CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 833

November 2007

Innovation in Business Groups

Sharon Belenzon and Tomer Berkovitz

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
Using novel data on European firms, this paper examines the effect of business group affiliation on innovation. We find that business groups foster the scale and novelty of corporate innovation. Group affiliation is particularly important in industries that rely more on external finance and have a higher degree of information asymmetry. We also find that the innovation of affiliates is less sensitive to operating cash flows. We interpret our results as supporting the 'bright side' of business group internal capital markets and explain how legal boundaries between group affiliates mitigate the inefficiencies found in internal capital markets of US conglomerates.

Keywords: business groups, innovation, internal capital markets
JEL Classifications: G34, L22, L26, and O32

This paper was produced as part of the Centre's Productivity and Innovation Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
We express our special gratitude to Patrick Bolton, Francisco Perez-Gonzalez, Mark Schankerman and John Van Reenen for numerous helpful discussions. We are also grateful for valuable comments from Nick Bloom, Steve Bond, Bjorn Jorgensen, Tarun Khanna, Daniel Paravisini, Tano Santos, Catherine Thomas, Manuel Trajtenberg, Daniel Wolfenzon, Yishay Yafeh and seminar participants at Columbia University (GSB), Hebrew University, London Business School Transatlantic Conference, London School of Economics, NBER, and Oxford University. We thank Liat Oren for invaluable assistance with the programming of the ownership algorithm, Hadar Gafni for excellent research assistance, and the Institute for Fiscal Studies, especially Rachel Griffith and Gareth Macartney, for the EPO patent matching. All remaining errors are our own.
Sharon Belenzon is a Research Economist with the Productivity and Innovation Research Group at the Centre for Economic Performance, London School of Economics. He is also a Postdoctoral Research Fellow at Nuffield College, University of Oxford. Tomer Berkovitz is a Postdoctoral Research Fellow at Columbia University, Graduate School of Business, NY, USA.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

ISBN 978-0-85328-208-2

1. Introduction

Extensive empirical literature over the past three decades has investigated the incentives of firms to innovate and the effect of innovation on performance.[1] More recently, La Porta et al. (1999) showed that outside the US, legally independent firms are commonly tied together through ownership links to form business groups. Yet, little attention has been devoted to the relation between business groups and innovation. This is especially surprising in light of the long debate in the literature on the effect of firm boundaries on the allocation of resources (Coase (1937), Mullainathan and Scharfstein (2001)).[2] In this paper, we provide a new perspective on how the boundaries between firms may affect the allocation of internal funds to R&D activity. Using a novel database, we show that while large organizations comprised of a single legal entity have been found to stifle innovation (e.g. Seru (2007) on US conglomerates), business groups foster innovation via a more efficient internal capital market.

In a world with asymmetric information, Myers and Majluf (1984) and Greenwald et al. (1984) argue that external financing is more costly than internal financing. If the asymmetric information problem is mitigated within a business group, an affiliated firm can gain access to more capital than a standalone firm and at a lower cost. Internal capital markets are especially important for innovation, where there is typically a high degree of asymmetric information between the firm and outside investors (Himmelberg and Peterson (1994)). While a standalone firm seeking to externally finance its R&D activity needs to turn to financial market intermediaries (venture capital funds, banks, etc.), a group-affiliated firm can rely on the group internal capital market.

Internal capital markets, however, have also been shown to have a 'dark side' in US conglomerates. Seru (2007) finds that conglomerates stifle innovation and relates his findings to inefficiencies in their internal capital markets. Seru finds that conglomerates employing central budget allocation are exposed to an agency problem between the division managers, who seek to maximize their budget, and the CEO, who acts to maximize firm value. This agency problem is particularly pronounced in novel R&D

[1] See Griliches (1998) for a comprehensive survey of the empirical literature.

[2] More recent studies on the effect of organizational design on innovation include Guedj and Scharfstein (2004), Guedj (2006), and Seru (2007).

projects where there is high degree of information asymmetry between the CEO and the division managers.

Business groups are fundamentally different from conglomerates since groups are composed of legally independent companies, typically, with minority shareholders. While the CEO of a conglomerate can shift funds from one division to the other at a low cost, the ultimate owner of a business group incurs larger costs when shifting funds between group members, especially when expropriating the rights of minority shareholders. In addition, the legal boundaries between group members allow for enforceable intra-group lending contracts, whereas in conglomerates, the lack of ex-post commitment mechanism renders intra-division contracts unenforceable.

This paper studies the effect of business group affiliation on innovation. We make three key contributions: (i) establish a positive effect of business group affiliation on the scale and novelty of corporate innovation, (ii) link this effect to the existence of efficient internal capital markets in business groups, and (iii) explain how the structural differences between business groups and conglomerates can explain the observed differences in the effect of their internal capital markets on innovation.

Business group affiliation is endogenous and might be affected by unobserved firm characteristics. More specifically, if groups can identify standalone firms with higher expected success probability, they may engage in 'winner-picking'. We mitigate the potential group selection bias in four ways. First, our ownership structure is stable over the estimation period 1995-2004. We merge our ownership data with a database on mergers and acquisitions and exclude firms that experience a change in their business group affiliation status, pushing back the potential endogeneity problem to the beginning of the sample. Second, we mitigate the effect of unobserved heterogeneity in firm-level innovation quality at the beginning of our sample by controlling for the pre-sample average number of patents of each firm (Blundell et al. (1999)). Third, we construct exogenous industry variables to study a specific channel through which group affiliation positively affects innovation. Finally, we mitigate the selection problem by examining specifications with only affiliated firms and analyzing the effect of group size on innovation. In such specifications the 'winner-picking' bias is likely to be smaller as now all firms in the sample belong to business groups.

Our analysis is based on four novel datasets on private and public European firms. First, in order to determine business group affiliation and construct detailed measures of group characteristics, we develop a unique algorithm that builds the complete structure of business groups based on approximately one million ownership links from the Amadeus ownership database. Second, to proxy for innovation, we match all patent applications from the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO) to all firms in the Amadeus database. Third, we use academic publications as a proxy for the novelty of innovation. For this purpose, we match the firm name in the Amadeus database to the address field in Thomson's ISI Web of Science database. Finally, we use current and historical versions of the Amadeus accounting database to build a comprehensive panel data of firm characteristics.

Using our newly assembled database, we investigate whether affiliated firms patent more than standalones and examine the novelty of their innovation. Controlling for various firm characteristics, we find that group-affiliated firms have 30 percent more patents than standalone firms. Yet, the difference in the number of patents may not reflect a difference in the quality of innovation, but rather the relative advantage of business groups in issuing patents. Groups may utilize economies of scale to employ a specialized team of patent attorneys and intellectual property experts to advise them on the best patenting strategies and have broader experience, talent management, etc. To address this concern, we examine the relative quality of patents using two measures: number of citations a patent receives and the effect of the number of patents on the firm's productivity. Both measures reveal that affiliates engage in more novel innovation.

Patents are a common measure of innovation. Yet, not all inventions are patentable and in many cases basic and novel research is published in academic journals (Cockburn and Henderson (1998)). Thus, we also test the effect of group affiliation on the number of academic publications in 'hard science' journals. To proxy for the quality of publications, we use two measures: the number of publications weighted by the number of citations and the number of publications in high-impact journals. According to both measures, group affiliation has a positive and significant effect on innovation.

