means over bank-market-year observations.

Table 2.10: Sell-off Value and Entry Cost Parameters (in million of 2007 $)

	Potential Entrant Definition 1	Potential Entrant Definition 2
Mean Entry Cost		
Single-market bank	10.374***	10.434***
	(1.037)	(0.972)
Multimarket bank	7.058***	7.070***
	(0.632)	(0.623)
Single-market bank × market size	1.189***	1.197***
	(0.261)	(0.265)
Multimarket bank × market size	1.387***	1.396***
	(0.315)	(0.325)
Mean Sell-off value		
Single-market bank	0.000	0.000
	(0.180)	(0.181)
Multimarket bank	0.008	0.008
	(0.014)	(0.013)
Number of Observations	22,425	
Number of Markets	1,725	

Notes: [i] Sample: Small urban and rural markets with less than 100,000 inhabitants, less than 8 SM incumbent banks, and less than 8 MM incumbent banks. [ii] Estimation: Minimum distance estimator with weighting matrix that replicates GMM in POB. Optimization using Compass Search. [iii] Bootstrap standard errors with 500 simulations. [iv] Significance level: *** 1%, ** 5%, and * 10%.

Chapter 3

Industry equilibrium with open-source and proprietary firms (with Gastón Llanes)

3.1 Introduction

Collaboration in research enhances the chances of discovery and creation. This is true not only for scientific discoveries, but also for commercial innovations. However, innovators face incentives to limit the access of competitors to their innovations. According to the traditional view in the economics of innovation, innovators innovate because they obtain a monopolistic advantage over their competitors. Therefore, innovators should prevent others from gaining access to their discoveries, either by keeping them secret or by protecting them with patents.

This view contrasts with the open-source development model, which has been intensively used in the software industry and in other industries at various points in time, as documented in the next section. In open source, developers voluntarily choose to disclose their technological improvements so that they can be copied, used and improved by other innovators free of charge. But if everybody has access to the same technologies, then how do developers benefit from their collaborations? What do they receive in exchange for renouncing their monopolistic advantage? The answer is that open-source producers profit by selling goods and services which are complementary to the open-source good.

The case of the Linux operating system is a good example. Linux receives substantial contributions of commercial firms like IBM, HP and Red Hat, among others, which benefit from selling complementary goods and services. For example, IBM sells consulting services and complementary proprietary software, HP sells personal computers and computer servers, and Red Hat sells training and support services.

Still, this leaves open the questions of what determines the choice of development model for profit-maximizing firms, why open-source and proprietary firms coexist in the same markets, and what are the implications of coexistence on market structure and investments incentives. Existing literature has yet to address these questions, which are instead the main focus of this paper.

We present a model of industry equilibrium with endogenous technology sharing. Firms sell packages composed of a primary good, such as software, and a complementary good, such as a smartphone, tablet PC, or support and training services. Firms choose their development model (open-source or proprietary), how much to invest in product development, and the price of their products. Open-source firms share their improvements to the main product, while proprietary firms, develop their products independently from other firms. Consumers value the quality of both goods (vertical differentiation) but also have idiosyncratic tastes for the products of different firms (horizontal differentiation).

We find that the equilibrium may have both types of firms, or only open-source firms. In the equilibrium with coexistence, the market structure is asymmetric, with large proprietary firms selling a high quality product and small open-source firms selling a low quality product. This finding is consistent with the observations of recent surveys. Seppä (2006) compares both types of firms, and finds that open-source firms tend to be smaller than proprietary firms. Bonaccorsi and Rossi (2004) show that the most important motive for firms to participate in open-source projects is that it allows small firms to innovate.

The equilibrium depends on the resolution of a trade-off between free-riding and collaboration, which is governed by a parameter measuring the degree of public good of the investment in R&D (i.e. the effect of total vs. individual contributions on quality). When open-source firms invest in R&D, they increase quality for all firms in the project. As a consequence, open-source firms are able to appropriate a smaller fraction of their investment, in comparison with proprietary firms. Nevertheless, open-source firms share their advances on the primary good, which means that, even though each firm may individually invest less than a proprietary firm, the total investment of all firms in the project may be larger than the investment of a proprietary firm.

When the degree of public good of the investment in R&D is high, free-riding is very important, which leads to lower individual investments for open-source firms. As a consequence, proprietary firms have an advantage over open-source firms in terms of market share and price. On the other hand, open-source firms benefit from lower development costs. Therefore, both types of firms coexist in equilibrium: some firms choose to be proprietary, have a high investment in R&D, and benefit from high market shares and prices, and other firms choose to be open source and benefit from lower development costs.

When the degree of public good of the investment in R&D is low, the positive effects of collaboration are stronger than the negative effects of free-riding, and open-source firms have higher (total) investment than proprietary firms (individual investments are similar, but open-source firms share their investments). In this case, all firms choose the open-source development model to benefit from higher market shares and lower development costs than proprietary firms.

In the market equilibrium, welfare is suboptimal because of the public good problem in open source and the duplication of effort of proprietary firms. In Section 3.6, we show that a subsidy to open-source development can improve welfare not only because it increases the investment in R&D, but also because it encourages commercial firms to participate in open source, enhancing collaboration as a result.

We also study investment incentives and market structure under free entry. When entry costs are small, the number of firms is large and the market becomes monopolistically competitive. The equilibrium of the limiting economy depends on the limit of the ratio of investments in R&D of open-source and proprietary firms. Even though free-riding becomes more important as the number of firms increases, collaboration becomes more important too, so either type of firm may have an advantage.

We find that when the degree of public good of the investment is at its maximum level (all investment is shared), the effects of free-riding and collaboration are perfectly balanced, and the equilibrium of the limiting economy has both types of firms. In this case, as the degree of horizontal differentiation decreases, the aggregate market share of open-source firms decreases, but the proportion of open-source firms in the total of firms increases. Thus, there are fewer but bigger proprietary firms in equilibrium. When the degree of public good of the investment is less than maximal, on the other hand, collaboration dominates free-riding and all firms become open-source. Thus, we find conditions under which *large numbers favor cooperation*, i.e. open source does not disappear as the number of firms grows.

We present several extensions to the baseline model in section 3.7. The baseline model assumes symmetric consumer preferences for open-source and proprietary products. However, given that open-source packages are based on the same primary good, open-source products are likely to be more similar than the products of proprietary firms. We modify the baseline model to allow for a higher cross-price elasticity between open-source products. We find that the main result of the paper still holds: when open-source and proprietary firms coexist, the market share of proprietary firms is higher than that of open-source firms. However, in this case, we also find that if the substitution between open-source products is large enough, there are equilibria in which all firms choose the proprietary development model.

Chapter 3: Industry equilibrium with open-source and proprietary firms (with Gastón Llanes)

The equilibrium with open-source and proprietary firms is characterized by an asymmetric market structure, even though all firms are ex-ante symmetric. We argue that this result is even stronger if there are initial asymmetries in firm size. Larger firms ex-ante have more incentives to remain proprietary, and the difference in market shares between open source and proprietary will tend to increase.

In the baseline model, firms decide on total investment in R&D which is not fully appropriable for open-source firms. In an extension we analyze what happens when firms make a fully appropriable, direct investment in the complementary good. We show that as the importance of the direct investment in the complementary on total quality increases, the number of firms in open source decreases. If the effect of direct investment is high enough, the equilibrium has both kinds of firms for all parameter values. Therefore, coexistence becomes more likely when firms can invest directly in the complementary good.

Finally, we study the equilibrium effects of partial and full compatibility between the primary and complementary goods of different firms. In particular, proprietary firms can sell the primary and complementary goods for a positive price, while the price of the open-source primary good is zero. We find that as the degree of compatibility increases, the market share and profits of proprietary firms increase relative to those of open-source firms. This suggests that open source will be more successful when the complementary good is more specific to the primary good, like in the case of support and training services, customizations, platform-specific software, and mobile devices (like MP3 players, PDAs or cell phones).

The model and the results are interesting for a variety of reasons. First, endogenizing the participation decision is crucial for understanding the motivations of commercial firms to participate in open-source projects. Second, to the best of our knowledge, this is the first analysis of direct competition between for-profit open-source and proprietary firms. Third, we show there are forces leading to an asymmetric market structure, even when all firms are ex-ante symmetric. Fourth, we obtain conditions under which open source can overcome free-riding and produce a good of high quality, even without coordination of individual efforts. Finally, the model allows an analysis of welfare and optimal policy.

It is important to remark that even though the model is specially designed to analyze open source, it has wider applicability. In particular, it can be used to analyze industries where firms cooperating in R&D coexist with firms developing technologies on their own (read the literature review for more details on the relation of this paper with the literature of cooperation in R&D).

The main contribution of this paper is to present the first tractable model of competition between profit maximizing open-source and proprietary firms. As such, the model captures the main ingredient shaping the decision to share technologies with rivals or not: the trade-off between appropriability and collaboration. We believe our paper is an important first step in the analysis of the behavior of profit maximizing open-source firms. In section 3.8 we discuss interesting directions for further research.

3.1.1 Open source in detail

There are clear antecedents of open-source in the history of technological change and innovation. Well documented examples are the iron industry in Cleveland, UK (Allen 1983); the Cornish pumping engine (Nuvolari 2004); the silk industry in Lyon (Foray and Perez 2006); the Japanese cotton textile industry (Saxonhouse 1974); the paper industry in Berkshire, US (McGaw 1987); and the case of the Viennese chair (Kyriazidou and Pesendorfer 1999). In all these episodes, inventors shared their improvements with other inventors, which led to a fast technical advance.

One of the characteristics in common with open-source is the presence of complementarities. For example, in the case of the iron industry in Cleveland, entrepreneurs were also owners or had mining rights of the mines in the Cleveland district. Improvements in the efficiency of blast furnaces lead to an increase in the value of the iron ore deposit. In the case of the Cornish engine, technical advances were publicized by mine managers, stimulated to do so by the owners of these mines.

Open source has been used to develop software since the early years of computer science, but gained special relevance in the 1990s, with the success of Linux, Apache and Sendmail, among other programs. Software programmers started to develop software as open source to avoid the restrictions imposed by proprietary firms on the access to the source code.

The participation of individual developers in open source is still very important, but the same is true for commercial firms. In the case of embedded Linux, for example, 73.5% of developers work for commercial firms and contribute 90% of the total investment in code Henkel (2006). Lakhani and Wolf (2005) show that 55% of open-source developers contribute code at work, and these programmers contribute 50% more hours than the rest. Lerner, Pathak, and Tirole (2006) show that around 30% of open-source contributors work for commercial firms (however, they cannot identify non-US commercial contributors). Moreover, they show that commercial firms are associated with larger and more dynamic open-source projects (commercial contributors have four times more sensitivity to the growth of the project).

The coexistence of open source and proprietary in software markets is pervasive, as can be seen in Table 3.1. The server operating system market is a good example. According to IDC (2008), the market shares of server operating systems installed in new computer servers in 2008 were: Microsoft 38%, Unix 32.3%, Linux 13.7%, and other 16.1%. This shows that Linux has a significant market share in the market for server operating systems. However, there are reasons to think that Linux's market share is underestimated by IDC. First, the measurement is a flow, not a stock. Second, the operating system is very often changed by users in the years following the acquisition of a computer server and Linux is considered to run better on old computers. It is also interesting to notice that most Unix systems nowadays are also open source. If we sum the shares for Unix-like systems (Unix plus Linux), we get that open-source operating systems have the largest share in the server operating systems market.

Table 3.1: Coexistence of O and P software.

Software	Open Source	Proprietary
Operating Systems	Linux, OpenSolaris	Windows
Web browsers	Mozilla/Netscape	Internet Explorer
Web servers	Apache	MS Internet Information Server
Mail servers	Sendmail	IBM Lotus Domino MS Exchange Server
Databases	MySQL, PostgreSQL	Oracle 11g, MS SQL Server
Content management	Plone	MS Sharepoint, Vignette
Application servers	JBoss, Zope	IBM WebSphere, MS .net
Blog publishing	WordPress	Windows Live Writer

The decision to become open source is affected by dynamic factors. For example, the decision to open Netscape's source code was in part due to the loss of market share to Internet Explorer. However, it is important to remark that in many opportunities, open-source products were the first to be introduced in the market and then proprietary products appeared. Moreover, open-source and proprietary firms coexist even in newly developed software markets, like application servers, blog publishing applications and content management systems.

We think static models like ours can be used to study the *equilibrium* industry structure in this kind of markets. In particular, there are several factors affecting the decision to become open source which can be explained in the context of a simple static model, like the way in which commercial open-source firms profit from their collaborations, and the exact role of free-riding and duplication of effort in determining equilibrium market shares and cost of innovation.

Commercial firms participate in open-source projects because they sell goods and services complementary to the software. For example, IBM provides support for over 500 software products running on Linux, and has more than 15,000 Linux-related customers worldwide.[1]

The presence of complementarities in open source has been documented in recent

[1]www.ibm.com/linux/ (accessed May 15, 2012).

empirical work. Henkel (2006) presents results from a survey of embedded Linux developers and show that 51.1% of developers work for manufacturers of devices, chips or boards and 22.4% work for specialized software companies. Dahlander (2005) finds that the dominant trend for appropriating the returns of innovation in open source is the sale of a complementary service. Fosfuri, Giarratana, and Luzzi (2008) show that firms with larger stock of hardware patents and hardware trademarks are more likely to participate in open source.

The sale of a complementary service can indeed be profitable. The case of Red Hat is illustrative. According to its financial statements, in fiscal year 2009 Red Hat invested $130 million in R&D, and obtained $652 million in revenues for its subscription and training services.

Many firms develop open-source and proprietary software at the same time. For example, IBM contributes code to Linux, but makes most of its software revenue in the middleware segment, where most of its programs are proprietary. Even Microsoft is becoming increasingly open, with its participation in Cloud computing, for example. Our model can be used to address this issue, by noticing that the complementary good sold by open-source firms may be a complementary proprietary software. Another interesting issue would be to analyze multiproduct software firms, and to determine which software should be open source and which should be kept proprietary. For the purposes of this paper, however, we concentrate in the analysis of a particular software segment, abstracting from the interactions with other segments.

Open-source licenses are the instruments guaranteeing the access of developers to the source code. Some licenses allow further modification of the source code without imposing any restriction on developers. Restrictive open-source licenses, on the other hand, require the disclosure of further improvements to the source code when programs are distributed (programmers are still allowed to keep their innovations private if the program is for personal use). The most popular O license is the General Public License (GPL), which is a restrictive license. The GPL is used by Linux, MySQL, Perl and Java, for example. It is true that some open-source contributors disclose improvements to the source code even when these modifications are for personal use. However, restrictive licenses are the most important means for the success of open-source projects. For example, the survey of embedded Linux developers finds that the main reason why developers disclose their contributions to the code is because they are forced to do so by the GPL (Henkel 2006).

3.1.2 Related literature

The first papers on open source were mainly concerned with explaining why individual developers contribute to open-source projects, apparently for free (read Lerner and Tirole 2005, von Krogh and von Hippel 2006, for excellent surveys). The initial answers were altruism, personal gratification, peer recognition and career concerns. The motivations of commercial open-source firms, on the other hand, have been studied less intensively.

Lerner and Tirole (2001, 2002, 2005) identify directions for further research. Some of the questions related with the present paper are: (i) what are the incentives of for-profit firms to participate in open source, (ii) what development model provides higher quality and welfare, and (iii) what is the influence of the competitive environment in open source. More importantly, these authors remark that direct competition between proprietary and open-source firms has received little attention. For more recent surveys, see Maurer and Scotchmer (2006) and Fershtman and Gandal (2011).

Early papers addressing competition between the two paradigms studied duopoly models of a profit-maximizing proprietary firm and a community of not-for-profit/non-strategic open-source developers, selling at zero price (Mustonen 2003, Bitzer 2004, Gaudeul 2007, Casadesus-Masanell and Ghemawat 2006, Economides and Katsamakas 2006). In these papers, however, open-source firms have no profits, and the choice of development model is exogenous. Introducing profit-seeking open-source firms is important because it allows us to analyze the incentives to invest in R&D and the decision to become open source.

Later papers introduced profit-maximizing open-source firms (Johnson 2002, Henkel 2004, Bessen 2006, Schmidtke 2006, Haruvy, Sethi, and Zhou 2008, Casadesus-Masanell and Llanes 2011), but assumed an exogenous market structure. Other papers study industry dynamics when there is competition between open-source and proprietary firms but assume that open-source projects are formed by altruistic contributors (Athey and Ellison 2014) or that open-source firms are non-strategic (Arora and Bokhari 2007). Likewise, recent papers show that the organizational structure of open-source projects may lead to better development incentives than the organizational structure of proprietary firms (Johnson 2006, Polanski 2007, Niedermayer 2013), but study each model in isolation and do not study direct competition between the two paradigms.

The main contributions of our paper are (i) to analyze an oligopoly model with direct competition between for-profit open-source and proprietary firms, in which (ii) the choice of development model is endogenous, and (iii) the market structure is determined endogenously as a result of firms' decisions. In this sense, the closest papers to ours are Jansen (2009) and von Engelhardt (2010).

Jansen (2009) studies a duopoly model with Cournot competition, in which firms may choose to share their knowledge to signal a low cost position, thereby reducing competition. In contrast with Jansen, we study an oligopoly with n firms, consider Bertrand competition, and focus on the effects of technology sharing on investment incentives. Von Engelhardt (2010) studies a Cournot oligopoly in which firms may invest in open-source software, proprietary software, or both. The focus of von Engelhardt's paper differs from ours, since he is more interested in studying the effects of the type of open-source license on the equilibrium. For most of his analysis, von Engelhardt focuses on studying symmetric equilibria (i.e. all firms are of the same type), but he also presents simulation results which show that under coexistence, proprietary firms are larger than open-source firms. We provide theoretical results which formalize these findings.

Finally, our paper is also related to the literature of cooperation in R&D and research joint ventures. A first strand of papers analyzed the effects of sharing R&D on the incentives to perform such investments (D'Aspremont and Jacquemin 1988, Kamien, Muller, and Zang 1992, Suzumura 1992). In particular, Kamien, Muller, and Zang show that free-riding incentives are so strong that a joint venture where firms share R&D but do not coordinate their R&D levels has a lower total investment than the individual investment of each of these firms when there is no cooperation in R&D. We show that this result can be reversed when firms profit from the sale of complementary private good.