Having found that affiliates systematically innovate more than standalones, we proceed to examine whether the group internal capital market might be a channel through which this effect takes place. Our main empirical strategy is to examine whether the effect of group affiliation systematically varies with exogenous industry conditions that are consistent with the internal capital markets theory. We focus on three industry characteristics: dependence on external funds, investment intensity, and the degree of asymmetric information.

If group affiliation affects innovation by providing cheaper external funding, we would expect this effect to be stronger in industries where firms (for exogenous reasons) invest more and rely more on external funds. We follow Rajan and Zingales (1998) and rank industries by their investment intensity and their dependence on external funding. Consistent with the internal markets hypothesis, we find that the positive effect of business group affiliation on innovation is more pronounced in industries that have higher investment intensity, higher external finance dependence, and higher external equity dependence.

Internal capital markets reduce the cost of asymmetric information between innovating firms and outsiders. Hence, we would expect group affiliation to have a stronger effect on innovation in industries where it is harder for outsiders to learn about the idiosyncratic value of firms. We construct two industry measures to capture this: Productivity Growth Dispersion and average Tobin's Q (Lee (1992), Gompers (1995)). High productivity dispersion within an industry means that firm performance in that industry is affected more by idiosyncratic components than by aggregate shocks. High average Tobin's Q implies that a larger fraction of firm value in that industry is associated with intangible assets, which makes it harder for the outside investor to evaluate those firms. Our findings support the internal capital markets hypothesis. We find that firms that operate in industries with a higher level of asymmetric information (according to both measures) benefit more from group affiliation.

Hoshi et al. (1991) show that the investment of Japanese group-affiliated firms is less sensitive to liquidity than unaffiliated firms because they rely on the group internal capital market. Assuming that affiliated European firms rely on the group internal market to finance their investments in R&D, while standalone firms rely more on their own internal liquidity, we would expect the effect of liquidity on innovation to be stronger for standalone firms. Our findings confirm this prediction: firm's liquidity positively affects innovation in standalone firms, but has no statistically significant effect on affiliated firms.

While we focus our analysis on internal capital markets, other hypotheses may explain the difference in innovation between affiliates and standalones. We discuss several alternative hypotheses including knowledge spillovers, quality of governance, and multinationals. We do not find an alternative hypothesis that is consistent with all of our findings.

The rest of the paper is organized as follows: Section 2 presents the data, Section 3 provides descriptive statistics, Section 4 describes the econometric specification, Section 5 reports the results, Section 6 discusses issues raised by our findings, Section 7 discusses alternative hypotheses, and Section 8 concludes.

2. Data

This paper combines data from four main sources: (1) ownership data on business-groups from Amadeus, (2) information on patents and citations from the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO), (3) academic publication data from Thomson Web of Knowledge, and (4) accounting data from Amadeus and Compustat. In this section, we explain our methodology for constructing the four datasets and describe our sample.

(1) Business groups - following previous literature including Almeida and Wolfenzon (2006), we define a business group as an organizational form in which at least two legally independent firms are controlled by the same ultimate owner. In order to fully characterize groups, we determine group affiliation status for all firms in the Amadeus database. For this purpose, we refer to the Amadeus ownership database, which includes detailed information on direct ownership links between firms in Europe. To ensure all ownership links represent control, we make the following assumptions: for private subsidiaries, we keep only links where the shareholder has at least 50 percent of the voting rights and for public firms, we keep only links where the shareholder has at least 20 percent of the voting rights.[3] These two assumptions leave us with close to one

[3]For reasons of conservatism and simplicity, we define control of a private firm as owning more

million ownership links. In order to infer group structure from these links, we develop an algorithm that constructs the corporate control chains, and then group together firms controlled by the same ultimate owner. Appendix A1 discusses the algorithm in greater detail.

(2) Patents - In order to generate a firm-level measure of innovation, we look at patent based measures which capture technological advances by firms and were shown to be a better measure of research productivity than R&D investments (Griliches (1990) and Trajtenberg (1990)). We constructed a unique novel database of European firm patents by matching all granted patent applications from the EPO and the USPTO to the complete list of Amadeus firms (about 8 million firm names) for the period 1979-2004. The sample of innovating firms used in the paper is the set of firms for which we find a match with the name on a patent applicant record.

In addition to patenting information, we also use patent citations data to measure the quality of patents. Patent quality is highly skewed where only few patents have significant economic value. A common method to proxy for the quality of patents is by counting the number of citations they receive (Trajtenberg (1990) and Hall et al. (2005)). Following this methodology, we compute a citations-weighted patent count where the weights are computed as the ratio between the number of citations patent i receives and the total number of citations made to other patents granted in the same year as patent i.

(3) Academic publications - Another measure of innovation is publication in academic journals (e.g., Cockburn and Henderson (1998)). To measure the publication activity of our sample firms, we develop a dataset on publications in academic journals. The world's largest source of information on academic publications is the Thomson's ISI Web of Knowledge, which includes publication records on hundreds of international journals in 'hard' sciences (such as natural or physical sciences). Each publication has an address field which contains the authors' affiliation. We match all patenting firms by name to the complete ISI database. For each publication, we also have information

than 50% of the firm's voting rights (excluding non-voting shares). Following previous literature on public firms (La Porta et al. (1999), Faccio and Lang (2002) and others), which have a more dispersed ownership, we set the threshold for public firms at 20%. All the results of this paper are robust to different plausible specifications of these thresholds.

on the number of citations, which we use to control for the quality of the publication. With similar data at the journal level, we are able to control for the importance of the journal in which the article was published (this is done by using the impact factor from the Journal Citations Report). We find that 1,094 patenting firms also publish at least one journal article in the period 1979-2004. During the estimation sample (1995-2004), these firms publish on average 0.82 articles per year and 4,852 articles in total. These articles received 43,440 citations from the scientific community.

(4) Accounting - accounting information is taken from Amadeus. The source of the accounting information is the Company Register House in each of the twelve countries included in our sample. The key advantage of these data is their large coverage of firms and unique accounting information on private firms with a wide size distribution. Yet, the accounting data has some limitations. First, countries differ in reporting requirements. For example, very small firms (fewer than 10 employees) in Great Britain are not obliged to disclose accounting information including number of employees, sales, or total assets. On the other hand, French firms must provide such information regardless of their size. Second, firms that do not report accounting information for four consecutive years are dropped from the Amadeus database. Thus, firms that existed in the 2003 sample might be dropped in the 2004 update. In order to capture all the firms that were dropped from the Amadeus sample we purchased older publications of Amadeus and added all firms that appeared in the older publications but were dropped from the 2004 update. We find that about 5 percent of firms are dropped from the Amadeus database each year.

3. First Look at the Data

12,749 firms in Amadeus which report lagged sales also have at least one patent from the European Patent Office between 1979 and 2004. To avoid double-counting of accounting variables, we drop all firms that do not report unconsolidated accounts, which leaves us with 12,389 firms. Panel A of Table 1 provides descriptive statistics on the key firm-level characteristics. On average, a firm in our sample has approximately one patent every two years. Our sample covers firms from a wide size distribution: the median number of employees per firm-year in our sample is 109, the 90th percentile is 982, and the 10th percentile is 7. For sales, the median is about \$16 million, the 90th percentile is \$203 million, and the 10th percentile is only \$0.6 million. Out of the 12,389 firms in our sample, 1,094 firms publish in academic journals. They publish on average 0.81 articles a year.

Panel B of Table 1 presents the group-level characteristics of the innovating firms in our sample. Out of the 12,389 innovating firms, 6,289 firms are affiliated with 4,970 different groups (the remaining firms are standalones). The distribution of the number of firms in each group is highly skewed with a median of 12 firms and an average of 72 firms (the 90th percentile is 200 firms). Group affiliates are spread across different industries - the average HHI of industry concentration is 0.26 with a median of 0.17. Finally, 51 percent of the innovating firms are affiliated with business groups.

Consider, for example, the Nasi-Agnelli group described in Figure 1. The Nasi and Angnelli families control Ifil Finanziaria, the investment company of the group which is located at the apex of the organizational structure. The families control 186 firms through complex ownership chains of up to 9 layers. Overall, the Nasi and Agnelli families control firms with total assets of \$172 billion and annual cash flows (estimated by net income plus depreciation) of \$2 billion. These figures demonstrate the substantial size of the group's internal capital market which can be used to finance R&D investments. Indeed, 8 firms in the group conduct innovation activity and have a total of 795 patents.

Table 2 reports summary statistics separately for affiliates and standalones. On average, an affiliated firm has close to twice as many patents per year compared to standalone (0.57 versus 0.31). The first graph in Figure 2 shows that this difference is consistent across most countries in our sample and is prominent in all major European economies for which we have more observations (e.g. Great Britain, France, and Germany). The second graph shows a similar pattern for the number of citations. The average affiliated firm is larger (677 versus 393 employees and \$211 million versus \$83 million in sales), and 6 years older than the average standalone. Mean comparison tests show that these differences are statistically significant at the 1% level.

4. Econometric Modeling

4.1. Baseline specification

We use the Negative Binomial model to analyze our patent count data. Models for count data assume a first moment of the form:[4]

$$E(P_{it}|X_{it}) = \exp(x'_{it}\beta)$$

where $E(.|.)$ is the conditional expectations operator and P_{it} is a count of the number of patents. We introduce firm fixed effects into the count data model using the 'mean scaling' method of Blundell et al. (1999).[5] This relaxes the strict exogeneity assumption underlying Hausman et al. (1984). Essentially, we exploit the fact that we have a long pre-sample history (of up to 15 years per firm) on patenting activity to construct its pre-sample average. This is then used as an initial condition to proxy for unobserved heterogeneity. The conditional expectation of the estimator is:[6]

$$E(P_{it}|X_{it}) = \exp\{\beta_1 Group_i + \beta_2 \ln Sales_{it-1} + Z'_{it}\beta_4 + \varphi_j + \tau_t + \eta_i\} \tag{4.1}$$

where $Group_i$ is a dummy with value 1 if the firm belongs to a business group and value 0 if the firm is standalone. $Sales_{it-1}$ is used to control for firm size (we use lagged value to mitigate transitory shocks that can affect both the incentive to innovate and sales), Z_{it} is a vector of control such as a complete set of country dummies, φ_j and τ_t are complete sets of three-digit industry SIC and year dummies and η_i is the firm fixed-effect.

Due to the panel structure of our data, we correct the standard errors for serial correlation. This is especially important since the group dummy is constant over time

[4]See Blundell et al. (1999) and Hausman et al. (1984) for discussions of count data models of innovation.

[5]See also Blundell, Griffith, and Windmeijer (2002).

[6]The variance of the Negative Binomial under our specification is:

$$V(P_{it}) = \exp(x'_{it}\beta) + \alpha \exp(2x'_{it}\beta)$$

where the parameter, α, is a measure of 'overdispersion', relaxing the Poisson restriction that the mean equals the variance ($\alpha = 0$).

within firms. Therefore, the reported standard errors are always robust to arbitrary heteroskedasticity and allow for arbitrary serial correlation.

4.2. Dealing with potential biases

4.2.1. Endogeneity bias – 'Winner-picking'

In our baseline specification, we measure the effect of group affiliation on innovation (β_1). If the selection into business groups is endogenous, then our coefficient estimates are likely to be biased. More specifically, if groups engage in winner-picking (i.e., group affiliation is positively correlated with an unobserved 'quality' variable), then β_1 would be upwards biased. We mitigate this potential bias in several ways.

First, since ownership structure is rather consistent over time, we keep in our sample only firms that maintained their affiliation status between 1995 and 2004. We use BvD's mergers and acquisition database - Zephyr, to examine changes in group affiliation status. About five percent of firms in our sample experience a change in their ownership structure – these firms are excluded from our estimation sample.

Second, we compute the average number of patents each firm had prior to our sample period (from 1979 until the first time the firm appeared in our sample). This variable is used as a proxy for the unobserved heterogeneity at the beginning of our sample (Blundell et al. (1999)). Having applied the aforementioned measures, winner-picking would affect our results only in the following case. A standalone firm joins a business group before 1995 (since changes in ownership structure from 1995 onwards were eliminated from our sample). This firm has some unobserved 'quality', which is not captured by the number of its patents up to that time, and still has an effect on the number of patents of that firm 10 years ahead. While we cannot completely discard such a scenario, it seems unlikely to have a significant effect on our results.

Third, we suggest a specific channel through which the group affiliation positively affects innovation, internal capital markets, and test it empirically using interactions with exogenous industry variables. Finally, we mitigate the potential business group selection bias by testing specifications with only affiliated firms. In these specifications, we drop the standalone firms from our sample and divide the affiliated firms into subsamples according to the size of their group. Then we test how the size of the group affects the innovation of affiliated firms.

4.2.2. Unit of observation

In the econometric analysis, we use observations at the firm level. Namely, we compare a group affiliated firm to a standalone firm (and not an entire group to a standalone firm) because they are legally, empirically, and economically comparable. Since this issue may relate to the more general discussion about the boundaries of the firm, we specify the reasons leading us to this choice. First, similar to standalone firms, affiliated firms are legally independent entities. This means that each firm has its own CEO, board of directors, financial statements, etc. Second, groups are highly diversified across industries and are larger by an order of magnitude than standalones (e.g. the average business group has aggregate annual sales of over $3 billion compared to $63 million of standalone). These differences make groups and standalone econometrically and economically incomparable. Finally, we use various observed firm characteristics to control for the remaining differences between affiliates and standalones.

4.2.3. Sample selection – innovating firms

Testing the effect of group affiliation on innovation by focusing only on innovating firms may lead to biased estimators if, for example, the R&D of all firms in a group is conducted by one firm. To deal with this potential sample selection bias, we test different specifications in which we also include non-innovating firms in our sample.[7] We show that both the level of innovation and the decision whether to innovate are positively affected by group affiliation.

4.3. Empirical strategy and definitions of variables

4.3.1. External funds dependence and asymmetric information

Our main empirical strategy is to examine how the effect of group affiliation varies with exogenous industry variables. Specifically, we test whether this variation is consistent with the theoretical predictions of the internal capital markets hypothesis. We focus

[7] The non-innovating firms are the set of firms that did not match to the EPO, UPSTO, or the ISI Web of Knowledge.

mainly on two industry characteristics: dependence on external funds and the degree of asymmetric information.