A second strand of papers analyzed the endogenous formation of research coalitions. Bloch (1995) presents a model in which firms decide sequentially whether to join the association or not, and compete in quantities after associations are formed. In equilibrium, two associations are formed. However, firms do not decide their optimal investments in R&D, so this model cannot be used to analyze the free-riding incentives created by association. Poyago-Theotoky (1995) and Yi and Shin (2000) assume that firms set their R&D levels cooperatively after associating. In this case, they show that firms in the joint venture invest more in R&D, and have higher profits than outsiders. We show this result is reversed when firms do not coordinate their R&D levels.

A third strand of papers analyzed the endogenous determination of spillovers among firms conducting R&D. Katsoulacos and Ulph (1998) show that firms selling complementary goods may choose maximal spillovers (i.e. decide to be open source), even when they take their decisions non-cooperatively. However, firms are not competing in the same industry. In our model, firms are direct competitors in the markets for the primary and complementary goods. Amir, Evstigneev, and Wooders (2003) present a duopoly model in which firms set cooperatively their R&D levels and the strength of the spillover. In their model, firms choose maximal spillovers, but this is due to the fact that they make their decisions cooperatively.

As can be seen, the literature of cooperation in R&D is an important precedent for our paper. Nevertheless, to the best of our knowledge, previous papers have not analyzed the case of endogenous formation of a coalition cooperating in R&D, when R&D levels are determined non-cooperatively. In particular, our contribution to this literature is the result that the equilibrium in which some firms decide to cooperate and others do not is characterized by an asymmetric market structure, where firms cooperating in R&D have smaller market shares.

3.2 The model

3.2.1 Technology

There are n firms selling packages composed of a primary good (which is potentially open-source) and a complementary private good. This assumption fits particularly well those cases in which the complementary good is essential (or almost essential) for the primary good, such as the cases of embedded systems in electronic devices (mobile phones, DVD and MP3 players, smartphones, tablet PCs, medical equipment, printers, etc.), server software, enterprise solutions, IT technical support, consulting services, etc.

Firms may improve the quality of their packages by investing in R&D. Let x_i be the investment in R&D of firm i. The cost of investment is $c\,x_i$, which is an endogenously determined fixed cost. The marginal cost of producing packages is zero.

Firms may choose to develop their primary goods under the open-source or proprietary development models, which we denote as O and P, respectively. O firms collaborate in the development of the primary good, and quality is given by $q_i = \alpha\, \ln(\Sigma_{i\in O}\, x_i) + (1-\alpha)\, \ln(x_i)$, where $\Sigma_{i\in O}\, x_i$ is the sum of investments of O firms, and $\alpha \in [0,1]$ represents the degree of public good of the investment in R&D. Proprietary firms perform their investment individually, and quality is given by $q_i = \ln(x_i)$.

Our specification for the quality of O packages implies that O firms benefit more from their own investment than from the investment of other firms in the open-source project. There are several reasons why this may be the case. First, O firms have incentives to contribute code to sections of the program that benefit them more directly than what they benefit other firms. Second, even though firms are compelled to share their improvements to the primary good, there may be a delay until these improvements are made available to other firms. Third, there may be a within-firm synergy through which firms gain valuable knowledge and expertise when they increase their participation in the open-source project, and thus are able to offer a better complementary good.

Finally, α may also indicate the degree of restrictiveness of the open-source license. Restrictive licenses, such as the General Public License, force developers to share their contributions to the code if they distribute the modified program. Permissive licenses, such as the BSD License, on the other hand, allow developers to keep their contributions private. Therefore, as α increases, the open-source license becomes more restrictive and a higher fraction of the source code is shared.

3.2.2 Preferences

There is a continuum of consumers. Each consumer has income y and buys one package. Consumer j's indirect utility from consuming package i is

$$v_{ij} = q_i + y - p_i + \varepsilon_{ij}, \tag{3.2.1}$$

where q_i is the quality of the package of firm i, p_i is its price, and ε_{ij} is an idiosyncratic shock (unobservable by firms) representing the heterogeneity in tastes between consumers. This specification for preferences allows for vertical (q_i) and horizontal (ε_{ij}) product differentiation.

Each consumer *observes* prices and qualities and then chooses the package that yields the highest indirect utility. The total mass of consumers is 1, so aggregate demands are equivalent to market shares. To obtain closed-form solutions for the demands we make the following assumption, which corresponds to the multinomial logit model (McFadden 1974):[2]

Assumption 1. *The idiosyncratic taste shocks ε_{ij} are i.i.d. according to the double exponential distribution:*

$$\Pr(\varepsilon_{ij} \leq z) = \exp\left(-\exp\left(-\nu - \delta\, z\right)\right)$$

where ν is Euler's constant ($\nu \approx 0.5772$) and δ is a positive constant. We assume that $\delta \leq 1$.

Under Assumption 1, the market share (demand) of firm i is

$$s_i = \frac{\exp\left(\delta\left(q_i - p_i\right)\right)}{\sum \exp\left(\delta\left(q_i - p_i\right)\right)}. \tag{3.2.2}$$

The taste shocks have zero mean and variance $\pi^2/(6\,\delta^2)$. As δ increases, consumers become less differentiated and the degree of horizontal differentiation among varieties decreases. The assumption on δ requires certain degree of horizontal differentiation relative to vertical differentiation in the model.

3.2.3 Game and equilibrium concept

The model is a two-stage non-cooperative game. The players are the n firms. In the first stage firms decide their type (O or P), and in the second stage they choose investment and prices (x_i, p_i).

Given investments (quality) and prices, each consumer chooses her optimal package. These decisions are summarized by consumer demands (s_i) and embedded into the firms' payoffs: $\pi_i = s_i\, p_i - c\, x_i$.

The equilibrium concept is subgame perfect equilibrium, and we focus on equilibria where all firms deciding to be of the same type in the first stage play the same

[2] The logit is a common model in discrete choice theory (Ben-Akiva and Lerman 1985), and has been widely used in econometric applications (see Train, McFadden, and Ben-Akiva 1987, and references therein), in marketing (McFadden 1986), and in theoretical work (see, for example, Besanko, Perry, and Spady 1990, Anderson and de Palma 1992, Anderson and Leruth 1993). See Anderson, de Palma, and Thisse (1992) for a detailed presentation of its main properties.

equilibrium strategy in the second stage. We call this *symmetric equilibria conditional on firm type* simply a *symmetric continuation equilibria.*

3.3 Solution of the model

3.3.1 Second stage

Let n_o be the number of firms deciding to be O in the first stage. In the second stage, firms choose p_i and x_i to maximize $\pi_i = s_i\, p_i - c\, x_i$, taking as given the decisions of other firms. Working with the first order conditions and imposing symmetry we obtain the optimal price:

$$p_i = \frac{1}{\delta\,(1 - s_i)}, \tag{3.3.1}$$

and the optimal investment in R&D for O and P firms:

$$x_o = \frac{1}{c}\, s_o \left(1 - \alpha\, \frac{n_o - 1}{n_o(1 - s_o)}\right), \tag{3.3.2}$$

$$x_p = \frac{1}{c}\, s_p. \tag{3.3.3}$$

The term inside the parenthesis of (3.3.2) represents free-riding. If $s_o = s_p$, O firms would invest less than P firms because they can appropriate a smaller fraction of their investment.

Further intuition on the main differences between the two development models can be obtained by computing the individual demand sensitivity to changes in investment. In the case of O firms is

$$\frac{\partial s_i}{\partial x_i} = \delta\, s_i(1 - s_i) \left(\frac{\alpha}{\sum_{i \in O} x_i} \frac{1 - \sum_{i \in O} s_i}{1 - s_i} + \frac{(1 - \alpha)}{x_i}\right),$$

and in the case of P firms is

$$\frac{\partial s_i}{\partial x_i} = \delta\, s_i(1 - s_i)\frac{1}{x_i}.$$

The expression within parenthesis for O firms shows the free-riding problem for O firms and it is affected by α. If $\alpha = 0$, investment is fully appropriable for O firms, there is no free-riding problem and all firms are alike in the model. If $\alpha = 1$, the only incentive to invest for an O firm is the gain in market share to the P firms. In the extreme when all firms are O and $\alpha = 1$, investment is a public good and no firm invest in R&D.

From (3.2.2), we can get the ratio of market shares s_o/s_p. Introducing equations (3.3.1) to (3.3.3), taking logs and rearranging terms we obtain

$$(1-\delta)\ \ln\left(\frac{s_o}{s_p}\right) + \frac{1}{1-s_o} - \frac{1}{1-s_p} = \delta\ \ln\left(1 - \alpha\,\frac{n_o-1}{n_o(1-s_o)}\right) + \alpha\,\delta\ \ln\left(n_o\right). \quad (3.3.4)$$

Equation (3.3.4) shows that the difference in market shares depends on the resolution of the conflict between *free-riding* and *collaboration*. To see this, notice that the left hand side is increasing in s_o and decreasing in s_p, so the difference in market shares will increase if the right hand side does. The first term on the right hand side arises from the difference in individual investments (free-riding). The second term arises because individual investments in open source are multiplied by the number of O firms (collaboration).

The trade-off between free-riding and collaboration is determined by α and n_o. On one hand, as α increases, the degree of public good of the investment in R&D increases, and thus the individual investments of O firms (x_o) decrease. On the other hand, as α increases, the joint investment of O firms ($n_o\, x_o$) has a greater effect on quality. Likewise, as n_o increases, the public good problem becomes more important (there are more firms sharing), but collaboration also becomes more important (there are more firms collaborating). Moreover, α and n_o are complementary. The effects of a higher n_o on free-riding and collaboration become more important when α is higher.

The second-stage equilibrium is completely characterized by (3.3.4) and the condition that the sum of the market shares is equal to 1:

$$n_o\, s_o + (n - n_o)s_p = 1. \tag{3.3.5}$$

Proposition 1. *A second-stage symmetric equilibrium exists and is unique. Given n_o, the equilibrium market shares solve (3.3.4) and (3.3.5).*

All proofs are relegated to the Appendix. In what follows we study the comparative statics of the second-stage equilibrium. In Lemma 1 we present a simple condition to determine which type of firm will have higher market share (quality and price).

Lemma 1. *$s_p > s_o$ if and only if $\alpha > \hat{\alpha}(n_o, n)$, where $\hat{\alpha}(n_o, n)$ is increasing in n_o and n and solves:*

$$\alpha\, \frac{n_o^\alpha}{n_o^\alpha - 1}\, \frac{n_o - 1}{n_o} = \frac{n-1}{n}.$$

The comparison of prices and quality is equivalent to the comparison of market shares: if $s_o > s_p$, then $p_o > p_p$ and $q_o > q_p$, and vice versa. Lemma 1 provides an important result: as n_o or n increase, it is more likely that O firms will have higher market share (quality and price) than P firms.

Lemmas 2 and 3 analyze the effects of changes in α and δ on s_o. The effects on s_p have the opposite sign.

Lemma 2. *There exists $\alpha_s \in (0, \hat{\alpha})$ such that s_o is increasing in α if and only if $\alpha < \alpha_s$.*

Lemma 2 implies that the graph of s_o with respect to α (the degree of public good of the investment) is hump-shaped. When α is close to zero, the investment of O firms is mostly private, and individual investments are close to the investments of P firms. Therefore, the positive effects of collaboration are more important than the negative effects of free-riding. For high values of α, free-riding becomes more important and the difference in individual investments between O and P firms increases. Therefore, for large α, free-riding dominates collaboration.

Lemma 3. *s_o is increasing in δ if and only if $\alpha < \hat{\alpha}(n_o, n)$.*

Lemma 3 shows that the effect of a higher δ depends on the value of α. When δ increases, vertical differentiation gets more important relative to horizontal differentiation. As a consequence, investing in R&D has a greater effect on demand, which benefits firms with higher quality products. If $\alpha < \hat{\alpha}$, O firms have higher quality products, and therefore, their market share increases relative to the market share of P firms. The opposite happens when $\alpha > \hat{\alpha}$.

3.3.2 First stage

In the first stage of the game, firms choose to be O or P, taking as given the decisions of the rest of firms and forecasting their equilibrium payoffs in the second stage. Let $\pi(n_o)$ be the second stage equilibrium payoffs when n_o firms decide to be O. Replacing the second stage equilibrium values of prices and investments for both types of firms we obtain

$$\pi_o(n_o) = \frac{s_o}{1-s_o}\left(\frac{1}{\delta} - (1-s_o) + \alpha\,\frac{n_o-1}{n_o}\right), \tag{3.3.6}$$

$$\pi_p(n_o) = \frac{s_p}{1-s_p}\left(\frac{1}{\delta} - (1-s_p)\right), \tag{3.3.7}$$

where $s_o = s_o(n_o)$ and $s_p = s_p(n_o)$ are the second stage equilibrium market shares. Equilibrium profits are always positive, given that α, δ, s_o and s_p are all between 0 and 1. Comparing equations (3.3.6) and (3.3.7), we can see that collaboration has a direct effect on profits (third term inside the parenthesis of equation 3.3.6): if $s_o = s_p$, the investment of O firms is lower than the investment of P firms ($x_o = x_p/n_o$), and the effect is larger, the larger is α.

A number n_o of firms in the open-source project is an equilibrium if and only if $\pi_o(n_o) \geq \pi_p(n_o-1)$ and $\pi_p(n_o) \geq \pi_o(n_o+1)$. These conditions were called internally stable and externally stable coalition conditions by D'Aspremont, Jacquemin, Gabszewicz, and Weymark (1983). The first inequality says that firms deciding to be O cannot gain by deviating and becoming P. The second inequality is a similar condition on the decision of being P. Using the function $f(n_o) = \pi_o(n_o) - \pi_p(n_o - 1)$, equilibrium conditions can be restated as $f(n_o) \geq 0$ and $f(n_o + 1) \leq 0$.

The equilibrium may be such that both types of firms coexist (interior equilibrium) or all firms choose to be of the same type. $n_o = 0$ is a trivial equilibrium because an industry with $(n-1)$ P firms and one O firm is equivalent to an industry with n P firms. Hence we focus on equilibria with $n_o \geq 1$. For $n_o = 1$ to be an equilibrium we need $f(2) \leq 0$. Likewise, for $n_o = n$ to be an equilibrium we need $f(n) \geq 0$.

Figure 3.1 shows an example of the $f(n_o)$ schedule when $\alpha = 1$, $\delta = 0.7$ and $n = 10$. In this case, there are 6 firms in the open-source project in equilibrium.

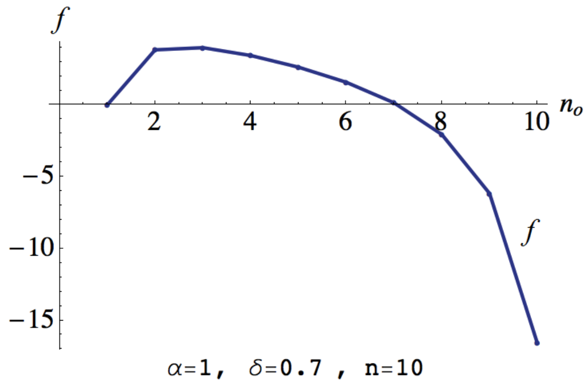


Figure 3.1: Equilibrium number of firms in open source.

When firms choose between O or P, they compare the relative benefits of collaboration and secrecy. There are two elements associated with this trade-off. On the one hand, free-riding and collaboration affect the equilibrium market shares, as analyzed in Section 3.3.1. On the other hand, O firms have a lower investment cost. Being P will be more profitable than being O only if free-riding is sufficiently strong as to overcome the positive effects of collaboration. Proposition 2 characterizes the equilibrium of the game.

Proposition 2. *A symmetric continuation equilibrium for the game exists. Given $n > 3$ and δ, there exist $0 < \bar{\alpha} < 1$, such that:*

(i) If $\alpha > \bar{\alpha}$, in the industry equilibrium both types of firms coexist.

(iii) If $\alpha \leq \bar{\alpha}$, in the industry equilibrium all firms decide to be O.

Proposition 2 shows there are two equilibrium regions. When α is large, the degree of public good of the investment in R&D is high, and free-riding is very important, leading to low individual investments for O firms (see Lemma 2). As a consequence, P firms have an advantage over O firms in terms of market share and price. On the other hand, O firms benefit from lower development costs. Therefore, there is room for both types of firms in the industry equilibrium: some firms choose to be P and have a high investment in R&D to benefit from high market shares and prices, and other firms choose to be O to benefit from low development costs. When α is small, the positive effects of collaboration on investment incentives are stronger than the negative effects of free-riding (see Lemma 2), and the total investment in the open-source project is larger than the investment of a P firm. Therefore, in the industry equilibrium all firms choose to be O. The next lemma characterizes the industry equilibrium when there is coexistence.

Lemma 4. *The industry equilibrium with coexistence is characterized by large P firms producing a high quality product and small O firms producing a low quality product.*

Figure 3.2 shows the equilibrium regions for different values of α and n, when $\delta = 1$. The area corresponding to equilibria with coexistence first increases but then decreases with n, which means that large numbers favor cooperation, even without coordination of individual investments. In the following section, we elaborate on this point.

3.4 Entry

In previous sections, we assumed a fixed number of firms. Implicitly, we were assuming an *exogenous* cost of entry, in addition to the *endogenous* cost of developing the

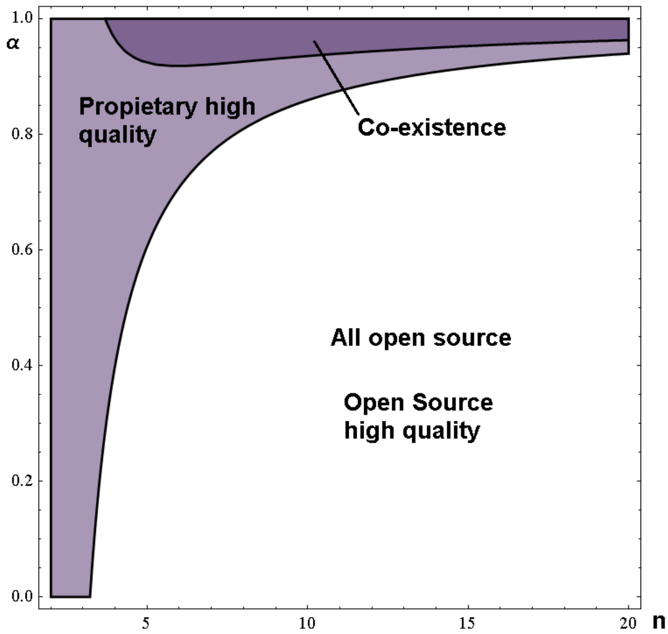


Figure 3.2: Equilibrium regions.

primary good. Concretely, suppose firms decide to enter the industry before deciding to become O or P. As the cost of entry decreases, the number of firms in equilibrium increases and we move along the curve of Figure 3.2.