If group affiliation affects innovation because it provides cheaper external funds, we would expect group affiliation to be more important for innovation in industries that (for exogenous reasons) rely more on external funds. We follow Rajan and Zingales (1998) and rank industries according to their dependence on external funds. In computing measures of external dependence, we use US Compustat firms. As discussed by Rajan and Zingales (1998), using US firms has important advantages: (i) Since the US market is one of the most advanced capital markets in the world, large publicly-traded firms face the least frictions in accessing finance. This means that the amount of external finance used by these companies is likely to be a pure measure of their demand for external finance. (ii) Disclosure requirements imply that data on external financing are comprehensive. (iii) While using US industry data is rather exogenous to European firms, it is likely that an industry's dependence on external funds in the US is a good measure of its dependence in European countries. The only two assumptions needed are that technological differences explain why some industries rely on external funds more than others and that these differences persist across countries. In addition, we face a practical limitation in computing the measures of external dependence from Amadeus since we have no information on capital expenditures for European firms. We compute two measures of external dependence: External Finance Dependence and External Equity Dependence. External Finance Dependence is the ratio between capital expenditures minus cash flow from operations and capital expenditures. External Equity Dependence is the net amount of equity issued over capital expenditures. An additional related measure is Investment Intensity which is the ratio between capital expenditures and net property, plant, and equipment.

Internal capital markets are more beneficial in industries with high level of information asymmetry between innovating firms and outside investors. Therefore, we would expect group affiliation to be more important for innovation in industries in which outsiders find it harder to learn about the idiosyncratic quality of the innovating firm. We construct two measures to capture this: Productivity Growth Dispersion and Average Tobin's Q. Productivity Growth Dispersion is computed as the difference in the three-year average productivity growth between the 90th and 10th percentiles in a specific industry. A high dispersion means that firm performance is affected more by idiosyncratic firm factors and less by aggregate shocks and systematic industry factors. A higher idiosyncratic component would make it more difficult for potential lenders to learn about the quality of an innovating firm by examining aggregate information about the industry in which the firm operates. We compute Productivity Growth Dispersion from the complete set of Amadeus firms for each industry.[8] In this case, we use data from Amadeus and not from Compustat since using Amadeus, productivity growth can be computed for a wide range of firms, including private ones (other industry measures are available only for public firms or not available at all in Amadeus).

Our second measure of asymmetric information is Average Tobin's Q (Lee (1992), Gompers (1995)). A higher value of average Tobin's Q implies that a larger fraction of the firm's value is associated with intangible assets and future growth opportunities. Industries in which the value of firm relies more on intangible assets and future growth opportunity would have a higher degree of information asymmetry. Tobin's Q is computed in the usual way using all Compustat firms for the period 1980-2004. Firm value is the sum of the values of common stock, preferred stock, long-term debt, and short-term debt net of assets. Book value of capital includes net plant, property and equipment, inventories, investments in unconsolidated subsidiaries and intangibles (other than R&D). Tobin's Q is then the ratio between market value and book value of capital.[9]

Our empirical specification now becomes:

$$\begin{aligned} E(P_{it}|X_{it}) \quad &= \quad \exp\{\beta_1 Group_i + \beta_2 Group_i \times Industry\ Measure_j \\ &\quad +\beta_2 \ln Sales_{it-1} + Z'_{it}\beta_4 + \varphi_j + \eta_i + \tau_t\} \end{aligned} \tag{4.2}$$

Our main interest in this section is in the coefficient β_2, where our industry measures are external finance dependence, external equity dependence, investment intensity, productivity growth dispersion, and average Tobin's Q. To reject the null hypothesis and

[8] We also experiment with computing the Productivity Growth Dispersion for each country separately. The pattern of results remains similar.

[9] We winsorize extreme values of Tobin's Q at 0.1 and 20.

show that internal capital markets is the channel through which group affiliation affects innovation, β_2 has to be significantly greater than zero.

Table 3 presents key summary statistics for all our industry measures. All industry measures, except the productivity growth dispersion, are calculated based on Compustat firms from 178 different industries (three-digit SIC level) in the period 1980-2004. Productivity growth dispersion is calculated for the same industries based on the complete set of Amadeus firms in the period 1995-2004. Panel A shows the distribution of the industry variables and Panel B presents the correlation matrix.

Figure 3 shows the difference in the number of patents between affiliates and standalones across our industry measures. It can be seen that, in general, there is a positive relation between our industry measures and the difference scale. While for some measures the relation is almost monotonic (e.g. external equity dependence), for others, most of the effect is in the highest deciles (e.g. Average Tobin's Q).

4.3.2. Liquidity

In the presence of asymmetric information in the capital markets, the availability of internal funds is an important determinant of firms' investment level. Hoshi et al. (1991) show that the investment of Japanese group affiliated firms is less sensitive to their liquidity than that of unaffiliated firms. The authors explain this finding with the existence of a group internal capital market. While an affiliated firm can rely on the group's internal capital market, a standalone firm has no access to such funds and therefore depends more on its own liquidity. Assuming that affiliated European firms rely on the group internal capital market to finance their innovation activity while standalone firms rely more on their own internal liquidity, we would expect the effect of liquidity on innovation to be stronger for standalone firms.

5. Results

5.1. The effect of group affiliation

Table 4 reports the estimation results of the effect of business group affiliation on innovation. In columns 1-5, we use the firm-year number of patents granted by the European Patent Office (EPO) as our dependent variable. In columns 6-10, we also include the number of patents granted by the (USPTO) to the European firms in our sample. In order to avoid double counting of patents on the same invention (granted both in the EPO and the USPTO), we use the Triadic database published by the OECD and exclude USPTO patents that appear in the same family as an EPO patent (where a family is a set of patents that cover the same invention). Our control variables include sales, pre-sample mean number of patents and a complete set of country, industry, and year dummies.

We can see in Table 4 that group affiliation positively affects the number of patents granted by the EPO (column 1) and both the EPO and USPTO (column 6). Evaluated at the sample mean, an affiliate has about 30 percent more patents that a standalone. The positive and significant relation remains when we use the number of firms in the group as our explanatory variable (columns 2 and 7). In order to understand better the relation between the size of the group and innovation, we divide all groups into three categories: small groups (2 or 3 affiliated firms), medium groups (between 4 and 50 affiliated firms), and large groups (more than 50 affiliated firms). We include a dummy variable for each category (columns 3 and 8), using standalones as the baseline category. We find that the effect of group affiliation increases with the size of the group (the coefficients of small, medium and large are 0.103, 0.260 and 0.418 respectively). The effect of group affiliation is not significant for small groups, but it is highly significant for medium and large groups.

Since firm patenting data is persistent over time, periodical shocks can have lasting effects on the dependent variable, which is not be captured by the other regressors. Therefore, we also test dynamic specifications by including lagged values of Patents (columns 4 and 9).[10] The pattern of results is similar to our previous estimations.

[10] In all dynamic specifications, we include a dummy variable for observations where lagged number of patents is zero.

In columns 5 and 10, we include only affiliated firms to test the robustness of our results to group selection (using the small business groups as the baseline category). Consistent with our previous results, we find that affiliation with medium and large groups positively affects innovation and the effect of large groups is more pronounced.

5.1.1. Quality of innovation

We have shown that group affiliation positively affects innovation, as measured by the number of patents. Yet, patents are an imperfect measure of innovation. For example, patenting an invention is a costly and lengthy process that requires professional expertise. Groups may employ a specialized team of patent attorneys and intellectual property experts to advise them on the best patenting strategies, have broader experience, talent management, etc. This means that for a given quality of research, an affiliated firm could issue more patents than a standalone. The prediction of this hypothesis is that the difference between standalones and affiliates would be in the number of patents, but not in the quality of their innovation. In addition, 'soft budget' constraints may lead firms affiliated with large business groups that have abundant resources to invest in low-quality research. This means that the extra resources provided by the group affiliation lead to more innovation being done, but not to an increase in its quality and importance.

We test whether business group affiliation affects the quality of innovation in three ways. First, we use patent citations to measure the quality of patents. Patents that receive more citations are assumed to be of a higher quality (e.g. Trajtenberg (1990)). If group affiliation negatively affects the quality of innovation, we would expect patents by group affiliates to receive a lower number of citations than patents by standalones. Second, we use academic publications by the firm as a measure of the novelty of innovation. We measure the quality of the publication in two ways: by the number of citations it receives and by counting the number of publications only in high-impact journals. Finally, higher quality innovation would have a stronger effect on the productivity of the firm. To test this, we estimate a production function equation and examine whether the effect of patenting by standalones and affiliates of small groups is stronger than that of affiliates of larger groups.

We control for the quality of patents by weighing the number of patents by the citations they receive. Following Trajtenberg (1990), we compute the weights as the ratio between the number of citations received by the patent and the average number of citations received by all patents granted in the same year. On average, a patent in our sample receives 2.44 citations. For patents held by standalones, the average number of citations per patent is 1.06 and for affiliates it is 3.28. Thus, a patent by a group affiliate receives on average 2.21 more citations. The difference is statistically significant at the 1 percent level (with t statistic of 7.68). When examining the difference in citations per patent across firm size, we find a similar pattern. Patents of a firm with fewer than10 employees (5701 patents) receive on average one additional citation when held by an affiliate (significant at the 1 percent level). The same difference holds for firms with fewer than 50 and fewer than100 employees.

In column 1 of Table 5, we find that the coefficient of the business group dummy is positive, highly significant, and of similar magnitude as in our previous estimations. The coefficients of the number of firms (column 2) and the dummies for group size (column 3) are also similar to our previous estimations – group size positively affects innovation and the effect is only significant for medium and large groups.

In columns 4-6 of Table 5, we use academic publications weighted by the number of citations as a measure of innovation. We find that the effect of the number of published articles in academic journals is positively affected by group affiliation. Consistent with the results of our previous estimations, we find that group size positively affects innovation, especially for medium and large groups.[11]

Table 6 reports the effect of group affiliation and patenting on the productivity of the firm. We find that patents positively affect productivity for all firms in our sample (column 1). However, the elasticity of the productivity of the firm with respect to patents stock is 0.067 and only 0.053 for standalones. We convert the elasticity to marginal return to patent stock by multiplying the elasticity by sales over patent

[11] Similar results also hold when we include only publications in high-impact scientific journals (as indicated by the Journal Citations Report (JCR) index).

stock (evaluated at the mean). We find that the marginal patent of affiliated firm contributes on average \$3.2M, while the return of the marginal patent of a standalone is only \$1.5M. The positive effect of group affiliation is even more pronounced when comparing medium and large groups to small groups and standalones (\$3.8M vs. \$1.2M, columns 4-5) or large groups to the rest of the firms (\$5.8M vs. \$1.5M, columns 6-7). When controlling for firm fixed effects ('within firm'), the marginal returns become \$4.3M and \$1.7M for affiliates of large groups and all other firms, respectively.

5.1.2. Sample selection

Thus far, our data have encompassed only innovating firms. Yet, the decision to innovate is endogenous. The ultimate owner of a business group may decide to concentrate all innovation efforts in one affiliated firm or assign all the patents of the group to a single subsidiary. If we include only innovating firms in the estimation sample, this may lead to an upward bias in the coefficient of group affiliation. In addition, if indeed business groups foster innovation, we would expect that group affiliation would not only affect the level of innovation, but would also affect the firm's decision whether to innovate.

Table 7 presents the results of a cross-sectional estimation including a random sample of 10% of the non-innovating firms in the Amadeus database that report employment and sales of over one million dollars. Our results show that group affiliation positively affects the level of innovation and that the effect is increasing in the size of the group (columns 1-4). In the Probit estimation (columns 5-8), we find that group affiliation and group size positively affect the firm's decision of whether to innovate.

The results outlined above show that group affiliation has a positive effect on the scale and novelty of innovation. To better understand the channel through which the effect operates, we turn to investigate in detail the internal capital markets hypothesis.

5.2. Internal capital markets

A possible explanation for the strong effect of group affiliation on innovation might be that groups provide internal capital markets that are especially important in financing risky innovative ventures. We would expect the positive effect of group affiliation on innovation to be more pronounced in industries that rely more on external funds and that are characterized by higher degree of asymmetric information.

Table 8 reports the estimation results for external dependence. Following Rajan and Zingales (1998), we examine the interaction of the group dummy with three variables: External Finance Dependence, External Equity Dependence and Investment intensity. The interaction term of group dummy with each of the three industry variables is positive and highly significant. The same pattern of results holds when we control for various firm and industry characteristics such as liquidity, capital, and competition.

According to the internal capital markets hypothesis, group affiliation has a stronger effect when information asymmetry is higher. To test this hypothesis, we interact our proxies for industry information asymmetry, Productivity Growth Dispersion, and Average Tobin's Q (computed at the three digit SIC level), with a dummy for business group affiliation. We control for potential biases by including the interaction of these two proxies with various firm and industry characteristics. Our results are reported in Table 9 and show a robust positive and significant effect of both interaction terms in all different specifications.

5.2.1. The effect of liquidity

In a study on Japanese business groups (keiretsu), Hoshi et al. (1991) find that investment is less sensitive to liquidity for affiliated firms. They find that the availability of internal funds is an important determinant of investment in a capital market with asymmetric information. In this section, we want to test whether a similar relation holds regarding innovation in European business groups. Namely, we examine whether the effect of liquidity on innovation is weaker for affiliated firms. Our proxy for liquidity is the lagged ratio of cash flows to sales. Since the firm's cash flows may be affected by intra-group loans, we use an income statement measure of cash flows: net income plus depreciation.[12]

Table 10 summarizes the results of our estimation. In column 1 we estimate the pooled effect of liquidity, which is positive and significant. In columns 2-3 and 4-5, we estimate this effect separately for affiliates versus standalones and small groups versus large groups, respectively. The effect of liquidity on innovation is much stronger for standalones and small groups, as expected under the internal capital markets hypothesis. For example, the liquidity coefficient is 0.019 and not significant for affiliates of large groups and 0.115 and significant at the 1 percent level for standalones and small groups.

6. Discussion

Our findings in this paper suggest that business groups foster innovation via internal capital markets. This result is especially interesting in light of recent findings by Seru (2007), who reports that inefficiencies in the internal capital market of US conglomerates stifle innovation. These counter results call for a discussion of three questions. First, why would the business group internal capital market be more efficient for innovation than the conglomerate internal capital market? Second, if indeed the business group internal capital market is more efficient, why are business groups not common in the US? Third, why are not all innovating firms in Europe affiliated with business groups (what is the trade-off)?