The fixed cost of entry may be a consequence of entry barriers in the market of complementary goods. The market of smartphones, for example, has high barriers to entry. Even though the Android operating system is free for firms wanting to develop a smartphone, this industry has few producers, because developing the hardware that goes with the software is expensive. In other markets of complementary goods, however, the cost of entry is small and a large number of firms is present. For example, the cost of entering the market of software training, support, and customization is relatively small, and a large number of independent software programmers provide such services. In fact, the open-source movement started precisely among these independent programmers.

For this reason, studying the case of a small cost of entry is of interest. In our model, as the cost of entry goes to zero, the number of firms in the industry goes to infinity and the industry structure becomes monopolistically competitive. As Besanko, Perry, and Spady (1990) show in their analysis of the logit model of monopolistic competition, price converges to a constant margin over marginal cost, which depends on the substitution parameter δ. The key assumption is horizontal differentiation, which allows any firm to enter and supply a differentiated product (Dixit and Stiglitz 1977).[3] Proposition 3 characterizes the equilibrium of the limiting economy.

Proposition 3. *A symmetric continuation equilibrium of the limiting economy ($n \to \infty$) exists and is unique. The equilibrium is characterized as follows:*

(i) If $\alpha < 1$, all firms decide to be O.

(ii) If $\alpha = 1$, both types of firms coexist, the ratio of market shares (s_o/s_p) is $1-\delta$, the aggregate market share of O firms ($n_o\, s_o$) is $1-(1-\delta)^{\frac{1-\delta}{\delta}}$ and the proportion of O firms (n_o/n) is $\frac{1-(1-\delta)^{(1-\delta)/\delta}}{\delta-(1-\delta)^{(1-\delta)/\delta}}$.

Proposition 3 shows the incentives to participate and to invest in an open-source project do not disappear when the number of firms goes to infinity. This result is somewhat surprising because as n_o increases, free-riding intensifies and individual incentives to invest in open source decrease. In the limit, we would expect the incentives to invest in open source to disappear completely. However, as n_o increases, collaboration between O firms also intensifies (individual investments of O firms are multiplied by a larger factor), which compensates for the negative effects of free-riding.

As $n \to \infty$, investment in R&D and market share converge to zero for *both* O and P firms, as firms become infinitesimally small. However, studying whether investments and market shares converge faster to zero for O or P firms (i.e., determining the limits of the ratios of investments and market shares) is interesting because doing so allows us to determine if large numbers favor cooperation in R&D.

As $n_o \to \infty$, the ratio of individual investments (x_o/x_p) converges to $1-\alpha$. When α is close to 1, individual incentives to invest in the open-source good are low. However,

[3]In the logit model, the support for consumer tastes (ε_{ij}) is the real line, which means a new firm can always grab *some* market share with a small investment in R&D, even under Bertrand competition.

as long as $\alpha < 1$, the ratio x_o/x_p is strictly positive and the individual investment of O firms is multiplied by a factor that goes to infinity. Therefore, $\frac{n_o x_o}{x_p} \to \infty$ and $s_o/s_p \to \infty$, which means all firms choose to be O.

When $\alpha = 1$, on the other hand, $n_o \to \infty$, $x_o/x_p \to 0$, and their product converges to a constant. In equilibrium, the ratio of market shares converges to $1 - \delta$ and the ratio of profits converges to 1, so both types of firms coexist. As in Proposition 2, in the equilibrium with coexistence, P firms are larger than O firms.

In the equilibrium with coexistence, the aggregate market share of O firms is decreasing in δ, but the proportion of O firms is *increasing* in δ. As δ increases, two effects occur: competition intensifies, which lowers mark-ups; and vertical differentiation becomes more important, which increases the returns to investment. As a consequence, P firms become larger, the aggregate market share of O firms becomes smaller, and there is room for fewer P firms.

3.5 Open membership and open-source licenses

In this section we analyze the effects of changes in n_o on individual profits for an O firm, and on individual profits for a P firm. Figure 3.3 shows the profit schedules of an O and a P firm for $\alpha = 1$, $\delta = 1$ and $n = 10$. We can see that π_o increases, keeps approximately constant for some values of n_o and finally increases again. Interestingly, at the equilibrium ($n_o = 6$), the profits of an O firm are increasing in n_o, which means that O firms would not find it profitable to limit access to the open-source project. This gives a rationale for open-source licenses, such as the GPL, guaranteeing open membership in open-source projects (as we explained above, the only restriction in the GPL is that whenever modifications to the program are distributed, they have to be made available to the rest of developers in the project).

Simulations show that the profit of an O firm is increasing in n_o at the equilibrium for any α and δ. Interestingly, the profit of an P firm is also increasing when the

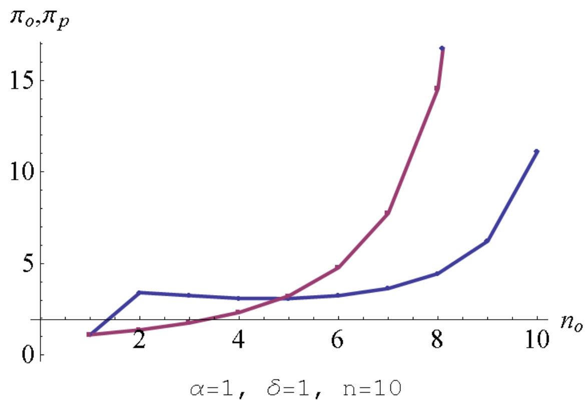


Figure 3.3: Firm profits as functions of n_o.

equilibrium has both kinds of firms. This means that both O and P firms prefer to compete against O firms rather than P firms.

Notice that the result holds even though O firms are direct competitors in the markets for the primary and complementary goods. If the firms were not direct competitors but were benefiting from the development of the open-source good, the result would be even stronger. However, the result also depends on the fact that the number of firms in the industry is fixed. If free-entry into the project would stimulate the entry of new firms in the industry the result could be reversed.

3.6 Welfare analysis

One of the advantages of the logit model is that it can be used to construct a representative consumer whose utility embodies the aggregate behavior of the continuum of users (Anderson, de Palma, and Thisse 1992).

Let s_i be the quantities of each package the representative consumer consumes, and let $\sum s_i = 1$. Total income is y, and z represents consumption of the numeraire. The utility of the representative consumer is

$$U = \sum q_i\, s_i - \frac{1}{\delta} \sum s_i \ln(s_i) + z.$$

The utility function embodies two different effects. The first term represents the direct effect from consumption of the n packages, in the absence of interactions. The second term introduces an entropy-effect, which expresses the representative consumer's preference to consume multiple packages.

The utility function is quasilinear, which implies transferable utility. Thus, social welfare is the sum of consumer utility and firm profits:

$$W = \sum q_i \, s_i - \frac{1}{\delta} \sum s_i \ln(s_i) + y - \sum c \, x_i. \tag{3.6.1}$$

Let us define formally the Social Planner's problem.

Definition 1. *The Social Planner's problem is*

$$\begin{aligned} \max_{\{s_i, x_i, \tau_i\}_{i=1}^{n}} & \sum q_i \, s_i - \frac{1}{\delta} \sum s_i \ln(s_i) + y - \sum c \, x_i \\ \textit{subject to} & \sum s_i = 1 \\ & s_i, x_i \geq 0, \tau_i \in \{O, P\} \quad \forall i \end{aligned}$$

where τ_i denotes whether firm i is O or P.

Given the number of firms n, the Social Planner's problem is to choose the quantity produced (s_i) for each firm, the investment in R&D (x_i) for each firm, and whether a firm should be O or P (τ_i). In the optimal solution, the Social Planner will have all firms sharing their investment in R&D. Also, given the concavity and symmetry of the utility function, the Social Planner will set $s_i = 1/n$ for all i. To determine the optimal investment, the Social Planner maximizes

$$W = \alpha \, \ln(n \, x^*) + (1 - \alpha) \, \ln(x^*) + \frac{1}{\delta} \ln(n) + y - n \, c \, x^*,$$

which leads to an optimal investment equal to $x^* = 1/(c \, n)$.

In the market equilibrium, product quality is suboptimal regardless of the number of O and P firms: O firms are subject to free-riding, which leads to a suboptimal investment in R&D, but P firms do not share their improvements on the primary good, generating an inefficient duplication of effort.

3.6.1 Government policy

Now we turn to an analysis of government policy. We will show the first best can be achieved by using a tax-subsidy scheme provided $\alpha < 1$. The cost of R&D for O is $c_o = (1-\kappa)\, c$, where κ is a proportional subsidy on the investment of O firms. This subsidy, in turn, is financed by proportional or lump-sum taxes paid by consumers.

The ratio of investments of O and P firms becomes

$$\frac{x_o}{x_p} = (1-\kappa)^{-1} \frac{s_o}{s_p} \left(1 - \alpha \frac{n_o - 1}{n_o} \frac{1}{1 - s_o}\right),$$

and the equation characterizing the equilibrium market shares becomes

$$(1-\delta) \ln\left(\frac{s_o}{s_p}\right) + \frac{1}{1-s_o} - \frac{1}{1-s_p} = \delta \ln\left(1 - \alpha \frac{n_o - 1}{n_o(1-s_o)}\right) + \alpha\, \delta \ln(n_o) - \delta \ln(1-\kappa).$$

An increase in the subsidy increases the difference in investments and market shares between O and P, and also decreases the cost of investment for O, so firms are more tempted to become O. Lemma 5 shows that if the subsidy is high enough, O firms will have a higher market share than P in a second-stage equilibrium.

Lemma 5. *$s_o > s_p$ in a second-stage equilibrium if and only if*

$$\kappa > 1 - \left(1 - \alpha \frac{n}{n-1} \frac{n_o - 1}{n_o}\right) n_o^{\alpha}.$$

In particular, if $\kappa > 1 - (1 - \alpha n(n-2)(n-1)^{-2})\,(n-1)^{\alpha}$, then $s_o > s_p$ for $n_o = n - 1$; therefore, all firms want to be O. Proposition 4 shows the optimal policy.

Proposition 4. *If $\alpha < 1$ then the optimal subsidy is $\kappa^* = \alpha$, which attains the first-best levels of investment. In equilibrium, all firms decide to be O.*

The optimal subsidy is computed such that the investment of an O firm (when all firms are O) $x_o = (1-\alpha)/(n\, c_o)$ equals the first best investment $x^* = 1/(c\, n)$.

Proposition 4 has an intuitive interpretation. Provided that an O firm can appropriate a portion of its investment in R&D, a R&D government subsidy can provide enough incentives to achieve the first best solution. Moreover, the optimal subsidy increases in the degree of public good of the investment in R&D.

The subsidy has a double effect: it increases the investment of O firms, and it encourages P firms to become O (to share R&D). The optimal subsidy is increasing in the degree of public good of the investment in R&D. In other words, the subsidy should be higher for projects for which the direct appropriability of the investment in R&D of O firms is not very high.

Finally, note that in our model, lump-sum or proportional taxes are equivalent, because each consumer buys one product; therefore, proportional taxes do not affect the quantities sold. Thus financing the subsidy with proportional taxes does not cause a deadweight loss and the policy maker can achieve the first best.

3.7 Extensions to the baseline model

In this section of the paper we discuss the modeling assumptions of the baseline model, and we present extensions showing the results that are not dependent on the basic assumptions.

3.7.1 Modeling assumptions

In this section we discuss the main assumptions of the model, and the consequences of relaxing them.

Bundling and compatibility. We have assumed that firms sell packages (bundles) composed of one unit of the primary good and one unit of the complementary good. Under this assumption, the two goods become effectively one and each firm sets only one price, which greatly simplifies the model and allows us to focus on the effects of technology sharing on the decision to be O or P.

Implicitly, we are assuming that (i) primary and complementary goods are perfect complements, and (ii) complementary goods designed for one primary good are incompatible with other primary goods. Under these assumptions, each consumer must choose a primary good and a complementary good from the same firm.

Industry examples and economic theory indicate that the incompatibility assumption may be a good description for several markets in which open source is important.

On the theory side, Carmen Matutes (1992) present a duopoly model to study compatibility and bundling decisions. Each firm sells two perfectly complementary components (consumers need one component of each kind), and components may be compatible or incompatible across firms. Matutes and Regibeau show there are equilibria where firms choose to make their components incompatible in order to commit to pure bundling.

There are many examples of incompatible components in the software industry. For example, Red Hat specializes in providing support services for Linux, and this support service has little value for Windows users (likewise, Microsoft's support service has little value for Linux users). Also, many applications that run in Mac OS X cannot run in Windows, (and many applications for the iPhone do not run on other mobile devices).

Clearly, the case of compatible components is also pervasive. For example, MS Office can be used in Macs, and Sun servers can run a variety of operating systems (although they are designed to run better on Sun's OpenSolaris). For this reason, in section 3.7.5 we modify the model to analyze the partial compatibility and full compatibility cases. Firms become multiproduct firms, and set separate prices for the primary and complementary goods. P firms obtain revenues from selling the primary and complementary goods, whereas O firms only obtain revenues from selling the complementary good (the price of the primary good is zero).

For tractability, we assume there is only one P firm competing against several O firms,[4] and we focus on the analysis of the equilibrium prices, investments and market shares as a function of the degree of compatibility.

We find that as goods become more compatible, the market share of the open-source complementary goods falls. Also, the profit of the P firm increases relative to the profit of an O firm. Therefore, open source will tend to perform better when the complementary good is more specific to the primary good, like in the case of support and training services, customizations, platform-specific software, and mobile devices (like MP3 players, PDAs or cell phones).

Investment in the complementary good. In the baseline model, firms cannot invest directly to increase the quality of the complementary good. Instead, the quality of the complementary good increases with individual investments in the primary good. This may happen because there is a synergy effect (the support service of a firm which contributes more code to Linux is likely to be better than the support service of a firm that contributes less), or because firms contribute in areas of the primary good that are more relevant for their own complementary goods (a firm selling financial software will be more interested in developing Linux's mathematical capabilities, and will benefit more from these contributions than the rest).

As we will see, the assumption of a synergy effect is *not essential* for our key result –the existence of an equilibrium with both kinds of firms, in which P firms have a larger market share and quality than O firms– which obtains when the synergy effect is small (even zero). Instead, the synergy effect allows for the possibility that O firms have higher quality and market share than P firms, and for the existence of equilibria with only O firms, which will happen if the synergy effect is large enough.

However, it is interesting to ask what would happen if firms could invest in the complementary good without contributing to the primary good. In section 3.7.4 we extend the baseline model to allow for this possibility. Now, the quality of the comple-

[4] In general, models with multiproduct firms are usually very difficult to solve. They usually consist of a duopoly game, and only symmetric continuation equilibria are analyzed, which limits applicability for the present case (n firms of two different types).

mentary good is a weighted average of the individual investment in the primary good (synergy effect) and the direct investment in the complementary good. We show that as the importance of the direct investment increases relative to the synergy effect, the incentive to be O decreases, and the model converges to the equilibrium with both kinds of firms.

Demand specification. Logit demands follow from our assumption of the double exponential distribution for the idiosyncratic taste term. This distribution is similar to the normal distribution (which would yield the probit model when applied to equation (3.2.1)), but has the advantage of providing an analytically tractable demand system whereas the normal does not.[5]

We have assumed a specific demand structure due to the difficulty of analyzing asymmetric equilibria. While we are confident that our results will continue to hold with other models of product differentiation, the formal extension of our analysis to a general demand system is by no means straightforward. As a first step, it therefore seems reasonable to study a simple case, such as the logit, which provides explicit expressions for the demand functions.[6]

Given that taste shocks are distributed in the real line, every firm will have some consumers with strong preferences for its products, and therefore even firms with very low quality products will end up with some positive demand. It could be argued that this is the reason why O firms may subsist in equilibrium, even when they have lower quality than P firms. However, this argument would be only relevant if O

[5] The logit is a common model in discrete choice theory (see, for example Ben-Akiva and Lerman 1985), and has been widely used in econometric applications (see Train, McFadden, and Ben-Akiva 1987, and references therein), in marketing (McFadden 1986), and in theoretical work (Anderson, de Palma, and Thisse 1992, Besanko, Perry, and Spady 1990, Anderson and de Palma 1992, Anderson and Leruth 1993).

[6] Other alternatives for modeling an oligopoly with vertically and horizontally differentiated goods are the linear demand model and Salop's circular city with exogenous locations. We have studied the linear demands case, and have found that results still hold under this alternative specification. However, the analysis becomes much more complex. An appendix with the analysis of the linear demand case may be obtained from the authors upon request. The circular city model would add additional complications because the equilibrium would depend on whether O and P firms were selling neighboring goods or not.

firms were selling an extremely low quality product, which is never a optimal decision whenever: (i) Inada conditions hold for the quality improving technologies, (ii) firms can appropriate some fraction of their investment.[7]

Finally, the assumption that the taste shocks are i.i.d. across packages implies that the differentiation between O and P firms is symmetric. However, O packages share the same primary good, so they are likely to be more similar than P packages. In section 3.7.2 we introduce a nested logit model to introduce a difference in the substitutability between O and P firms.

We find that as O packages become more similar, the equilibrium number of firms in O decreases. Depending on parameter values, the equilibrium may have only O firms, only P firms, or both kinds of firms. If the difference in substitutability is high enough, all firms will decide to be P in equilibrium. This extension provides an important result: we should expect to see a higher proportion of O firms in industries where firms have more possibilities to differentiate their complementary goods.

3.7.2 Lower differentiation for open-source products

Given that O packages share the same primary good, they are likely to be more similar to each other than P packages. To introduce this difference in the degree of substitution, we use a nested logit model (Ben-Akiva 1973), which adds an element of *endogenous* horizontal differentiation to the trade-off between collaboration and secrecy. By becoming P, firms are able to differentiate their product more than O firms.