6.1. Innovation in conglomerates vs. business groups

Seru (2007) describes the R&D investment decision in a conglomerate. Understanding the source of inefficiencies in a conglomerate can help us explain why the business group internal capital market is more efficient. Seru considers a conglomerate with multiple divisions, where the headquarters centrally allocates the budget across the divisions. The agency problem between the headquarters, who acts to maximize firm

[12] Another reason for using this measure is that private European firms do not file cash flow statements.

value, and the division managers, who seek to maximize their division's budget, leads to inefficient allocation of capital and less innovation. The division managers hold valuable information about the status of the project, but have no incentive to truthfully reveal it to the headquarters, since then resources will be reallocated. Since more novel projects are associated with higher degree of asymmetric information between the headquarters and the division manager, risky projects (above a certain threshold) would not be financed ex ante. The intuition is that the headquarters rejects novel projects since it will not be able to evaluate the project and optimally decide whether to shut it down or continue with its funding.

The agency problem is embedded in the ex-post commitment problem. The headquarters cannot commit not to transfer funds from one division to the other. In contrast to conglomerates, such agency problems do not arise in business groups due to their different structure. Business groups are composed of legally independent companies which often have minority shareholders.

There are two major benefits of the business group structure relative to a conglomerate. First, in a business group, the ultimate owner cannot shift funds from one company to the other at no cost because she would be expropriating the rights of the minority shareholders. The legal boundaries in a business group mitigate the ex-post commitment problem of a conglomerate. The boundaries generate a legal commitment of the ultimate owner not to shift fund in the middle of a project.[13] Conglomerates can try to mimic this structure by decentralizing their budget. Indeed, Seru (2007) reports that conglomerates that decentralized their budget to the divisions are able to mitigate the inefficiencies.

Second, the legal boundaries between different members of the business groups enable the group members to sign enforceable contracts with each other. This means that instead of inefficient reallocation of funds, the business group could utilize its internal capital markets using intra-group lending contracts. For example, a cash-rich company with little investment opportunities (e.g. utility and insurance companies)

[13] Even if in some cases the diversion of funds between two firms in the same group is legal (Johnson et al. (2000)), there is a cost associated with such action (Almeida and Wolfenzon (2006)). For example, there is a waste involved in the diversion when the ultimate owner tries to shift the funds in sophisticated ways so the minority shareholders would not notice it.

could lend funds to other firms in the group which are cash-constrained but have better investment opportunities. Since the two companies are two different legal entities, the contract is enforceable by court.

6.2. Business groups in the US

If business groups are better in facilitating innovation, then why are they not prevalent in the US today? Business groups were common in the US until the 1930's (Morck (2005)), but the American tax reform of the 1930's disentangled business groups. The Roosevelt administration believed that business groups facilitate governance problems, tax avoidance, market power, and dangerously concentrate political influence. In reaction, the American Congress enacted inter-corporate dividend taxes, all but abolished consolidated tax filing for business groups, eliminated capital gains taxes on liquidated controlled subsidiaries, and explicitly banned large pyramidal groups from controlling public utilities companies. The explicit goal of these and other policies was to break up large US business groups.

We explored the current European laws for similar rules regarding pyramidal business groups. We found that the European laws concerning inter-company dividend tax, consolidated tax returns, and capital gains tax are different from those in the US and impose lower costs on business groups. In addition, we did not encounter any explicit restriction on utilities companies in European business group. Hence, legal and tax differences between Europe and the US lead to these differences in ownership structures. Appendix B summarizes these main legal differences within Europe and between Europe and the US.

6.3. What is the cost of business group affiliation?

If business groups facilitate innovation, why are not all innovating European firms affiliated with business groups? What is the cost associated with business group affiliation from the perspective of a standalone firm?

Business group affiliation may be costly for the owners of a standalone firm. Suppose, for example, that an entrepreneur forms a start-up company and that she realizes the potential benefits from business group affiliation for a risky and cash-constrained firm. Assuming that the entrepreneur has private benefits of control (Grossman and Hart (1988), Harris and Raviv (1988)), she faces a trade-off between the benefit of the business group internal capital market and the cost of losing control. In order to join a business group, founders of standalone firms have to sell a significant amount of equity holdings and give up their control. Therefore, for some entrepreneurs, it is optimal not to join a business group.

7. Alternative Hypotheses

In the previous section, we provide robust empirical findings supporting the theoretical prediction that the business group internal capital market positively affects innovation. In this section, we test several alternative hypotheses for our results. While alternative hypotheses may explain some of the differences between the number of patents of affiliated firms and standalones, we find that a significant part of this difference is explained by the existence of a group internal capital market. For example, it is hard to find an alternative hypothesis that would explain why affiliated firms systematically innovate more in industries that rely more on external dependence, are more dispersed and have a higher Tobin's Q ratio.

7.1. Knowledge spillovers

In this paper, we focus on the effect of group internal capital markets on innovation. However, affiliated firms may benefit from knowledge spillovers from research of other firms in the same group ('internal R&D market'). Since the seminal works of Griliches (1979) and Scherer (1982), an expanding literature has documented the positive effect of knowledge spillovers on innovative activity. Assuming that the frequency and magnitude of knowledge spillovers is higher within the same group, this can lead to a positive effect of group affiliation on innovation. To test this hypothesis, we examine the degree of similarity in the research conducted by firms in the same group. Under the knowledge spillovers hypothesis, we would expect firms in the same group to have similar R&D focus. For this purpose, we follow Jaffe (1986) and compute a measure of technological similarity between firm pairs which is based on the extent to which two firms patent in the same technology fields. The research similarity index for each pair of firms (i and j) is computed in the following way:

$$TEC_{ij} = \frac{(T_i T_j^{/})}{(T_i T_i^{/})^{\frac{1}{2}} (T_j T_j^{/})^{\frac{1}{2}}} \tag{7.1}$$

where T is a vector representing the firm's share of patents in the four-digit technology sectors. The technology space information is provided by the allocation of all patents by the EPO into 623 different technology classes (International Patent Classification). We use the average share of patents per firm in each technology class over the period 1979 to 2004 to create the following vector for each firm: $T_i = (T_{i,1}, T_{i,2}, ...T_{i,426})$, where $T_{i,m}$ is the share of patents of firm i in technology class m. TEC_{ij} is then the co-sinus between vector i and j.

Intuitively, the technological similarity is computed as the co-sinus between vector pairs that represent the distribution of firm patenting across four-digit technology fields. A higher similarity of these vectors means the two firms are technologically closer to one another. The average measure of technological similarity between a firm and other firms in its group is 0.024. While it is significantly higher than the average technological similarity between a firm and all firms outside its group (0.014), it is very close to zero and does not indicate that knowledge spillovers are a key determinant of the group's innovation.

7.2. Quality of governance

Affiliated firms may benefit from better governance compared to standalones. For example, all group affiliates have a dominant shareholder (the ultimate owner) who monitors the company's activity - including its innovation activity. However, a standalone firm may have a dispersed ownership and thus, suffer from lack of monitoring and a potential conflict of interests between the manager and the atomic shareholders. Alternatively, the controlling shareholder of a standalone firm may be reluctant to invest in risky R&D activity since a large fraction of his wealth is invested in the company (as opposed to the ultimate owner of a business group whose investment is more diversified). We test this hypothesis by including two control variables in our baseline specification: a block-holder dummy and an ownership concentration variable. Our results are robust to these specifications.

The difference in the scale and novelty of innovation between standalones and affiliates may also be explained by family ownership. Most standalone firms are controlled by an individual or a family. Cadbury (2000) claims that a family CEO focuses more on the long-term compared to an unrelated chief executive. This may imply that family firms invest more in R&D. However, a family CEO may have less expertise in innovation compared to a professional CEO. We test this hypothesis by including a family ownership dummy in our baseline specification. The family variable is negative and insignificant, and the coefficients of the group affiliation variables remain positive and significant.

7.3. Multinationals

Group affiliates are more likely to be part of a multinational organization than standalones. In fact, 95 percent of the multinational firms in our sample are group affiliated. Multinationals are known to be different than domestic firms across various dimensions. For example, Bloom et al. (2007) show that firms affiliated with multinational organizations are more likely to invest in 'soft' innovation such as information communication technologies. In case multinationals are also better in innovating than domiciles, the group dummy may be capturing this effect. To control for this, we include a multinational dummy in our baseline regression. A firm is defined as a multinational if its ultimate owner is from a different country than the firm itself. We do not find a significant effect of multinationals and the group affiliation variables remain robust.

8. Conclusion

This paper uses a novel dataset including ownership structure, financial reports, patenting, and academic publications to study the effect of group affiliation on innovation. Our results indicate that group affiliation has a strong positive effect on innovation. We find that affiliates of large groups patent more than standalones or affiliates of small groups, controlling for size and other observable characteristics. Affiliates of large groups also conduct more novel innovations, as measured by the number of patent citations and academic publications. The patents of large-group affiliates have greater impact on productivity than those of standalones or affiliates of small groups.

Our findings are consistent with the hypothesis that the group internal capital market is a key channel through which group affiliation positively affects innovation. We show that the effect of group affiliation is more pronounced in industries that rely more on external funding and have a higher level of information asymmetry. Moreover, the effect of liquidity on innovation is much weaker for affiliates of large groups than for standalones and affiliates of small groups.

The results presented in this paper contribute to the recent debate on the effect of ownership structure and internal capital market on innovation. Seru (2007) finds that US conglomerates stifle innovation due to agency problems that lead to inefficiencies in their internal capital markets. Morck (2005) claims that conglomerates are more common than business groups in the US as a result of legal barriers to the creation of business groups introduces in the 1940's. We show that in Europe, where such legal barriers do not exist, most innovating firms choose to form business groups and that these groups (especially the larger ones) foster the scale and novelty of corporate innovation. We explain how fundamental differences between the structure of business groups and conglomerates may lead to more efficient internal capital markets in business groups.

Our findings indicate that in an economy with both business groups and standalone firms, affiliates are more innovative than standalones. We do not claim, however, that business groups necessarily increase the aggregate level of innovation in the economy. In fact, business groups may have a negative effect on competition, which would, in turn, lead to a negative effect on the aggregate level of innovation. We leave the analysis of the macro implications of business groups to future research.

References

Almeida, Heitor, and Daniel Wolfenzon. 2006. "A Theory of Pyramidal Ownership and Family Business Groups". *Journal of Finance* 61, 2637-2680.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2007. "Americans do I.T. Better: US Multinationals and the Productivity Miracle". Working paper.

Blundell, Richard, Rachel Griffith, and John Van Reenen. 1999. "Market Shares, Market Value and Innovation in a Panel of British Manufacturing Firms". *Review of Economic Studies* 66, 529-554.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer. 2002. "Individual Effects and Dynamic Feedback in Count Data Models". *Journal of Econometrics* 108, 113-131.

Cadbury, Adrian. 2000. "Family Firms and their Governance: Creating Tomorrow's Company from Today's". London: Egon Zehnder International.

Coase, Ronald H. 1937. "The Nature of the Firm". *Economica* 4, 386-405.

Cockburn, Iain and Rebecca M. Henderson. 1998 "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery". *Journal of Industrial Economics* 46(2), 157-182

Faccio, Mara, and Larry H.P. Lang. 2002. "The Ultimate Ownership of Western European Corporations". *Journal of Financial Economics* 65, 365-395.

Gompers, Paul A. 1995. "Optimal Investment, Monitoring, and the Staging of Venture Capital". *Journal of Finance* 50(5), 1461-1489.

Greenwald, Bruce, Joseph E. Stiglitz, and Andrew Weiss. 1984. "Informational Imperfections in the Capital markets and Macro-Economic Fluctuations". *American Economic Review* 74(1), 194-199.

Griliches, Zvi. 1979. "Issues in Assessing the Contribution of Research and Development to Productivity Growth". *Bell Journal of Economics*, 10(1), 92-116.

Griliches, Zvi. 1990. "Patent Statistics as Economic Indicators: A Survey". *Journal of Economic Literature* 28, 1661-1707.

Griliches, Zvi. 1998. "R&D and Productivity: The Econometric Evidence". The University of Chicago Press.

Grossman, Sanford J., and Oliver D. Hart. 1988. "One Share-One Vote and the Market for Corporate Control". *Journal of Financial Economics* 20, 175-202.

Guedj, Ilan and David S. Scharfstein. 2004. "Organizational Scope and Investment: Evidence from the Drug Development Strategies and Performance of Biopharmaceutical Firms". Working paper.

Guedj, Ilan. 2006. "Ownership vs. Contract: How Vertical Integration Affects Investment Decisions in Pharmaceutical R&D". Working paper.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg. 2005. "Market Value and Patent Citations". *The Rand Journal of Economics* 36, 16-38.

Harris, Milton, and Artur Raviv. 1988. "Corporate Control Contests and Capital Structure". *Journal of Financial Economics* 20, 55-88.

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship". *Econometrica* 52(4), 909-938.

Himmelberg, Charles and Bruce Petersen. 1994. "R&D and Internal Finance: A Panel Study of Small Firms in High-Tech Industries". *Review of Economics and Statistics* 76(1), 38-51.

Hoshi, Takeo, Anil Kashyap and David Scharfstein. 1991. "Corporate Structure, Liquidity, and Investment: Evidence from Japanese Industrial Groups". *The Quarterly Journal of Economics* 106(1), 33-60.

Jaffe, Adam B. 1986. "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value". *American Economic Review* 76(5), 984-1001.

Johnson, Simon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2000. "Tunneling". *American Economic Review* 90, 22-27.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 1999. "Corporate Ownership around the World". *Journal of Finance* 54(2), 471-517.

Lee, Scott D. 1992. "Management buyout proposals and inside information" *Journal of Finance* 47, 1061-79.

Morck, Randall. 2005. "How to Eliminate Pyramidal Business Groups: The Double-Taxation of Intercorporate Dividends and Other Incisive Uses of Tax Policy". *Tax Policy and the Economy* 19, 135-179.

Mullainathan,Sendhil and David Scharfstein. 2001. "Do Firm Boundaries Matter?". *The American Economic Review Papers and Proceedings* 91, 195–199.