The main consequences are that (i) the equilibrium number of firms in O is smaller than in the previous model, (ii) equilibria with only P firms exist, and (iii) parameter values exist that lead to multiple equilibria.

Consumers are heterogeneous in two different dimensions: they have idiosyncratic

[7]This condition holds if there is at least one P firm or $\alpha < 1$. There is zero appropriability only if all firms are O and $\alpha = 1$, but in this case the comparison between O and P quality is irrelevant.

tastes for the primary good and idiosyncratic tastes for the complementary good. The relative strength of these two forces drives the differences in substitution. Following the nested logit representation of Cardell (1997), consumer j's indirect utility from consuming package i, based on primary good k, is

$$v_{ikj} = q_i + y - p_i + \eta_{kj} + (1-\sigma)\,\varepsilon_{ij},$$

where q_i is defined as in section 3.2, η_{kj} is a primary good idiosyncratic component, and $\sigma \in [0,1]$ weighs the different idiosyncratic components. Assumption 2 replaces Assumption 1 for the standard logit case.

Assumption 2. *The idiosyncratic components ε_{ij}, corresponding to complementary good i, are i.i.d. according to the double exponential distribution with scale parameter δ. The idiosyncratic components η_{kj}, corresponding to primary good k, are i.i.d. according to a distribution such that $\eta_{kj} + (1-\sigma)\,\varepsilon_{ij}$ is distributed double exponential with scale parameter δ.*

Assumption 2 implies the horizontal differentiation term $\eta_{kj} + (1-\sigma)\,\varepsilon_{ij}$ has the same distribution as ε_{ij} in the previous model. The variance of η_{kj} is $\sigma(2-\sigma)\pi^2/(6\,\delta^2)$. Cardell shows a unique distribution for η_{kj} exists such that Assumption 2 holds.

Parameter σ determines the relative strength of the horizontal differentiation forces. As σ increases, consumers become more differentiated in their tastes for the primary good and less differentiated in their tastes for the complementary good. When $\sigma = 0$, consumers only have idiosyncratic preferences for the complementary good, and the model becomes the standard logit model of previous sections. When $\sigma = 1$, consumers only have idiosyncratic preferences for the primary good, and O firms sell a homogeneous good.

The proportion of consumers choosing open-source variant i can be decomposed in the following way:

$$s_i = s_{i|o}\,S_o, \tag{3.7.1}$$

where S_o is the aggregate market share of the open-source primary good, and $s_{i|o}$ is the share of variant i within the open-source project.

Under Assumption 2, i's market share within the open-source project depends on its individual contribution to the project:

$$s_{i|o} = \frac{\exp\left(\delta \frac{(1-\alpha)\,\ln x_i - p_i}{(1-\sigma)}\right)}{\sum_{i\in O}\exp\left(\delta \frac{(1-\alpha)\,\ln x_i - p_i}{(1-\sigma)}\right)}.$$

The aggregate market share S_o depends on the average value of the O varieties (the expected value of the maximum of the utilities), V_o:

$$S_o = \frac{\exp\left(\delta\, V_o\right)}{\exp\left(\delta\, V_o\right) + \sum_{i\in P}\exp\left(\delta(q_i - p_i)\right)},$$
$$V_o = \frac{(1-\sigma)}{\delta}\,\ln\left(\sum_{i\in O}\exp\left(\frac{\delta(q_i-p_i)}{(1-\sigma)}\right)\right).$$

P nests are composed only of one P product, so the average value of the nest is the value of its only component. Therefore, the market share of a P firm is simply

$$s_i = \frac{\exp\left(\delta(q_i - p_i)\right)}{\exp\left(\delta\, V_o\right) + \sum_{i\in P}\exp\left(\delta(q_i - p_i)\right)}. \tag{3.7.2}$$

Further intuition on the substitution patterns implied by the assumptions on preferences can be obtained by computing the demand sensitivity to changes in price. The slope of the demand function for a P firm is

$$\frac{\partial s_i}{\partial p_i} = -\delta s_i(1-s_i),$$

and for an O firm, it is

$$\frac{\partial s_i}{\partial p_i} = -\delta s_i\left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} s_{i|o} - s_i\right).$$

These expressions summarize the substitution patterns in the nested logit model: O firms face a more elastic demand than P firms, and the price elasticity for O firms is increasing in σ. If $\sigma = 0$, the slope of the demand is the same for O and P firms, as in the standard logit model. If $\sigma = 1$, the slope of the demand for O firms goes to infinity.

The optimal price and investment of P firms have the same functional forms as before. The optimal price and investment of O firms become:

$$p_o = \frac{1}{\delta\left(1 - s_o + \frac{\sigma}{1-\sigma}\frac{n_o-1}{n_o}\right)}, \tag{3.7.3}$$

$$x_o = \frac{1}{c}\, s_o \left(1 - \frac{\alpha}{1-\sigma}\,\frac{(n_o-1)/n_o}{\left(1 - s_o + \frac{\sigma}{1-\sigma}\frac{n_o-1}{n_o}\right)}\right). \tag{3.7.4}$$

Conditional on market shares, optimal prices for O firms are decreasing in σ, whereas optimal prices for P firms are independent of σ.

From (3.7.2) and (3.7.1), we obtain the ratio of market shares s_o/s_p. Introducing prices and investments, taking logs and rearranging terms, we obtain

$$\begin{aligned}(1-\delta)\ln\left(\frac{s_o}{s_p}\right) + \frac{1}{1 - s_o + \frac{\sigma}{1-\sigma}\frac{n_o-1}{n_o}} - \frac{1}{1-s_p} \\ = \delta\,\ln\left(1 - \frac{\alpha}{1-\sigma}\,\frac{(n_o-1)/n_o}{\left(1 - s_o + \frac{\sigma}{1-\sigma}\frac{n_o-1}{n_o}\right)}\right) + (\alpha\,\delta - \sigma)\,\ln(n_o).\end{aligned} \tag{3.7.5}$$

As in the standard logit case, to guarantee the existence of a symmetric equilibrium we need enough horizontal differentiation relative to vertical differentiation. We assume $\sigma \leq 1 - \delta$, which is a sufficient condition. Proposition 5 summarizes the equilibrium of the second-stage of the game.

Proposition 5. *A second-stage symmetric equilibrium for the nested model exists and is unique. Given n_o, the equilibrium market shares solve (3.7.5) and (3.3.5).*

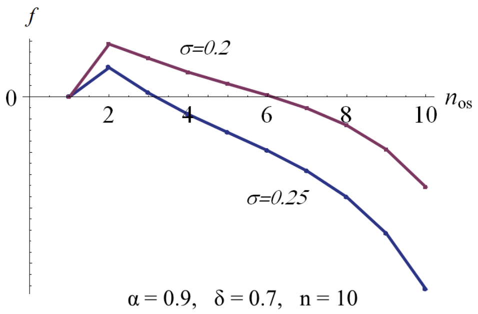
Comparing equations (3.3.4) and (3.7.5), we can see the higher substitution between O varieties has three effects on equilibrium market shares. First, a lower investment of O firms occurs due to the lower return to investment (first term on the right-hand side of 3.7.5). Second, the higher substitution directly and negatively affects the average value of the complementary good (second term on the right-hand side of 3.7.5). Consumers care for variety; therefore, the value of choosing an O package decreases when the complementary good becomes less differentiated. Third, the equilibrium price of O firms is smaller because of the higher substitution (second term on the left-hand side of 3.7.5). The first two effects tend to reduce the market share of O relative to P, and the third effect tends to increase it.

To solve the first stage of the game, we calculate $f(n_o) = \pi_o(n_o) - \pi_p(n_o - 1)$, where $\pi(n_o) = p_i\, s_i - c\, x_i$. Equilibrium conditions are the same as in the standard logit case.

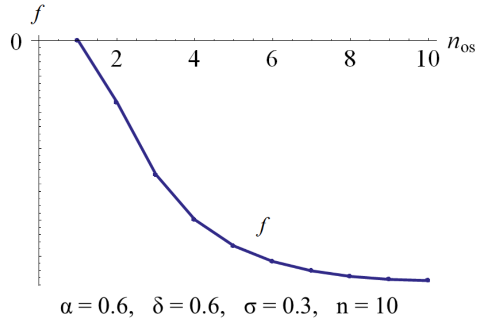
Figure 3.4 shows the graph of $f(n_o)$ for different parameter values. We can make three interesting observations. First, as σ increases for given α and δ (O varieties become more similar), the equilibrium number of firms choosing O decreases (Figure 3.4a). Second, if σ is high enough, equilibria with only P firms exist (Figure 3.4b). Third, for some parameter values, the model exhibits multiple equilibria (in Figure 3.4c an equilibrium exists with $n_o = 2$ and another with $n_o = 10$). In this case, coordination failures may imply a desirable open-source project fails to form.

Figure 3.5 shows the equilibrium regions for different values of α and σ, given $\delta = 0.6$ and $n = 10$ (in case of multiple equilibria, we take the equilibrium with highest n_o). Open source will subsist if the differentiation between O varieties is high enough (σ is low enough). Also, values of α exist such that as σ increases, the equilibrium goes from all O, to coexistence and then to all P. For coexistence, we need a combination of low σ and high α. Finally, our simulations show that whenever O and P coexist, the quality and market share of P firms is larger than that of O firms, which means that the main result of the paper still holds.

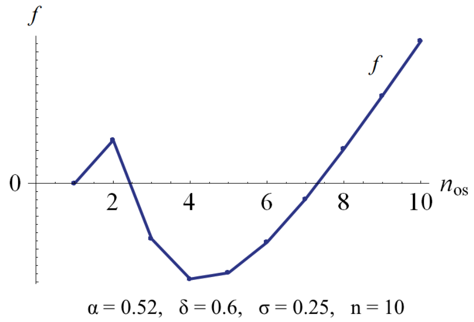
As in the case of the standard logit, studying what happens when the number of firms goes to infinity is interesting. Proposition 6 characterizes the equilibrium of the limiting economy for the nested logit model and formalizes the previous intuition on the effects of σ on the equilibrium.



(a) Decrease in n_o when σ increases.



(b) All P Equilibrium.



(c) Multiple Equilibria.

Figure 3.4: Equilibrium of the nested logit model.

Proposition 6. *A symmetric continuation equilibrium of the limiting economy exists, but the model may exhibit multiple equilibria. A threshold* $0 \leq \tilde{\sigma}(\alpha, \delta) < \alpha\,\delta$ *exists such that*

(i) If $\alpha < 1$ *and* $\sigma < \alpha\,\delta$*, an equilibrium exists in which all firms decide to be O.*

(ii) If $\sigma \geq \tilde{\sigma}(\alpha, \delta)$*, an equilibrium exists in which all firms decide to be P.*

(iii) If $\alpha = 1$ *and* $\sigma < \tilde{\sigma}(1, \delta)$*, an equilibrium exists in which both types of firms coexist.*

As σ increases, O products become more similar. When $\alpha < 1$, the ratio of equilibrium market shares depends on the factor ${n_o}^{\alpha\,\delta-\sigma}$. When $\alpha\,\delta - \sigma > 0$, the effects of collaboration in R&D are stronger than the effects of lower substitution, and the difference in market shares between O and P firms grows large as n_o goes to infinity. As a consequence, an equilibrium exists in which all firms decide to be O. When $\alpha\,\delta - \sigma \leq 0$, collaboration in R&D is not strong enough to compensate for the effects

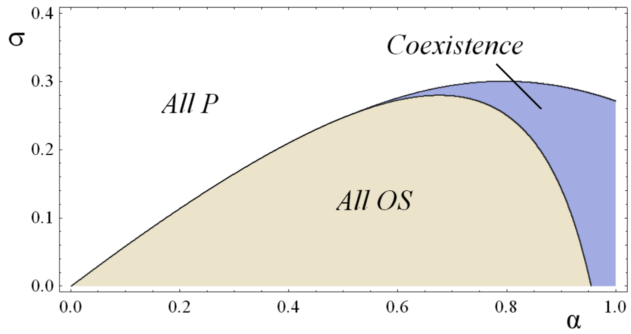


Figure 3.5: Equilibrium regions for the nested logit.

of having a lower substitution; therefore, P firms obtain an advantage in market share. Thus, no equilibrium with O firms exists. For intermediate values of σ, two equilibria exist, one in which all firms choose to be O and another in which all firms choose to be P.

Finally, when $\alpha = 1$, two types of equilibria exist. If $\sigma < \tilde{\sigma}(1, \delta)$, the unique equilibrium has both types of firms. As σ gets closer to 0, the number of O firms in equilibrium increases.[8] If $\sigma \geq \tilde{\sigma}(1, \delta)$, on the other hand, the unique equilibrium has only P firms.

3.7.3 Initial asymmetries

We have assumed firms are ex-ante symmetric, which allowed us to concentrate on ex-post differences arising endogenously in the model. However, it would be interesting to analyze what happens when there are initial asymmetries, which could be due to initial differences in the stock of R&D or installed base.

Specifically, suppose all firms are initially P and have different stocks of R&D. Firms can increase quality by investing in R&D, and have to choose to become O or remain P. Firms deciding to become O will have to share not only their current investment in R&D, but also their initial stocks.

For P firms, initial differences will persist ex-post. Larger firms may invest more or

[8]Unlike the case of $\sigma = 0$, when $\sigma > 0$ and $\alpha = 1$, the number of O firms in equilibrium is finite. See the proof of Proposition 6 for more details.

less than smaller firms, but will finish with a larger stock of R&D. Firms deciding to be O, on the other hand, will tend to be more similar because they have to share their initial R&D stocks. This means that larger firms will have less incentives to become O. In equilibrium, firms deciding to be O will be smaller ex-ante and ex-post. The reason is twofold: larger firms ex-ante have more incentives to remain P, and P firms have more incentives to maintain a larger stock of R&D ex-post.

3.7.4 Direct investment in the complementary good

In the baseline model, the quality of the complementary good is determined by individual investments in the primary good (x_i). This means that firms cannot increase the quality of the complementary good without increasing the quality of the primary good at the same time.

In this section, we extend the baseline model to allow for direct investment in the complementary good. The indirect utility of consumer j from consuming package i is still:

$$v_{ij} = \alpha\, a_i + (1-\alpha)\, b_i + y - p_i + \varepsilon_{ij},$$

where $a_i = \ln(x_i)$ for P firms and $a_o = \ln(\Sigma_{i \in o}\, x_i)$ for O firms, but now, $b_i = \omega \ln(x_i) + (1-\omega)\ln(z_i)$ for all firms, where z_i is the direct investment in the complementary good, and $\omega \in [0,1]$ measures the importance of the synergy effect vis-a-vis the direct investment. The total cost of the investment is $c\,(x_i + z_i)$.

Demands are still given by (3.2.2), and firms maximize $\pi = p_i\, s_i - c\,(x_i + z_i)$. Equilibrium price is:

$$p_i = \frac{1}{\delta(1-s_i)}$$

for all firms, and equilibrium investments in the primary good for O and P firms are:

$$
\begin{aligned}
x_o &= \frac{\alpha + (1-\alpha)\omega}{c}\, s_o \left(1 - \frac{\alpha}{\alpha + (1-\alpha)\,\omega}\, \frac{n_o - 1}{n_o(1 - s_o)}\right), \\
x_p &= \frac{\alpha + (1-\alpha)\,\omega}{c}\, s_p.
\end{aligned}
$$

Finally, equilibrium investment in the complementary good is:

$$
z_i = \frac{(1-\alpha)\,(1-\omega)}{c}\, s_i
$$

for all firms.

An increase in the importance of direct investment with respect to the synergy effect will decrease the optimal investment in the primary good and increase the optimal investment in the complementary good, for given s_i. Substituting prices and investments in the ratio of market shares, taking logs and rearranging terms we get:

$$
\begin{aligned}
&(1-\delta)\,\ln\left(\frac{s_o}{s_p}\right) + \frac{1}{1-s_o} - \frac{1}{1-s_p} \\
&\quad = \delta\,(\omega + (1-\omega)\alpha)\ln\left(1 - \frac{\alpha}{\omega + (1-\omega)\alpha}\, \frac{n_o - 1}{n_o(1-s_o)}\right) + \delta\,\alpha\,\ln(n_o),
\end{aligned} \tag{3.7.6}
$$

where α and δ are defined as in section 3.2.2. Notice that we can get the baseline model by making $\omega = 1$ in equation (3.7.6). As ω decreases, the market share of O firms decreases for given n_o. Thus, the change in the composition of investments as ω decreases has a higher impact on O firms because of the lower appropriability of the investment in the primary good.

To solve the first stage of the game, we calculate $f(n_o) = \pi_o(n_o) - \pi_p(n_o - 1)$, where $\pi(n_o) = p_i\, s_i - c\, x_i$. The equilibrium conditions are the same than in the baseline model.

Figure 3.6 shows the equilibrium for $\alpha = 0.9$, $\delta = 0.7$, $n = 10$, and different values of ω. We can see that as ω decreases, the equilibrium number of firms in O decreases. In the limit, when $\omega = 0$, the equilibrium has both kinds of firms (coexistence). Simulations show that this result holds for any value of δ and α.

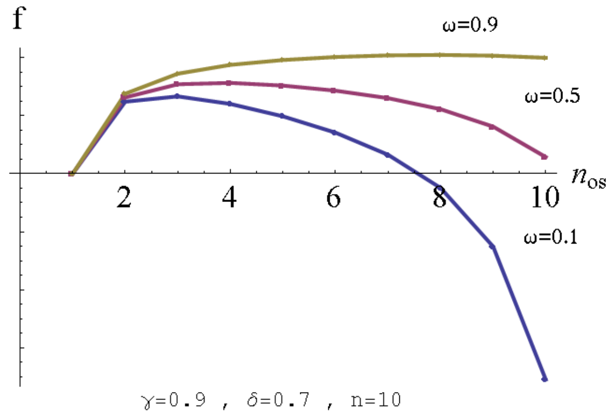


Figure 3.6: Equilibria with direct investment in the complementary good.

The analysis of this section shows the effects of allowing firms to invest solely in the complementary good. Direct investment in the complementary good is fully appropriable. As the importance of this kind of investment relative to the synergy effect increases, the quality differential between O and P firms decreases, reducing the incentives to participate in O. However, the main result of the paper still holds: there are parameter values for which both kinds of firms coexist in equilibrium, and these equilibria are characterized by an asymmetric market structure, with a few large P firms and many small O firms.

3.7.5 Compatibility between open source and proprietary

In previous sections, we assumed that complementary goods could only be used with the primary good for which they were developed. In this section, we analyze what happens when complementary goods can be combined with any primary good, that is, when O and P goods are compatible.

A direct implication is that primary and complementary goods will have separate prices. Let r_k be the price of primary good k, and p_i be the price of complementary good i. P firms can set a positive price for both goods. O firms, on the other hand, are selling an homogeneous primary good, and price competition implies that $r_o = 0$. Therefore, O firms can set a positive price only for the complementary good.

As a result of compatibility, the model becomes highly complex. Therefore, to keep the model tractable, we will assume there is only one P firm, competing against n_o O firms. In other words, we will fix the number of O and P firms and we will focus on studying equilibrium investment and pricing decisions on the second stage of the previous models.

The model is based on the nested logit. There are two nests, one corresponding to users of the O primary good, and the other corresponding to users of the P primary good. Within each nest, consumers can choose between the complementary goods of all firms.

The quality of the O primary good is $a_o = \ln(\Sigma_{i \in o} x_i)$ and the quality of the P primary good is $a_p = \ln(x_i)$, where x_i are individual investments in the primary good.

The quality of the complementary good is determined by a synergy effect as in the baseline model, but this synergy effect is discounted when the complementary good is used with the primary good for which it was not developed. Specifically, the quality of complementary good i developed for primary good k is $b_{ik} = \ln(x_i)$ when used with k, and $b_{ik'} = \ln(\theta\, x_i)$ when used with $k' \neq k$.

$\theta \in [0, 1]$ measures the degree of compatibility between the primary and complementary goods of different firms. When $\theta = 0$, goods of different firms are incompatible, and we obtain the model of previous sections. $0 < \theta < 1$ implies partial compatibility, and $\theta = 1$ implies full compatibility.

Consumer j's indirect utility from consuming good i with primary good k is:

$$v_{kij} = \alpha\, a_k + (1 - \alpha)\, b_{ik} + y - r_k - p_i + \sigma\, \eta_{kj} + (1 - \sigma)\, \varepsilon_{ij}.$$

The distributions of the taste shocks η_{kj} and ε_{ij} are given in Assumption 2.

Let S_o and S_p be the market shares of the O and P primary goods. The market share of complementary good i can be decomposed in the following way:

$$s_i = S_o\, s_{i|o} + S_p\, s_{i|p},$$

where $s_{i|o}$ is the market share of firm i inside the O nest, and $s_{i|p}$ is the market share of firm i inside the P nest.

The market share of complementary good i inside nest k is:

$$s_{i|k} = \frac{\exp\left(\dfrac{\delta((1-\alpha)\,b_{ik} - p_i)}{1-\sigma}\right)}{\sum \exp\left(\dfrac{\delta((1-\alpha)\,b_{ik} - p_i)}{1-\sigma}\right)}.$$

As in section 3.7.2, the market share inside a nest depends only on the relative quality and price of the different complementary goods. This is true for O and P firms. The only difference is that now all firms sell complementary goods in both nests.

The market share of the O primary good is:

$$S_o = \frac{\exp(\delta V_o)}{\exp(\delta V_o) + \exp(\delta V_p)},$$

and $S_p = 1 - S_o$, where V_o and V_p are the average values of the complementary goods within each nest:

$$\begin{aligned} V_o &= \alpha\, a_o + \frac{1}{\delta}(1-\sigma)\,\ln\left(\sum \exp\left(\frac{\delta((1-\alpha)\,b_{io} - p_i)}{1-\sigma}\right)\right), \\ V_p &= \alpha\, a_p - r + \frac{1}{\delta}\,(1-\sigma)\,\ln\left(\sum \exp\left(\frac{\delta((1-\alpha)\,b_{ip} - p_i)}{1-\sigma}\right)\right). \end{aligned}$$

The profit of O firms is $\pi_o = s_i\, p_i - c\, x_i$, whereas the profit of the P firm is $\pi_p = s_p\, p_p + S_p\, r_p - c\, x_p$. Notice that when $\theta = 0$, $S_p = s_p$ and the P firm chooses a single price equal to $p_p + r_p$, as in previous sections. As in section 3.7.2, we assume $\sigma < 1 - \delta$, which guarantees the quasiconcavity of the maximization problem of O and P firms.

The optimal price and investment for O firms are:

$$\begin{aligned} p_o &= \frac{\frac{1}{\delta}\,(1-\sigma)}{1 - s_o - \sigma\, S_o S_p\, \frac{(s_{i|o} - s_{i|p})^2}{s_o}} \\ x_o &= \frac{\alpha}{c}\, S_o\, S_p\, (s_{i|o} - s_{i|p})\, \frac{\delta p_o}{n_o} + \frac{(1-\alpha)}{c}\, s_o. \end{aligned}$$

The optimal price for the primary and complementary good of the P firm, and its optimal investment are:

$$p_p = \frac{(1-\sigma)\, s_{i|o}}{\delta S_o\, s_{i|o}(1-s_{i|o}) + S_p\, s_{i|p}(1-s_{i|p})}$$
$$r_p = \frac{1}{\delta(1-S_p)} - (s_{i|p} - s_{i|o})\, p_p$$
$$x_p = \frac{\alpha}{c} S_p + \frac{(1-\alpha)}{c} s_p;$$

Replacing optimal prices and investments into the market share equations, and noticing that $S_p = 1 - S_o$, $s_{p|o} = 1 - n_o\, s_{o|o}$, and $s_{p|p} = 1 - n_o\, s_{o|p}$, we can get a system of three equations with three unknowns characterizing the equilibrium market shares. These equilibrium expressions are difficult to analyze because of their analytical complexity. Therefore, it is useful to begin by analyzing the full compatibility case ($\theta = 1$), which provides a tractable set of equilibrium conditions.

Full compatibility

When $\theta = 1$, $s_i = s_{i|o} = s_{i|p}$ and equilibrium prices and investments become:

$$\begin{aligned}
p_o &= \frac{(1-\sigma)}{\delta(1-s_o)}, \\
p_p &= \frac{(1-\sigma)}{\delta(1-s_p)}, \\
r_p &= \frac{1}{\delta(1-S_p)}, \\
x_o &= \frac{1-\alpha}{c}\, s_o, \\
x_p &= \frac{\alpha}{c}\, S_p + \frac{1-\alpha}{c}\, s_p.
\end{aligned}$$

The ratios of market shares are:

$$\begin{aligned}
\frac{s_o}{s_p} &= \exp\left(\frac{\delta(1-\alpha)}{(1-\sigma)} \ln\left(\frac{x_o}{x_p}\right) - \frac{\delta(p_o - p_p)}{(1-\sigma)}\right), \\
\frac{S_o}{S_p} &= \exp\left(\delta\alpha \ln\left(\frac{n_o\, x_o}{x_p}\right) + \delta r\right).
\end{aligned}$$

Operating, we get:

$$\ln\left(\frac{s_o}{s_p}\right) = -\frac{\delta(1-\alpha)}{(1-\sigma)}\ln\left(\frac{\alpha}{1-\alpha}\frac{S_p}{s_o}+\frac{s_p}{s_o}\right) - \frac{1}{1-s_o} + \frac{1}{1-s_p}, \tag{3.7.7}$$

$$\ln\left(\frac{S_o}{S_p}\right) = \delta\alpha\,\ln(n_o) - \delta\alpha\,\ln\left(\frac{\alpha}{1-\alpha}\frac{S_p}{s_o}+\frac{s_p}{s_o}\right) + \frac{1}{1-S_p}. \tag{3.7.8}$$

The equilibrium is characterized by the previous equations and the condition that the sum of market shares is equal to 1:

$$S_o + S_p = 1, \tag{3.7.9}$$

$$n_o\, s_o + s_p = 1. \tag{3.7.10}$$

Proposition 7. *A symmetric continuation equilibrium for the full compatibility case exists and is unique. Equilibrium market shares solve (3.7.7) to (3.7.10). In equilibrium, $s_p \geq s_o$ and $S_o \geq S_p$, and the profit of the P firm is always higher than the profit of an O firm.*

It is important to remark that the result that $S_o > S_p$ for all parameter values depends on the fact that there is only one P firm, and may not longer hold if we introduce more P firms in the market.

Lemma 7 provides simple results which can be compared with those of the nested logit model of section 3.7.2, where it is assumed that $\theta = 0$. When $\theta = 0$, O firms could have a higher market share and profits than P firms, depending on α, δ and σ. When $\theta = 1$, on the other hand, the complementary goods market share and profit of the P firms will always be higher than the market share of O firms.

These results suggest that O will perform better when the complementary good is more specific to the primary good. Good examples of these kinds of complementary goods are support and training services, customizations, platform-specific software, and mobile devices (like MP3 players, PDAs or cell phones), among others.

Partial compatibility

We now turn to the analysis of the effects of changes in θ on the equilibrium. As θ increases, the degree of specificity of the complementary goods decreases. This means that O firms tend to have similar market shares in the O and P primary goods markets, and therefore have less incentives to invest in the primary good. The P firm, on the other hand, keeps its incentives to invest in the primary good because it sells this good for a positive price. Therefore, as θ increases, the investment of the P firm increases relative to the investment of the O firms.

The above argument implies that as θ increases, we should see an increase in the market shares of the primary and complementary goods of the P firm, and also an increase in the profits of the P firm relative to the profit of the O firms.

Figure 3.7 shows the effects of changes in θ on the natural logarithm of the ratios s_o/s_p, S_o/S_p and π_o/π_p, for $n_o = 10$, $\alpha = 0.3$, $\delta = 0.8$, $\sigma = 0.1$ and $c = 1$. We can see that these three ratios decrease with θ, which confirms our previous assertions. Our simulations for different values of the parameters indicate that $S_o/S_p > 1$ for all θ.

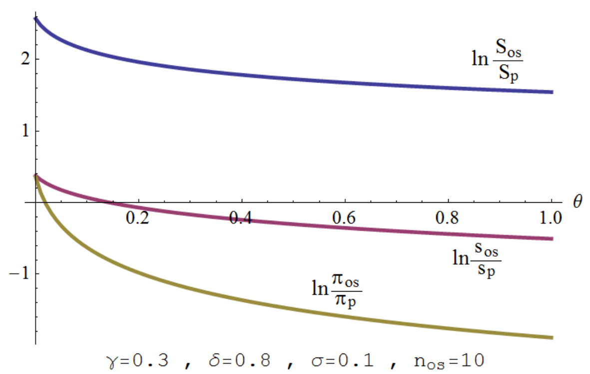


Figure 3.7: Effects of changes in θ on equilibrium.

In Figure 3.7, O firms have higher market shares in the complementary goods and higher profits than the P firm for small θ, but have lower market shares and profits for high θ. More generally, we know that s_o/s_p and π_o/π_p may be larger or smaller than 1 when $\theta = 0$, which depends on α, δ and σ. However, what is more important is that these two ratios are decreasing in θ and that they will always be smaller than 1 when

$\theta = 1$, by Lemma 7.

Finally, it is important to remark that introducing compatibility strengthens our results of an asymmetric market structure, where P firms have a higher market share than O firms (in the complementary goods market).

3.8 Conclusions

This paper investigates the motivations of commercial firms to participate in open source, and the implications of direct competition between open-source and proprietary firms on R&D investments and equilibrium market shares. We present a model in which firms decide whether to become open source or proprietary, and their investment in R&D and price. Both types of firms sell packages composed by a primary good (e.g., software) and a complementary private good (e.g., support and training services or hardware). The difference between both types of firms is that open-source firms share their investments in R&D, whereas proprietary firms develop their products on their own.

Our main contribution is to determine conditions under which open-source and proprietary firms coexist in equilibrium. An asymmetric market structure characterizes such equilibria: proprietary firms invest more in R&D and obtain a larger market share than open-source firms. Open-source firms, on the other hand, benefit from lower development costs. This result is robust to the introduction of a lower differentiation among open-source varieties. We also study a limiting economy, and show conditions under which large numbers favor cooperation in R&D.

Our model points to several important characteristics of open source. In particular, the success of open source will depend on (i) the strength of the complementarity between primary and complementary goods, (ii) the possibility to differentiate the firm's open-source variant from other open-source and proprietary products, and (iii) the degree of appropriability of investments in R&D.

The welfare analysis shows the equilibrium with coexistence is suboptimal for two reasons: too little collaboration (caused by proprietary firms) and too little investment in R&D (caused by open-source firms). We show a subsidy to open-source development can improve welfare not only because it increases the investment in R&D, but also because it encourages commercial firms to participate in open source, thereby enhancing collaboration. This finding explains the active involvement of governments in promoting open source.

Our objective was to present a tractable model analyzing the coexistence of open-source and proprietary firms. We believe our paper is an important first step in the analysis of the behavior of profit-maximizing firms in open source. Several directions for further research are possible. First, the model could be modified to study bundling and compatibility decisions. Second, consumer preferences could be modified to introduce network effects. Third, an important technological difference between open-source and proprietary firms is that the former benefit more from user innovation than the latter. In open source, users can access the source code, which allows them to customize the software program to their needs and to correct bugs at a faster rate.

Appendix: proofs of theorems in text

Proof of Proposition 1. The first order conditions with respect to p_i and x_i are

$$\frac{\partial \pi_i}{\partial p_i} = \frac{\partial s_i}{\partial p_i} p_i + s_i \leq 0 \qquad \text{with equality if } p_i > 0, \tag{3.8.1}$$

$$\frac{\partial \pi_i}{\partial x_i} = \frac{\partial s_i}{\partial x_i} p_i - c \leq 0 \qquad \text{with equality if } x_i > 0. \tag{3.8.2}$$

For the moment, assume that $p_i > 0$ and $x_i > 0$ in equilibrium, so the first order conditions hold with equality. Later, we will show there are no corner equilibria. Working with equation (3.8.1) we obtain the optimal price:

$$p_i = \frac{1}{\delta(1 - s_i)}. \tag{3.8.3}$$

Equation (3.8.3) holds for both types of firms (O and P). To find the optimal investment in R&D

we need to calculate $\partial s_i/\partial x_i$, which in the case of O firms is

$$\frac{\partial s_i}{\partial x_i} = \delta\, s_i(1-s_i)\left(\frac{\alpha}{\sum_{i\in O} x_i}\frac{1-\sum_{i\in O} s_i}{1-s_i} + \frac{(1-\alpha)}{x_i}\right),$$

and in the case of P firms is

$$\frac{\partial s_i}{\partial x_i} = \delta\, s_i(1-s_i)\frac{1}{x_i}.$$

Imposing symmetry and introducing these expressions into (3.8.2) we obtain

$$x_o = \frac{1}{c}\, s_o \left(1-\alpha\,\frac{n_o-1}{n_o}\,\frac{1}{1-s_o}\right), \tag{3.8.4}$$

$$x_p = \frac{1}{c}\, s_p, \tag{3.8.5}$$

and the ratio of optimal investments in equilibrium,

$$\frac{x_o}{x_p} = \frac{s_o}{s_p}\left(1-\alpha\,\frac{n_o-1}{n_o}\,\frac{1}{1-s_o}\right). \tag{3.8.6}$$

From (3.2.2) we obtain the ratio of market shares between O and P firms:

$$\frac{s_o}{s_p} = \exp\left(\delta(q_o - q_p + p_p - p_o)\right), \tag{3.8.7}$$

$$\ln\left(\frac{s_o}{s_p}\right) = \delta\,(q_o - q_p) + \frac{1}{1-s_p} - \frac{1}{1-s_o}, \tag{3.8.8}$$

and from the definition of q_o and q_p, we obtain

$$q_o - q_p = \ln\left(\frac{x_o}{x_p}\right) + \alpha\,\ln(n_o). \tag{3.8.9}$$

From equations (3.8.6), (3.8.8) and (3.8.9), we obtain equation (3.3.4), which is an implicit equation determining the relation of market shares between O and P firms in equilibrium. This equation, together with the equation establishing that the sum of the market shares is equal to 1, completely characterizes the equilibrium.

Now we will show there are no corner solutions ($x_i > 0$ and $p_i > 0$ in the symmetric continuation equilibrium). If $p_i = 0$, then profits are zero and the firm would find profitable to increase p_i. To analyze $x_i = 0$ we have to specify what happens with s_i when $x_i = 0$. Assume that if $x_i = 0$ and $x_j > 0$ for at least one $j \neq i$, then $s_i = 0$. When $x_i = 0$ for all i, on the other hand, $s_i = \frac{\exp(-\delta\, p_i)}{\sum \exp(-\delta\, p_i)}$.

There are 3 cases: $x_p = 0$ and $x_o > 0$, $x_p > 0$ and $x_o = 0$, and $x_p = 0$ and $x_o = 0$. If $x_p = 0$ and $x_o > 0$, then $s_p = 0$ and a P firm makes zero profits. But a P firm can deviate to $p_i = \frac{1}{\delta\,(1-s_i)}$ and $x_i = \frac{s_i}{c}$ with $s_i > 0$. Such a deviation is profitable if $s_i > 1 - \frac{1}{\delta}$ which always holds. If $x_p > 0$ and $x_o = 0$, then $s_o = 0$ and an O firm makes zero profits. But an O firm can deviate to $p_i = \frac{1}{\delta(1-s_i)}$ and $x_i = \frac{s_i}{c}$ obtaining positive profits. If $x_p = 0$ and $x_o = 0$, then $s_p = s_o = \frac{1}{n}$. An O or a P firm can deviate to $x_i = \epsilon > 0$ obtaining a discontinuous jump in revenue ($s_i = 1$) and a small increase in costs.

Finally, to show existence and uniqueness, we need to prove two things: (1) there is only one fixed point of the system of equations in Proposition 1 (there is only one symmetric continuation equilibrium), and (2) the profit function is concave at the equilibrium (the second order conditions for optimality hold).

Let us first show there is only one fixed point in term of equilibrium market shares. Define the function $g(s_o)$ by plugging equation (3.3.5) in equation (3.3.4):

$$\begin{aligned} g(s_o) = (1-\delta)\ln\left(\frac{(n-n_o)s_o}{1-n_o s_o}\right) - \delta\,\ln\left(1-\alpha\frac{n_o-1}{(1-s_o)n_o}\right) + \\ -\,\alpha\,\delta\,\ln(n_o) - \frac{n-n_o}{n-1-n_o(1-s_o)} + \frac{1}{1-s_o}. \end{aligned} \tag{3.8.10}$$

By construction, s_o solves equations (3.3.4) and (3.3.5) if and only if $g(s_o) = 0$. Existence of at least one s_o such that $g(s_o) = 0$ follows from continuity of g and the fact that $\lim_{s_o \to 0} g(s_o) = -\infty$ and $\lim_{s_o \to \frac{1}{n_o}} g(s_o) = +\infty$. To show that there exists only one such s_o, it is sufficient to show that g is strictly increasing. Computing this derivative, we obtain

$$\begin{aligned} \frac{\partial g}{\partial s_o} = & \frac{1-\delta}{s_o(1-n_o\,s_o)} + \frac{\alpha\,\delta\,(n_o-1)/(1-s_o)}{(1-\alpha)(n_o-1)+1-s_o\,n_o} + \\ & + \frac{(n-n_o)\,n_o}{(1+n_o(1-s_o)-n)^2} + \frac{1}{(1-s_o)^2}. \end{aligned}$$

All terms are positive because $s_o n_o \leq 1$. It follows that there exists a unique (s_o, s_p) solving the system of equations.

To prove that the profit function is concave at the equilibrium candidate, we will show that the Hessian of the profit function (at the equilibrium price and market share) is negative definite. A necessary and sufficient condition for negative definiteness is that the leading principal minors alternate sign. To simplify the exposition here we show the determinants of the Hessian of both firms:

$$|H_p| = \frac{\delta\, s_p^2}{x_p^2}\left(1 - \delta\,(1 - s_p)^2\right),$$
$$|H_o| \geq \delta \frac{s_o^2}{x_o^2}\left(\left(\frac{1 - n_o\, s_o}{(1 - s_o)n_o^2}\alpha + (1 - \alpha)\right) - \delta\,(1 - s_o)^2\left(\frac{1 - n_o\, s_o}{(1 - s_o)n_o}\alpha + (1 - \alpha)\right)^2\right).$$

A sufficient condition for both determinants to be positive is $\delta \leq 1$, which has been assumed throughout the paper. Thus, the concavity of the profit function at the equilibrium is guaranteed for both types of firms. ■

Proof of Lemma 1. To prove the first part of the lemma we only have to check the sign of $g\left(\frac{1}{n}\right)$, where g is defined in (3.8.10). If $g\left(\frac{1}{n}\right) < 0$, then $s_o > 1/n$ and therefore $s_o > s_p$. Then,

$$g\left(\frac{1}{n}\right) = -\delta\left(\ln\left(1 - \alpha\frac{n}{n-1}\frac{n_o - 1}{n_o}\right) + \alpha \ln(n_o)\right),$$

and $g\left(\frac{1}{n}\right) < 0$ if and only if

$$\alpha\,\frac{n_o^\alpha}{n_o^\alpha - 1}\,\frac{n_o - 1}{n_o} > \frac{n-1}{n}.$$

Let $h(\alpha, n_o) = \alpha\,\frac{n_o^\alpha}{n_o^\alpha - 1}\,\frac{n_o - 1}{n_o}$. Given that $h(\alpha, n_o)$ is increasing in α, and $h(0, n_o) < (n-1)/n < h(1, n_o)$ there exists only one $\hat{\alpha}$ such that $h(\hat{\alpha}, n_o) = (n-1)/n$. Moreover, for $\alpha > \hat{\alpha}$ we have $h(\alpha, n_o) > (n-1)/n$ and $s_p > s_o$, and viceversa.

Finally, the proof that $\hat{\alpha}(n_o, n)$ is increasing in n and n_o follows from applying the implicit function theorem to $F(\alpha, n_o, n) = h(\alpha, n_o) - (n-1)/n$, and by observing that $\partial h/\partial n_o < 0$. ■

Proof of Lemma 2. By the implicit function theorem, $\frac{\partial s_o}{\partial \alpha} = -\frac{\partial g/\partial \alpha}{\partial g/\partial s_o}$. We know that $\partial g/\partial s_o > 0$. Let us now compute $\partial g/\partial \alpha$:

$$\frac{\partial g}{\partial \alpha} = \ln(n_o) - \frac{n_o - 1}{\alpha + (1 - s_o - \alpha)n_o}$$

Therefore, $\partial s_o/\partial \alpha = 0$ when $\partial g/\partial \alpha = 0$. Solving for the value $\hat{s}_o$ that makes $\partial g/\partial \alpha = 0$ we obtain

$$\hat{s}_o = \frac{\ln(n_o)(n_o(1 - \alpha) + \alpha) + 1 - n_o}{n_o \ln(n_o)}$$

Substituting $\hat{s}_o$ in $g = 0$, we obtain an equation determining the value α_s that makes the derivative equal to zero. To prove that to the right of α_s the graph of $s_o(\alpha)$ is decreasing, assume this is not the case, so $\partial g/\partial \alpha > 0$. Then, for $\alpha > \alpha_s$ it has to be the case that $s_o > \hat{s}_o$, but this implies that

$\partial g/\partial\alpha < 0$, which is a contradiction. This means that $\partial s_o/\partial\alpha < 0$ for $\alpha > \alpha_s$. A similar reasoning implies that $\partial s_o/\partial\alpha > 0$ for $\alpha < \alpha_s$. ■

Proof of Lemma 3. By the implicit function theorem, $\frac{\partial s_o}{\partial\delta} = -\frac{\partial g/\partial\delta}{\partial g/\partial s_o}$. In the proof of Proposition 1 we showed that $\partial g/\partial s_o > 0$. Next, we will determine the sign of $\partial g/\partial\delta$. Computing this derivative, we obtain

$$\begin{aligned}\frac{\partial g}{\partial\delta} &= -\alpha\ln n_o - \ln\left(1-\alpha\frac{n_o-1}{n_o(1-s_o)}\right) - \ln\left(\frac{(n-n_o)s_o}{1-n_os_o}\right)\\ &= -\frac{1}{\delta}\left[(1-\delta)\ln\left(\frac{s_o}{s_p}\right)+\frac{1}{1-s_o}-\frac{1}{1-s_p}\right]-\ln\left(\frac{(n-n_o)s_o}{1-n_os_o}\right).\end{aligned} \tag{3.8.11}$$

In the second row we use the expression for $g(s_o) = 0$. If $\alpha > \hat{\alpha}$ then $s_o < s_p$ (by Lemma 1), $\partial g/\partial\delta > 0$ (by equation (3.8.11)), and $\partial s_o/\partial\delta < 0$. Conversely, if $\alpha < \hat{\alpha}$ then $s_o > s_p$, $\partial g/\partial\delta < 0$, and $\partial s_o/\partial\delta > 0$. ■

Proof of Proposition 2. We begin by showing existence. For $n_o = 1$ to be an equilibrium we only need $f(2) \le 0$. Likewise, for $n_o = n$ to be an equilibrium we only need $f(n) \ge 0$. In order to have an equilibrium with both types of firms ($1 < n_o < n$), we need that $f(n_o) \ge 0$ and $f(n_o{+}1) \le 0$ at the equilibrium n_o. Suppose there is no equilibrium with $n_o = 1$ or $n_o = n$. Then, $f(2) > 0$ and $f(n) < 0$ so $f(n_o)$ goes from positive to negative at least once when going from $n_o = 1$ to $n_o = n$. Therefore, existence of an equilibrium is guaranteed.

Next, we show that $f(2) > 0$ for any n, α and δ, which means that the equilibrium always has at least two O firms. From the definition of $f(n_o)$, we obtain

$$\delta\, f(2) = \frac{s_o(2)}{1-s_o(2)}\left(1-\delta\,(1-s_o(2))+\frac{\alpha\delta}{2}\right)-\frac{1}{n-1}\left(1-\delta\,\frac{n-1}{n}\right).$$

Let $\bar{s}_o$ be the maximum value of $s_o(2)$ for which $f(2) \le 0$. From the above equation, we obtain

$$\bar{s}_o = \frac{\alpha\, n-2}{4\,(n-1)\,n}-\frac{n\left(1-\delta\left(1-\frac{\alpha}{2}\right)\right)}{2\,\delta\,(n-1)}+\sqrt{\frac{\delta+(1-\delta)\,n}{\delta\,(n-1)\,n}+\left(\frac{n}{2\,\delta\,(n-1)}-\frac{\left(1+\left(1-\frac{\alpha}{2}\right)n\right)}{2\,n}\right)^2}.$$

Let $w = g(\bar{s}_o)$, where $g(s_o)$ is defined in (3.8.10). $w < 0$ implies that $s_o(2) > \bar{s}_o$, which means that $f(2) > 0$. w is strictly increasing in n, and has the following upper bound, as $n \to \infty$:

$$\bar{w} = -\alpha\,\delta\,\ln(2)+(1-\delta)\,\ln\left(\frac{1-\delta}{1-\delta\left(1-\frac{\alpha}{2}\right)}\right)-\delta\ln\left(1-\frac{\alpha}{2}\right).$$

$\bar{w}$ is strictly convex in δ and α, which means that the maximum is at $\delta = 0$ or $\delta = 1$, and $\alpha = 0$ or $\alpha = 1$. It is straightforward to show that $\bar{w}$ goes to zero as δ or α go to zero, and also when both δ and α go to 1. Given that w is strictly increasing in n, this means that $w(n, \delta, \alpha)$ is negative for any finite n. Therefore, $f(2) > 0$.

Lemma A1 will prove very important in characterizing the subgame perfect equilibrium of the game. In order for an O firm to find it profitable to become P ($f(n_o) < 0$), it has to be the case that the increase in market share from becoming P is large enough to compensate for the increase in cost. If $\alpha < \hat{\alpha}(n_o - 1, n)$, then O firms have a larger market share so it is not profitable for them to deviate ($f(n_o) > 0$). Corollaries A1 and A2 are two important implications of this lemma.

Lemma A1 (Sufficient condition for positive f)**.** *If $\alpha < \hat{\alpha}(n_o - 1, n)$ then $f(n_o) > 0$.*

Proof. Rearranging $f(n_o)$ and multiplying by δ we obtain

$$\delta\, f(n_o) = \frac{s_o}{1 - s_o}(1 - \delta(1 - s_o)) - \frac{\tilde{s}_p}{1 - \tilde{s}_p}(1 - \delta(1 - \tilde{s}_p)) + \alpha\delta \frac{s_o}{1 - s_o} \frac{n_o - 1}{n_o}$$

where $s_o = s_o(n_o)$ and $\tilde{s}_p = s_p(n_o - 1)$. The sign of f depends on the sign of the right-hand side of the equation. The first two terms have the same functional form and are increasing in s. The last term is always positive. Therefore, if $s_o(n_o) \geq s_p(n_o - 1)$, then $f(n_o) > 0$. A sufficient condition is that $s_o(n_o - 1) \geq 1/n$ and $s_o(n_o) \geq 1/n$, which is equivalent to $\alpha < \hat{\alpha}(n_o - 1, n)$ and $\alpha < \hat{\alpha}(n_o, n)$. However, $\hat{\alpha}(n_o, n)$ is decreasing in n_o, so $\alpha < \hat{\alpha}(n_o - 1, n)$ implies $f(n_o) > 0$. ■

Corollary A1 (Necessary condition for an interior equilibrium)**.** *At an interior equilibrium n_o it is necessary that $\alpha \geq \hat{\alpha}(n_o, n)$.*

Proof. For an interior equilibrium at n_o we need that $f(n_o) \geq 0$ and $f(n_o + 1) \leq 0$, but Lemma A1 implies that for $f(n_o + 1) \leq 0$ we need $\alpha \geq \hat{\alpha}(n_o - 1, n)$. ■

Corollary A2 (Sufficient condition for an equilibrium with $n_o = n$)**.** *If $\alpha \leq \hat{\alpha}(n - 1, n)$ then all firms decide to be O in equilibrium.*

Proof. If $\alpha \leq \hat{\alpha}(n - 1, n)$ then $f(n) \geq 0$, so if $n_o = n$ then no firm would gain by becoming a P firm. ■

Corollary A1 states that in any interior equilibrium it has to be the case that the P firms have a larger market share than O firms, and therefore a higher quality product. Corollary A2, on the other hand, shows that if the degree of public good of the investment is low enough, O firms have a larger market share for any n_o, and therefore all firms decide to collaborate in the open-source project. Lemma A2 complements Corollary A2, by providing the necessary and sufficient condition for an equilibrium with $n_o = n$.

Lemma A2 (Necessary and Sufficient condition for equilibrium with $n_o = n$)**.** *Given $n > 3$ and δ, there exists $\bar{\alpha} \in (\hat{\alpha}, 1)$ such that $f(n) \geq 0$ if and only if $\alpha \leq \bar{\alpha}$.*

Proof. The equilibrium condition with $n_o = n$ depends only on the sign of $f(n_o)$ so, to simplify things, we will work with a scaled version of $f(n_o)$, $\delta f(n_o)$, for the rest of this proof. We know that $f(n) > 0$ for $\alpha < \hat{\alpha}(n-1, n)$. We need to determine the sign of $f(n)$ for the rest of values of α. When $n_o = n$, $s_o = 1/n$. Therefore,

$$f(n) = \frac{1}{n-1} - \frac{\delta(1-\alpha)}{n} - \frac{\tilde{s}_p}{1-\tilde{s}_p}\left(1 - \delta(1-\tilde{s}_p)\right), \tag{3.8.12}$$

where $\tilde{s}_p = s_p(n-1)$. We need to find the value of $\tilde{s}_p$ that makes $f(n) = 0$. There are two roots of this equation. The only positive root is

$$\tilde{s}_p = \frac{-n^2(1-\delta) - (1-\alpha)\delta - n\alpha\delta + \sqrt{n^4 - 2(n-1)n^2(n-1-\alpha)\delta + z^2}}{2\delta\, n\,(n-1)}$$

where $z = \delta\,(n-1)\,(n-1+\alpha)$. The corresponding value for $s_o(n-1)$ is

$$\tilde{s}_o = \frac{n^2 + z - \sqrt{n^4 - 2(n-1)n^2(n-1-\alpha)\delta + z^2}}{2\delta\, n\,(n-1)^2}$$

Plugging this value in the equilibrium condition (3.8.10) and solving for α we obtain the value $\bar{\alpha}$ where $f(n) = 0$. Lemma A1 implies that $\bar{\alpha} \geq \hat{\alpha}(n-1, n)$. Lemma 2 implies that $\partial\tilde{s}_p/\partial\alpha > 0$ in the relevant area. This means that $\bar{\alpha}$ is the unique value of α such that $f(n) = 0$.

To finish the proof we need to show that $f(n) > 0$ for $\alpha < \bar{\alpha}$ and $f(n) < 0$ for $\alpha > \bar{\alpha}$. Given the continuity and monotonicity of s_p, it suffices to show there is some value to the right or to the left of $\bar{\alpha}$ such that these inequalities hold.

Consider first the case of $\alpha < \bar{\alpha}$. We know that at $\alpha = \hat{\alpha}(n-1, n)$, $f(n) > 0$. This proves that $f(n) > 0$ for $\alpha < \bar{\alpha}$. For $\alpha > \bar{\alpha}$, consider $\alpha = 1$. When $\alpha = 1$, the investment of O firms is very low, and P firms have the largest advantage. In this case, $f(n) < 0$, which proves that this inequality holds for any $\alpha > \bar{\alpha}$. ■

Proposition 2 follows directly from Corollaries A1 and A2, and Lemma A2. ■

Proof of Lemma 4. Lemma 4 follows directly from Corollary A1. ■

Proof of Proposition 3. When $n \to \infty$, there are three types of equilibria. Remember we defined $\pi_k(n_o)$ as the profit of firm type k when there are n_o O firms. For an equilibrium with coexistence, we need $\lim_{n\to\infty} \frac{\pi_o(n_o)}{\pi_p(n_o)} = 1$. For an equilibrium with only O firms, we need $\lim_{n\to\infty} \frac{\pi_o(n)}{\pi_p(n)} \geq 1$, and for an equilibrium with only P firms, we need $\lim_{n\to\infty} \frac{\pi_o(2)}{\pi_p(1)} \leq 1$.

Taking the limit of equation 3.3.4, we obtain a function determining the limit of the ratio of market shares,

$$(1-\delta)\ln \lim_{n\to\infty}\frac{s_o}{s_p} = \delta \ln \lim_{n\to\infty}\left[\left(1-\alpha\frac{n_o-1}{n_o(1-s_o)}\right)n_o^{\alpha}\right],$$

and from the definition of π_o and π_p in equations (3.3.6) and (3.3.7), we obtain the ratio of profits,

$$\lim_{n\to\infty}\frac{\pi_o}{\pi_p} = \lim_{n\to\infty}\frac{1-\delta\left(1-\alpha\frac{n_o-1}{n_o}\right)}{1-\delta}\lim_{n\to\infty}\frac{s_o}{s_p}.$$

In an equilibrium with coexistence, $n_o\to\infty$ as $n\to\infty$. There are two cases. If $\alpha<1$, then $\frac{s_o}{s_p}\to\infty$ and $\frac{\pi_o}{\pi_p}\to\infty$, which contradicts the necessary condition for an equilibrium with coexistence. $\alpha=1$, on the other hand,

$$(1-\delta)\ln \lim_{n\to\infty}\frac{s_o}{s_p} = \delta \ln \lim_{n_o\to\infty}(1-n_o\, s_o),$$

so $\frac{s_o}{s_p}$ may converge to a constant if $n_o\, s_o$ converges to a constant. Specifically, if $n_o\, s_o\to 1-(1-\delta)^{\frac{1-\delta}{\delta}}$, then $\frac{s_o}{s_p}\to 1-\delta$ and $\frac{\pi_o}{\pi_p}\to 1$, and there is an equilibrium with coexistence.

Let us now look for an equilibrium with only O firms. If $\alpha<1$, $\frac{s_o}{s_p}\to\infty$ and $\frac{\pi_o}{\pi_p}\to\infty$, which verifies the condition for an equilibrium with only O firms. If $\alpha=1$, $n_o\, s_o\to 1$, which means that $\frac{s_o}{s_p}\to 0$ and $\frac{\pi_o}{\pi_p}\to 0$, which contradicts the necessary condition for an equilibrium with only O firms.

Finally, let us look for an equilibrium with only P firms. We have to show that when $n_o=2$, O firms would gain by becoming P. When $n_o=2$,

$$\lim_{n\to\infty}\frac{s_o}{s_p} = \left[\left(1-\frac{\alpha}{2}\right)2^{\alpha}\right]^{\frac{\delta}{1-\delta}},$$

$$\lim_{n\to\infty}\frac{\pi_o(2)}{\pi_p(1)} = \frac{1-\delta\left(1-\alpha\frac{1}{2}\right)}{1-\delta}\left[\left(1-\frac{\alpha}{2}\right)2^{\alpha}\right]^{\frac{\delta}{1-\delta}}.$$

The above expression is larger than 1 for all α and δ. Therefore, there is no equilibrium with only P firms. ■

Proof of Lemma 5. The equilibrium condition for the second stage is fully characterized by

$$(1-\delta)\ln\left(\frac{s_o}{s_p}\right)+\frac{1}{1-s_o}-\frac{1}{1-s_p} = \delta\ln\left(1-\alpha\frac{n_o-1}{n_o(1-s_o)}\right)+\alpha\,\delta\ln(n_o)-\delta\ln(1-\kappa), \quad (3.8.13)$$

and the market clearing condition $n_o\, s_o+(n-n_o)\, s_p=1$. The left-hand side in equation 3.8.13 is zero when $s_o=s_p$ and it is strictly increasing in s_o. Then $s_o>s_p$ if and only if the right-hand side is positive. Finally, the right-hand side is positive if and only if $1-\kappa<\left(1-\alpha\frac{n}{n-1}\frac{n_o-1}{n_o}\right)n_o^{\alpha}$. ■

Proof of Proposition 4. Suppose that all firms are O and let us compute the optimal subsidy that achieves the first best solution. Later we will show the optimal subsidy indeed induced all firms to be O. When $n_o = n$ an individual O firm invest $x_o = (1-\alpha)/(n\,c_o)$. The first best optimal investment is $x^* = 1/(c\,n)$. A government achieves the first best solution if and only if $c = c_o/(1-\alpha) = c(1-\kappa)/(1-\alpha)$. Then the optimal subsidy is $\kappa^* = \alpha$. By lemma 5, if $\kappa^* = \alpha > 1 - \left(1 - \alpha\frac{n(n-2)}{(n-1)^2}\right)(n-1)^{\alpha}$ then all firms are O. But this condition holds for any $\alpha \in [0,1]$. ■

Proof of Proposition 5. The first order conditions are (3.8.1) (3.8.2). As in the standard logit model, there are no corner solutions so the first order conditions hold with equality. Equilibrium prices and investment for P firms are identical to the logit model so we will focus on the O firms. In the case of O firms the partial derivative of the market share with respect to the price is

$$\frac{\partial s_i}{\partial p_i} = -\frac{\delta}{(1-\sigma)} s_i(1 - \sigma\, s_{i|o} - (1-\sigma)\, s_i).$$

Then from the equation (3.8.1) and imposing symmetry we obtain the optimal price in equation (3.7.3). To find x_o we need to calculate $\partial s_i/\partial x_i$ for O firms:

$$\frac{\partial s_i}{\partial x_i} = \delta\,(1 - \sum_{i \in O} s_i)\frac{\alpha}{\sum_{i \in O} x_i} + \delta\frac{s_i(1 - \sigma\, s_{i|o} - (1-\sigma)\, s_i)}{(1-\sigma)}\frac{(1-\alpha)}{x_i}.$$

From equation (3.8.2) and imposing symmetry we obtain equation (3.7.4), and the ratio of optimal investments in equilibrium:

$$\frac{x_o}{x_p} = \frac{s_o}{s_p}\left(1 - \frac{\alpha}{1-\sigma}\,\frac{(n_o-1)/n_o}{\left(1 - s_o + \frac{\sigma}{1-\sigma}\,\frac{n_o-1}{n_o}\right)}\right). \tag{3.8.14}$$

The ratio of market shares between O and P firms is

$$\frac{s_o}{s_p} = n_o^{-\sigma}\exp\left(\delta(q_o - q_p + p_p - p_o)\right), \tag{3.8.15}$$

$$\ln\left(\frac{s_o}{s_p}\right) = -\sigma\ln n_o + \delta\,(q_o - q_p) + \frac{1}{1-s_p} - \frac{1}{1 - s_o + \frac{\sigma}{1-\sigma}\,\frac{n_o-1}{n_o}}, \tag{3.8.16}$$

and from the definition of q_p and q_o we obtain:

$$q_o - q_p = \ln\left(\frac{x_o}{x_p}\right) + \alpha\,\ln(n_o). \tag{3.8.17}$$

From equations (3.8.14), (3.8.16) and (3.8.17), we obtain equation (3.7.5), which is an implicit equation determining the relation of market shares between O and P firms in equilibrium. This equation, together with the equation establishing that the sum of the market shares is equal to 1, completely characterizes the equilibrium.

Finally, to show existence and uniqueness, we need to prove two things: (1) there is only one fixed point of the system of equations in Proposition 5 (there is only one symmetric continuation equilibrium), and (2) the profit function is concave at the equilibrium (the second order conditions for optimality hold).

To show (1), define $g(s_o)$ by plugging (3.3.5) in equation (3.7.5). Then, the result follows from an application of the mean value theorem as in the standard logit model.

To prove that the profit function is concave at the equilibrium candidate, we will evaluate the determinant of the Hessian of the profit function at the equilibrium price and market share, and show that it is positive definite. The determinant of the Hessian for O firms is

$$|H_o| \geq \delta \frac{s_o^2}{(1-\sigma)\, x_o^2}(1-\sigma s_{i|o}) \left(\frac{(1-\sigma)}{\delta} \left(\frac{(1-\sigma)(1-n_o\, s_o)}{(1-\sigma s_{i|o} - (1-\sigma)s_o)n_o^2}\alpha + (1-\alpha) \right) + \right.$$
$$\left. - \left(\frac{(1-\sigma)(1-n_o\, s_o)}{(1-\sigma s_{i|o} - (1-\sigma)s_o)n_o}\alpha + (1-\alpha) \right)^2 \right).$$

The determinant of the Hessian for P firms is equivalent to that of the standard logit model. A sufficient condition for both determinants to be positive is $(1-\sigma) \geq \delta$ or $\sigma \leq (1-\delta)$, which has been assumed for this section of the paper. Thus, the concavity of the profit function at the equilibrium is guaranteed for both types of firms. ■

Proof of Proposition 6. We have characterized the equilibrium when $\sigma = 0$ in Proposition 3. Therefore, for the rest of the proof, assume $\sigma > 0$. Taking the limit of equation (3.7.5) we obtain the ratio of market shares in equilibrium, and from the definition of π_o and π_p, we obtain the ratio of profits,

$$\lim_{n\to\infty} \frac{\pi_o}{\pi_p} = \lim_{n\to\infty} \frac{1}{1+\frac{\sigma}{1-\sigma}\frac{n_o-1}{n_o}} \lim_{n\to\infty} \frac{1-\delta+\frac{\delta}{1-\sigma}(\alpha-\sigma)\frac{n_o-1}{n_o}}{1-\delta} \lim_{n\to\infty} \frac{s_o}{s_p}.$$

First, we look for equilibria with $n_o \to \infty$ (coexistence or only O firms). If $\alpha < 1$, then

$$(1-\delta)\,\ln \lim_{n\to\infty} \frac{s_o}{s_p} = \sigma + \delta \ln \left[(1-\alpha) \lim_{n_o\to\infty} n_o^{\alpha-(\sigma/\delta)} \right].$$

There are three cases when $\alpha < 1$. If $\alpha > \sigma/\delta$, then $\frac{s_o}{s_p} \to \infty$ and $\frac{\pi_o}{\pi_p} \to \infty$ when $n_o \to \infty$; therefore, there is an equilibrium with only O firms. If $\alpha < \sigma/\delta$, then $\frac{s_o}{s_p} \to 0$ and $\frac{\pi_o}{\pi_p} \to 0$, so there is no equilibrium in which $n_o \to \infty$. Finally, if $\alpha = \sigma/\delta$, then $\frac{s_o}{s_p} \to e^{\frac{\sigma}{1-\delta}} \left(1-\frac{\sigma}{\delta}\right)^{\frac{\delta}{1-\delta}} < 1$ and $\frac{\pi_o}{\pi_p} \to \frac{s_o}{s_p}$, so there is no equilibrium in which $n_o \to \infty$. If $\alpha = 1$, then

$$(1-\delta)\ln \lim_{n\to\infty} \frac{s_o}{s_p} = \sigma + \delta \ln \lim_{n_o\to\infty} \left[(1-\sigma)\,(1-n_o\, s_o)\, n_o^{-\sigma/\delta} \right],$$

so $\frac{s_o}{s_p} \to 0$ and $\frac{\pi_o}{\pi_p} \to 0$ when $n_o \to \infty$, so there is no equilibrium in which $n_o \to \infty$. Therefore, there is an equilibrium with $n_o \to \infty$, only when $\sigma/\delta < \alpha < 1$, and in such equilibrium, all firms choose to be O.

Next, we look for an equilibrium with only P firms, for which we need $\lim_{n\to\infty} \frac{\pi_o(2)}{\pi_p(1)} \leq 1$. Given that firms are infinitesimal when $n \to \infty$, the market share of P firms is the same when $n_o = 1$ and $n_o = 2$. Taking the limit of equation (3.7.5), we obtain

$$\lim_{n\to\infty} \frac{s_o}{s_p} = 2^{\frac{\alpha\,\delta-\sigma}{1-\delta}}\, e^{\frac{\sigma}{(1-\delta)(2-\sigma)}} \left(1 - \frac{\alpha}{2-\sigma}\right)^{\frac{\delta}{1-\delta}},$$

and therefore, the limit of the ratio of profits is

$$\lim_{n\to\infty} \frac{\pi_o(2)}{\pi_p(1)} = \frac{2\,(1-\sigma)\,(1-\delta) + \delta\,(\alpha-\sigma)}{(1-\delta)\,(2-\sigma)}\, 2^{\frac{\alpha\,\delta-\sigma}{1-\delta}}\, e^{\frac{\sigma}{(1-\delta)(2-\sigma)}} \left(1 - \frac{\alpha}{2-\sigma}\right)^{\frac{\delta}{1-\delta}}.$$

Denote the right-hand side of the above expression by $F(\alpha,\delta,\sigma)$, and let $\tilde{\sigma}(\alpha,\delta)$ be the value of σ which solves $F(\alpha,\delta,\sigma) = 1$. $F(\alpha,\delta,\sigma)$ is decreasing in σ, which means that $\lim_{n\to\infty} \frac{\pi_o(2)}{\pi_p(1)} \leq 1$ (i.e., there is an equilibrium with only P firms) if and only if $\sigma \geq \tilde{\sigma}(\alpha,\delta)$. Note that this condition holds not only for $\alpha < 1$, but also for $\alpha = 1$.

Next, we show $0 < \tilde{\sigma} < \alpha\,\delta$. We know $\tilde{\sigma} > 0$ from Proposition 3. Also, $\tilde{\sigma} \geq \alpha\,\delta$ if and only if $F(\alpha,\delta,\alpha\,\delta) \geq 1$. Substituting $\sigma = \alpha\,\delta$ into $F(\alpha,\delta,\sigma)$, and rearranging, we obtain the following condition:

$$\frac{\alpha}{2-\alpha\,\delta} + \ln\left(1 - \frac{\alpha}{2-\alpha\,\delta}\right) \geq 0,$$

which is not possible, given that $\alpha/(2-\alpha\,\delta) > 0$. Therefore, $\tilde{\sigma} < \alpha\,\delta$.

Finally, we know that when $\alpha = 1$, there is no equilibrium such that $n_o \to \infty$, i.e. there is no equilibrium in which all firms choose to be O, and there is no equilibrium in which a proportion of firms chooses to be O. We also know there is an equilibrium with only P firms only when $\sigma \geq \tilde{\sigma}(\alpha,\delta)$. Therefore, it remains to show what is the equilibrium when $\alpha = 1$ and $0 < \sigma < \tilde{\sigma}(\alpha,\delta)$.

It is easy to see that in this case, there is an equilibrium with coexistence, in which the number of O firms converges to a constant. To see this, note that $\sigma < \tilde{\sigma}(\alpha,\delta)$ implies that $\frac{\pi_o(2)}{\pi_p(1)} > 1$, and that we have shown that $\lim_{n_o\to\infty} \frac{\pi_o}{\pi_p} < 1$ when $\alpha = 1$. Therefore, there is a constant $2 \leq n_o^* < \infty$ such that $\frac{\pi_o(n_o^*)}{\pi_p} \geq 1$ and $\frac{\pi_o(n_o^*+1)}{\pi_p} \leq 1$ (the market share of P firms does not change when n_o changes and n_o is finite, since in this case O firms have an aggregate market share equal to zero). ■

Proof of Proposition 7. To show existence, we begin by noticing that for any value of $S_o \in [0,1]$, there is a value of $s_o \in (0, 1/n_o)$ that solves equation 3.7.7. This is because the left hand side is continuous for $s_o \in (0, 1/n_o)$, and goes from $-\infty$ when $s_o \to 0$ to ∞ when $s_o \to 1/n$. By a similar argument, for any value of $s_o \in (0, 1/n_o]$ there is a value of $S_o \in [0,1]$ that solves equation 3.7.8. This implies that an equilibrium exists.

To show uniqueness, let us denote equation (3.7.7) by $f(s_o, S_o) = 0$ and equation (3.7.8) by $g(s_o, S_o) = 0$. By the implicit function theorem we can write $S_o = f_1(s_o)$ using the first equation, and $S_o = g_1(s_o)$ using the second equation. It is straightforward to show that both functions are continuously increasing, but the slope of f_1 is always less than 1, while the slope of g_1 is always larger than 1. This means that the two curves cross only once, so the symmetric continuation equilibrium is unique.

Now we show that $s_p > s_o$. In equilibrium, the following condition must hold:

$$\ln\left(\frac{s_o}{s_p}\right) + \frac{1}{1-s_o} - \frac{1}{1-s_p} = -\frac{\delta(1-\alpha)}{(1-\sigma)} \ln\left(\frac{\alpha}{(1-\alpha)}\frac{S_p}{s_o} + \frac{s_p}{s_o}\right). \tag{3.8.18}$$

When $s_p = s_o = 1/n$, the condition becomes

$$0 = -\frac{\delta(1-\alpha)}{(1-\sigma)} \ln\left(1 + \frac{\alpha}{(1-\alpha)} S_p\, n\right).$$

Define m as the right hand side of the previous equality. $s_p = s_o$ is not an equilibrium because $m < 0$ when $s_p = s_o$. For equation (3.8.18) to return to the equilibrium, s_p has to increase and s_o has to decrease relative to $s_p = s_o$. Therefore, the equilibrium has $s_p > s_o$.

Next, we show that $S_o > S_p$. We will prove this by contradiction. By a similar argument than before, for $S_o \leq S_p$ the following condition must hold:

$$-\ln(n_o) + \ln\left(\frac{1-n_o s_o}{s_o} + \frac{\alpha}{2(1-\alpha)s_o}\right) - \frac{2}{\delta\alpha} \geq 0.$$

This implies that we should have $s_o \leq \frac{\alpha+2(1-\alpha)}{2\beta\left(1+e^{2/(\delta\alpha)}\right)n_o}$ in equilibrium. If we introduce $s_o = \frac{\alpha+2(1-\alpha)}{2(1-\alpha)\left(1+e^{2/(\delta\alpha)}\right)n_o}$ in the equilibrium condition (3.7.7), and rearrange terms, we get that for this value of s_o to be an equilibrium the following should hold:

$$\begin{aligned} 0 \quad = \quad & -\frac{2(1-\alpha)}{\alpha+2(1-\alpha)}\left(1+e^{\frac{2}{(\delta\alpha)}}\right) + \log\left(\frac{\alpha+2(1-\alpha)}{2(1-\alpha)e^{\frac{2}{(\delta\alpha)}} - \alpha}\right) + \frac{2(1-\alpha)}{\alpha(1-\sigma)} \\ & + \frac{n_o}{n_o - \frac{(\alpha+2(1-\alpha))\left(1-\tanh\left(\frac{1}{(\delta\alpha)}\right)\right)}{4(1-\alpha)}} + \left(\frac{\delta(1-\alpha)}{(1-\sigma)} - 1\right)\log(n_o) \end{aligned} \tag{3.8.19}$$

Let $\hat{m}$ be the right hand side of the previous expression. It can be shown that $\hat{m} < 0$ for any value of the parameters. Therefore, $s_o > \frac{\alpha+2(1-\alpha)}{2(1-\alpha)\left(1+e^{2/(\delta\alpha)}\right)n_o}$ in equilibrium, and therefore, $S_o \leq S_p$ cannot hold.

Finally, to see that the P firm will always have a higher profit than an O firm, notice that the P firm can always choose the same investment and price than an O firm. The P firm will have the same revenues due to the complementary good, and the same cost of R&D, but it will also have some revenue on the primary good side. ■

References

ABOAL, D., P. GARDA, B. LANZILOTTA, AND M. PERERA (2015): "Innovation, Firm Size, Technology Intensity, and Employment Generation: Evidence from the Uruguayan Manufacturing Sector," *Emerging Markets Finance and Trade*, 51(1), 3–26.

ACEMOGLU, D. (1998): "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *The Quarterly Journal of Economics*, 113(4), 1055–1089.

——— (2003): "Patterns of Skill Premia," *The Review of Economic Studies*, 70(2), 199–230.

ADAMS, R. M., K. P. BREVOORT, AND E. K. KISER (2007): "Who Competes with Whom? The Case of Depository Institutions," *Journal of Industrial Economics*, 55(1), 141–167.

AGUIRREGABIRIA, V., R. CLARK, AND H. WANG (2015): "Diversification of Geographic Risk in Retail Bank Networks: Evidence from Bank Expansion after the Riegle-Neal Act," Working paper.

ALLEN, R. C. (1983): "Collective invention," *Journal of Economic Behavior and Organization*, 4(1), 1–24.

ALVAREZ, R., J. BENAVENTE, R. CAMPUSANO, AND C. CUEVAS (2011): "Employment Generation, Firm Size, and Innovation in Chile," IDB Publications 54258, Inter-American Development Bank, Washington, DC.

——— (2012): "Public Programs, Innovation, and Firm Performance in Chile," IDB Publications 62658, Inter-American Development Bank, Washington, DC.

Amel, D., and M. Starr-McCluer (2002): "Market Definition in Banking: Recent Evidence," *Antirust Bulletin*, 47(1), 63–89.

Amir, R., I. Evstigneev, and J. Wooders (2003): "Noncooperative versus cooperative R&D with endogenous spillover rates," *Games and Economic Behavior*, 42(2), 183–207.

Anderson, S., and A. de Palma (1992): "Multiproduct firms: a nested logit approach," *Journal of Industrial Economics*, 40(3), 261–276.

Anderson, S., A. de Palma, and J. Thisse (1992): *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA.

Anderson, S., and L. Leruth (1993): "Why firms may prefer not to price discriminate via mixed bundling," *International Journal of Industrial Organization*, 11(1), 49–61.

Arora, A., and F. Bokhari (2007): "Open versus closed firms and the dynamics of industry evolution," *Journal of Industrial Economics*, 55(3), 499–527.

Athey, S., and G. Ellison (2014): "Dynamics of Open Source Movements," *Journal of Economics & Management Strategy*, 23(2), 294–316.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?," *The Quarterly Journal of Economics*, 113(4), 1169–1213.

Barlevy, G. (2007): "On the Cyclicality of Research and Development," *American Economic Review*, 97(4), 1131–1164.

Basu, S., and J. Fernald (2001): "Why is productivity procyclical? Why do we care?," in *New developments in productivity analysis*, pp. 225–302. University of Chicago Press.

Ben-Akiva, M. (1973): "Structure of Passenger Travel Demand Models," Ph.D. thesis, Massachusetts Institute of Technology.

Ben-Akiva, M., and S. Lerman (1985): *Discrete choice analysis: theory and application to travel demand*. MIT press.

Benavente, J. M., and R. Lauterbach (2008): "Technological innovation and employment: complements or substitutes?," *The European Journal of Development Research*, 20(2), 318–329.

Berger, A. N., S. D. Bonime, L. G. Goldberg, and L. J. White (2004): "The Dynamics of Market Entry: The Effects of Mergers and Acquisitions on Entry in the Banking Industry," *Journal of Business*, 77(4), 797–834.

Berger, A. N., N. H. Miller, M. A. Petersen, R. G. Rajan, and J. C. Stein (2005): "Does function follow organizational form? Evidence from the lending practices of large and small banks," *Journal of Financial Economics*, 76(2), 237–269.

Besanko, D., M. Perry, and R. Spady (1990): "The logit model of monopolistic competition: brand diversity," *Journal of Industrial Economics*, 38(4), 397–415.

Bessen, J. (2006): "Open Source Software: Free Provision of Complex Public Goods," in *The Economics of Open Source Software Development*, ed. by J. Bitzer, and P. Schröder. Elsevier, Amsterdam.

Biehl, A. R. (2002): "The extent of the market for retail banking deposits," *Antitrust Bulletin*, 47(1), 91–106.

Binelli, C., and A. Maffioli (2007): “A Micro-econometric Analysis of Public Support to Private R&D in Argentina,” *International Review of Applied Economics*, 21(3), 339–359.

Bitzer, J. (2004): “Commercial versus open source software: the role of product heterogeneity in competition,” *Economic Systems*, 28(4), 369–381.

Blanchflower, D. G., and S. M. Burgess (1998): “New Technology And Jobs: Comparative Evidence From A Two Country Study,” *Economics of Innovation and New Technology*, 5(2-4), 109–138.

Bloch, F. (1995): “Endogenous Structures of Association in Oligopolies,” *RAND Journal of Economics*, 26(3), 537–556.

Blundell, R., and S. Bond (1998): “Initial conditions and moment restrictions in dynamic panel data models,” *Journal of Econometrics*, 87(1), 115–143.

Bonaccorsi, A., and C. Rossi (2004): “Altruistic individuals, selfish firms? The structure of motivation in open source software,” *First Monday*, 9(1).

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002): “Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence,” *The Quarterly Journal of Economics*, 117(1), 339–376.

Bresnahan, T. F., and P. C. Reiss (1991): “Entry and Competition in Concentrated Markets,” *Journal of Political Economy*, 99(5), 977–1009.

Brickley, J. A., J. S. Linck, and C. W. Smith Jr. (2003): “Boundaries of the firm: evidence from the banking industry,” *Journal of Financial Economics*, 70(3), 351–383.

Canales, R., and R. Nanda (2012): “A darker side to decentralized banks: Market power and credit rationing in SME lending,” *Journal of Financial Economics*, 105(2), 353–366.

Cardell, N. (1997): "Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity," *Econometric Theory*, 13(2), 185–213.

Carmen Matutes, P. R. (1992): "Compatibility and Bundling of Complementary Goods in a Duopoly," *The Journal of Industrial Economics*, 40(1), 37–54.

Caroli, E., and J. Van Reenen (2001): "Skill-biased organizational change? Evidence from a panel of British and French establishments," *The Quarterly Journal of Economics*, 116(4), 1449–1492.

Casadesus-Masanell, R., and P. Ghemawat (2006): "Dynamic Mixed Duopoly: A Model Motivated by Linux vs. Windows," *Management Science*, 52(7), 1072–1084.

Casadesus-Masanell, R., and G. Llanes (2011): "Mixed source," *Management Science*, 57(7), 1212–1230.

Castillo, V., A. Maffioli, S. Rojo, and R. Stucchi (2014): "The effect of innovation policy on SMEs employment and wages in Argentina," *Small Business Economics*, 42(2), 387–406.

Cohen, A., and M. J. Mazzeo (2007): "Market Structure and Competition among Retail Depository Institutions," *Review of Economics and Statistics*, 89(1), 60–74.

Czarnitzki, D., and S. Thorwarth (2012): "Productivity effects of basic research in low-tech and high-tech industries," *Research Policy*, 41, 1555–1564.

Dahlander, L. (2005): "Appropriation and Appropriability in Open Source Software," *International Journal of Innovation Management*, 9(3), 259–285.

D'Aspremont, C., and A. Jacquemin (1988): "Cooperative and noncooperative R&D in duopoly with spillovers," *American Economic Review*, 78(5), 1133–1137.

D'Aspremont, C., A. Jacquemin, J. Gabszewicz, and J. Weymark (1983): "On the stability of collusive price leadership," *Canadian Journal of Economics/Revue canadienne d'Economique*, 16(1), 17–25.

Dick, A. (2007): "Market Size, Service Quality and Competition in Banking," *Journal of Money, Credit and Banking*, 39(1), 50–81.

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

Doms, M., T. Dunne, and K. R. Troske (1997): "Workers, Wages, and Technology," *The Quarterly Journal of Economics*, 112(1), 253–290.

Dunne, T., S. D. Klimek, M. J. Roberts, and D. Y. Xu (2013): "Entry, exit, and the determinants of market structure," *The RAND Journal of Economics*, 44(3), 462–487.

Economides, N., and E. Katsamakas (2006): "Two-sided competition of proprietary vs. open source technology platforms and the implications for the software industry," *Management Science*, 52(7), 1057–1071.

Entorf, H., and W. Pohlmeier (1990): "Employment, Innovation and Export Activity," in *Microeconometrics: surveys and Applications*, ed. by J. L. J. Florens, M. Ivaldi, and F. Laisney, pp. 349–415. Basil Blackwell, London.

Fershtman, C., and N. Gandal (2011): "A Brief Survey of the Economics of Open Source Software," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan, Basingstoke.

Foray, D., and L. Perez (2006): "The economics of open technology: collective organisation and individual claims in the 'fabrique lyonnaise' during the old regime," in *New Frontiers in the Economics of Innovation And New Technology: Essays in*

Honor of Paul A. David, ed. by C. Antonelli, P. A. David, D. Foray, B. H. Hall, and W. E. Steinmueller. Edward Elgar Publishing.

Fosfuri, A., M. Giarratana, and A. Luzzi (2008): "The Penguin Has Entered the Building: The Commercialization of Open Source Software Products," *Organization Science*, 19(2), 292–305.

Gaudeul, A. (2007): "Do open source developers respond to competition? The (La)TeX case study," *Review of Network Economics*, 6(2), 239–263.

Giuliodori, D., and R. Stucchi (2012): "Innovation and job creation in a dual labor market: Evidence from Spain," *Economics of Innovation and New Technology*, 21(8), 801–813.

Greenan, N. (2003): "Organisational change, technology, employment and skills: an empirical study of French manufacturing," *Cambridge Journal of Economics*, 27(2), 287–316.

Griliches, Z. (1969): "Capital-Skill Complementarity," *The Review of Economics and Statistics*, 51(4), 465–468.

Guner, N., G. Ventura, and Y. Xu (2008): "Macroeconomic implications of size-dependent policies," *Review of Economic Dynamics*, 11(4), 721–744.

Hall, B. H., F. Lotti, and J. Mairesse (2008): "Employment, innovation, and productivity: evidence from Italian microdata," *Industrial and Corporate Change*, 17(4), 813–839.

Harrison, R., J. Jaumandreu, J. Mairesse, and B. Peters (2014): "Does innovation stimulate employment? A firm-level analysis using comparable micro-data from four European countries," *International Journal of Industrial Organization*, 35, 29–43.

Haruvy, E., S. Sethi, and J. Zhou (2008): "Open source development with a commercial complementary product or service," *Production and Operations Management*, 17(1), 29–43.

Heitfield, E., and R. A. Prager (2004): "The Geographic Scope of Retail Deposit Markets," *Journal of Financial Services Research*, 25(1), 37–55.

Henkel, J. (2004): "The Jukebox Mode of Innovation - A Model of Commercial Open Source Development," CEPR Discussion Papers 4507, C.E.P.R.

Henkel, J. (2006): "Selective revealing in open innovation processes: The case of embedded Linux," *Research Policy*, 35(7), 953–969.

Ho, K., and J. Ishii (2011): "Location and competition in retail banking," *International Journal of Industrial Organization*, 29(5), 537–546.

IDC (2008): "Worldwide Quarterly Server Tracker," http://www.idc.com/.

Ishii, J. (2005): "Compatibility, Competition, and Investment in Network Industries: ATM Networks in the Banking Industry," Working paper.

Jansen, J. (2009): "Share to scare: technology sharing in the absence of intellectual property rights," Discussion Paper 2009/36, Max Planck Institute for Research on Collective Goods, Bonn.

Johnson, J. (2002): "Open source software: private provision of a public good," *Journal of Economics and Management Strategy*, 11(4), 637–662.

——— (2006): "Collaboration, peer review and open source software," *Information Economics and Policy*, 18(4), 477–497.

Kamien, M. I., E. Muller, and I. Zang (1992): "Research Joint Ventures and R&D Cartels," *American Economic Review*, 82(5), 1293–1306.

Katsoulacos, Y., and D. Ulph (1998): "Endogenous Spillovers and the Performance of Research Joint Ventures," *Journal of Industrial Economics*, 46(3), 333–357.

Kroszner, R. S., and P. E. Strahan (1999): "What Drives Deregulation? Economics and Politics of the Relaxation of Bank Branching Restrictions," *The Quarterly Journal of Economics*, 114(4), 1437–1467.

Kwast, M. L., M. Starr-McCluer, and J. D. Wolken (1997): "Market definition and the analysis of antitrust in banking," *Antitrust Bulletin*, 42(4), 973–995.

Kyriazidou, E., and M. Pesendorfer (1999): "Viennese Chairs: A Case Study for Modem Industrialization," *Journal of Economic History*, 59(1), 143–166.

Lachenmaier, S., and H. Rottmann (2011): "Effects of innovation on employment: A dynamic panel analysis," *International Journal of Industrial Organization*, 29(2), 210–220.

Lakhani, K., and R. Wolf (2005): "Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects," in *Perspectives on Free and Open Source Software*, ed. by J. Feller, B. Fitzgerald, S. Hissam, and K. R. Lakhani, pp. 3–22. MIT Press.

Lerner, J., P. Pathak, and J. Tirole (2006): "The Dynamics of Open Source Contributors," *American Economic Review*, 96(2), 114–118.

Lerner, J., and J. Tirole (2001): "The open source movement: Key research questions," *European Economic Review*, 45(4-6), 819–826.

——— (2002): "Some Simple Economics of Open Source," *Journal of Industrial Economics*, 50(2), 197–234.

——— (2005): "The Economics of Technology Sharing: Open Source and Beyond," *Journal of Economic Perspectives*, 19(2), 99–120.

Maurer, S. M., and S. Scotchmer (2006): “Open Source Software: The New Intellectual Property Paradigm,” NBER Working Papers 12148, National Bureau of Economic Research.

Mazzeo, M. J. (2002): “Product Choice and Oligopoly Market Structure,” *RAND Journal of Economics*, 33(2), 221–242.

McFadden, D. (1974): “Conditional logit analysis of qualitative choice behavior,” in *Frontiers in Econometrics*, ed. by P. Zarembka. Academic Press, New York.

——— (1986): “The Choice Theory Approach to Market Research,” *Marketing Science*, 5(4), 275–297.

McGaw, J. (1987): *Most Wonderful Machine: Mechanization and Social Change in Berkshire Paper Making, 1801-1885*. Princeton University Press, Princeton, NJ.

Meschi, E., E. Taymaz, and M. Vivarelli (2011): “Trade, technology and skills: Evidence from Turkish microdata,” *Labour Economics*, 18, Supplement 1, S60 – S70.

Monge-Gonzalez, R., J. Rodriguez-Álvarez, J. Hewitt, J. Orozco, and K. Ruiz (2011): “Innovation and Employment Growth in Costa Rica: A firm-level Analysis,” IDB Technical Note 318, Inter-American Development Bank, Washington, DC.

Mustonen, M. (2003): “Copyleft - the economics of Linux and other open source software,” *Information Economics and Policy*, 15(1), 99–121.

Niedermayer, A. (2013): “On platforms, incomplete contracts, and open source software,” *International Journal of Industrial Organization*, 31(6), 714–722.

Nuvolari, A. (2004): “Collective invention during the British Industrial Revolution: the case of the Cornish pumping engine,” *Cambridge Journal of Economics*, 28(3), 347–363.

Pagés, C. (2010): *The age of productivity: transforming economies from the bottom up*. Palgrave Macmillan.

Pakes, A., M. Ostrovsky, and S. Berry (2007): "Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry Exit Examples)," *RAND Journal of Economics*, 38(37), 619–640.

Pesendorfer, M., and P. Schmidt-Dengler (2008): "Asymptotic Least Squares Estimators for Dynamic Games," *Review of Economic Studies*, 75(3), 901–928.

Petersen, M. A., and R. G. Rajan (1994): "The Benefits of Lending Relationships: Evidence from Small Business Data," *The Journal of Finance*, 49(1), 3–37.

Petit, P. (1995): "Employment and Technological Change," in *Handbook of the Economics of Innovation and Technological Change*, ed. by P. Stoneman, pp. 366–408. Wiley-Blackwell.

Pianta, M. (2005): "Innovation and Employment," in *The Oxford Handbook of Innovation*, ed. by J. Fagerberg, D. C. Mowery, and R. R. Nelson, pp. 568–598. Oxford University Press, Oxford.

Piva, M., and M. Vivarelli (2005): "Innovation and employment: Evidence from Italian microdata," *Journal of Economics*, 86(1), 65–83.

Polanski, A. (2007): "Is the General Public License a rational choice?," *Journal of Industrial Economics*, 55(4), 691–714.

Poyago-Theotoky, J. (1995): "Equilibrium and Optimal Size of a Research Joint Venture in an Oligopoly with Spillovers," *Journal of Industrial Economics*, 43(2), 209–226.

Rice, T., and P. E. Strahan (2010): "Does Credit Competition Affect Small-Firm Finance?," *Journal of Finance*, 65(3), 861–889.

Sapienza, P. (2002): "The Effects of Banking Mergers on Loan Contracts," *The Journal of Finance*, 57(1), 329–367.

Saxonhouse, G. (1974): "A Tale of Japanese Technological Diffusion in the Meiji Period," *Journal of Economic History*, 34(1), 149–165.

Schmidtke, R. (2006): "Private Provision of a Complementary Public Good," CESifo Working Paper Series 1756, CESifo GmbH.

Seim, K. (2006): "An Empirical Model of Firm Entry with Endogenous Product-Type Choices.," *RAND Journal of Economics*, 37(3), 619–640.

Seppä, A. (2006): "Open Source in Finnish Software Companies," Discussion Papers 1002, The Research Institute of the Finnish Economy.

Smolny, W. (1998): "Innovations, Prices and Employment: A Theoretical Model and an Empirical Application for West German Manufacturing Firms," *The Journal of Industrial Economics*, 46(3), 359–381.

Stein, J. C. (2002): "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms," *The Journal of Finance*, 57(5), 1891–1921.

Stock, J. H., J. H. Wright, and M. Yogo (2002): "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business & Economic Statistics*, 20(4), 518–529.

Suzumura, K. (1992): "Cooperative and Noncooperative R&D in an Oligopoly with Spillovers," *American Economic Review*, 82(5), 1307–1320.

Train, K., D. McFadden, and M. Ben-Akiva (1987): "The demand for local telephone service: A fully discrete model of residential calling patterns and service choices," *RAND Journal of Economics*, 18(1), 109–123.

Van Reenen, J. (1997): "Employment and Technological Innovation: Evidence from U.K. Manufacturing Firms," *Journal of Labor Economics*, 15(2), 255–284.

Vivarelli, M. (1995): *The Economics of Technology and Employment: Theory and Empirical Evidence*. Elgar, Aldershot.

——— (2014): "Innovation, employment and skills in advanced and developing countries: a survey of economic literature," *Journal of Economic Issues*, 48(1), 123–154.

von Engelhardt, S. (2010): "Quality competition or quality cooperation? License-type and the strategic nature of open source vs. closed source business models," Discussion Paper 034, Jena Economic Research Papers.

von Krogh, G., and E. von Hippel (2006): "The Promise of Research on Open Source Software," *Management Science*, 52(7), 975–983.

Wang, H. (2009): "Consumer Valuation of Retail Networks: Evidence from the Banking Industry," Working paper.

Yi, S., and H. Shin (2000): "Endogenous formation of research coalitions with spillovers," *International Journal of Industrial Organization*, 18(2), 229–256.