Myers, Stewart and Nicholal S. Majluf. 1984. "Corporate Financing and Investment Decisions when Firms have Information that Investors Do Not Have". *Journal of Financial Economics* 13, 187-221.

Rajan, Raghuram and Luigi Zingales. 1998. "Financial Dependence and Growth". *American Economic Review* 88, 559-586.

Scherer, Frederick M. 1982. "Inter-Industry Technology Flows in the United States". *Research Policy* 11, 227-245.

Seru, Amit. 2007. "Do Conglomerates Stifle Innovation?". Working paper, University of Chicago.

Trajtenberg, Manuel. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Innovations". *The Rand Journal of Economics* 21(1), 172-187

Figure 1 - Innovation in the Nasi And Agnelli Business Group

Ultimate Owner

NASI AND AGNELLI FAMILIES

Main activity: Manufacture of motor vehicles

186 firms organized in 9 ownership layers.

Investment Company (apex):

IFIL FINANZIARIA

Group Assets: $172,453M
Group Cash flow: $2,116M

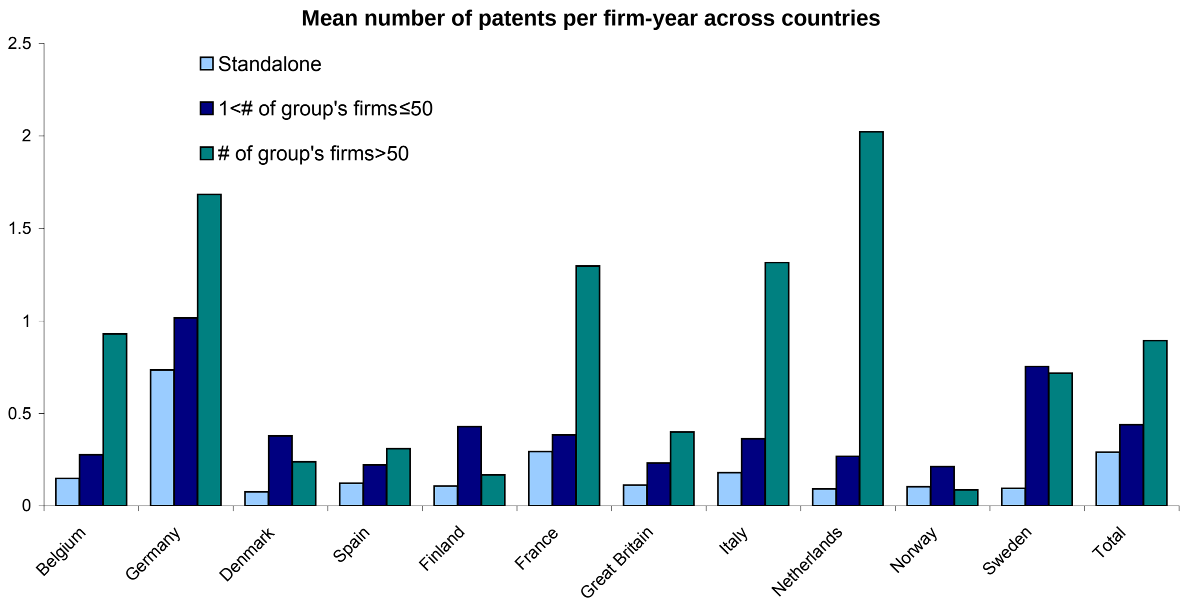
Innovating Firms:

- AUTOMOTIVE LIGHTING (IT) — Assets: $116M, # of patents: 18
- AUTOMOTIVE LIGHTING (DE) — Assets: $354M, # of patents: 11
- MAGNETI MARELLI (FR) — Assets: $20M, # of patents: 89
- SISTEMI SOSPENSIONI (IT) — Assets: $582M, # of patents: 1
- MAGNETI POWER TRIAN (IT) — Assets: $415M, # of patents: 87
- AUTOMOTIVE LIGHTING (FR) — Assets: $55M, # of patents: 4
- FIAT AUTO (IT) — Assets: $12,601M, # of patents: 582
- MAGNETI MARELLI ELEC (FR) — Assets: $227M, # of patents: 3

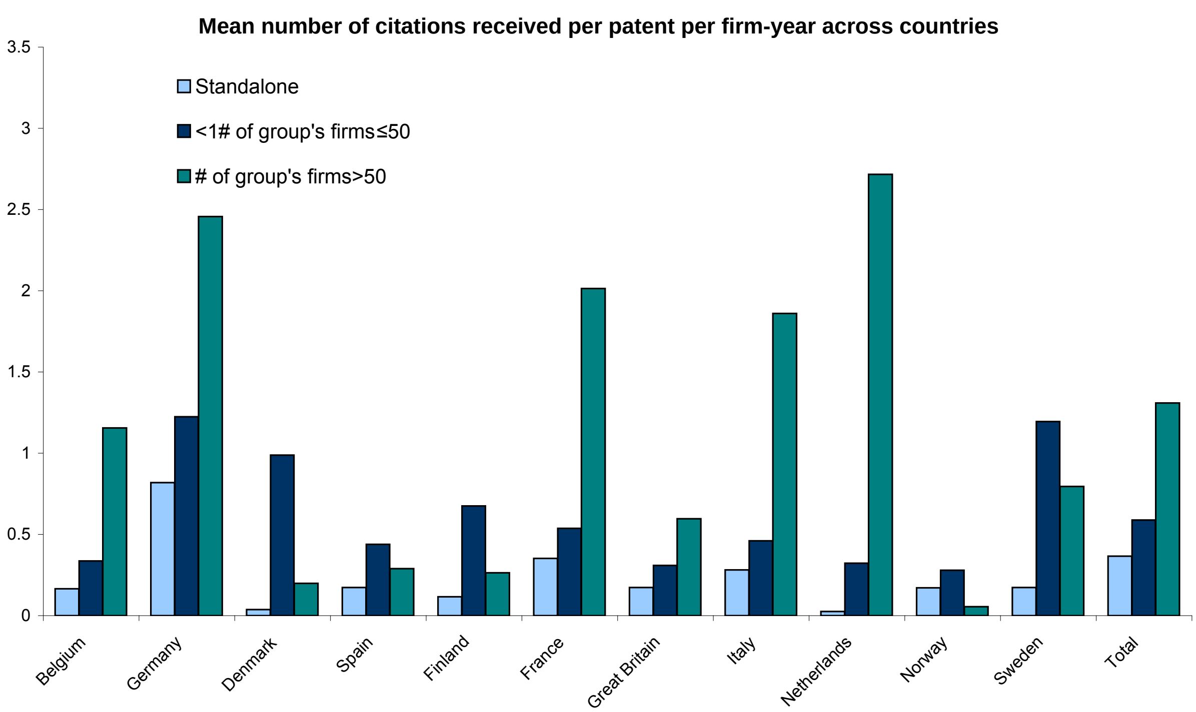
Non-Innovating Firms: 177 subsidiaries

Notes: This figure describes the innovation activity in an Italian business group controlled by the Nasi and Agnelli families. The group structure is based on ownership information from the complete Amadeus database for 2004. The group structure was generated based on the algorithm introduced in section 2. Total assets and cash flows are updated for the last financial statement available in Amadeus (typically 2004). The number of patents is updated for 2005.

Figure 2: Innovation across business group size and countries

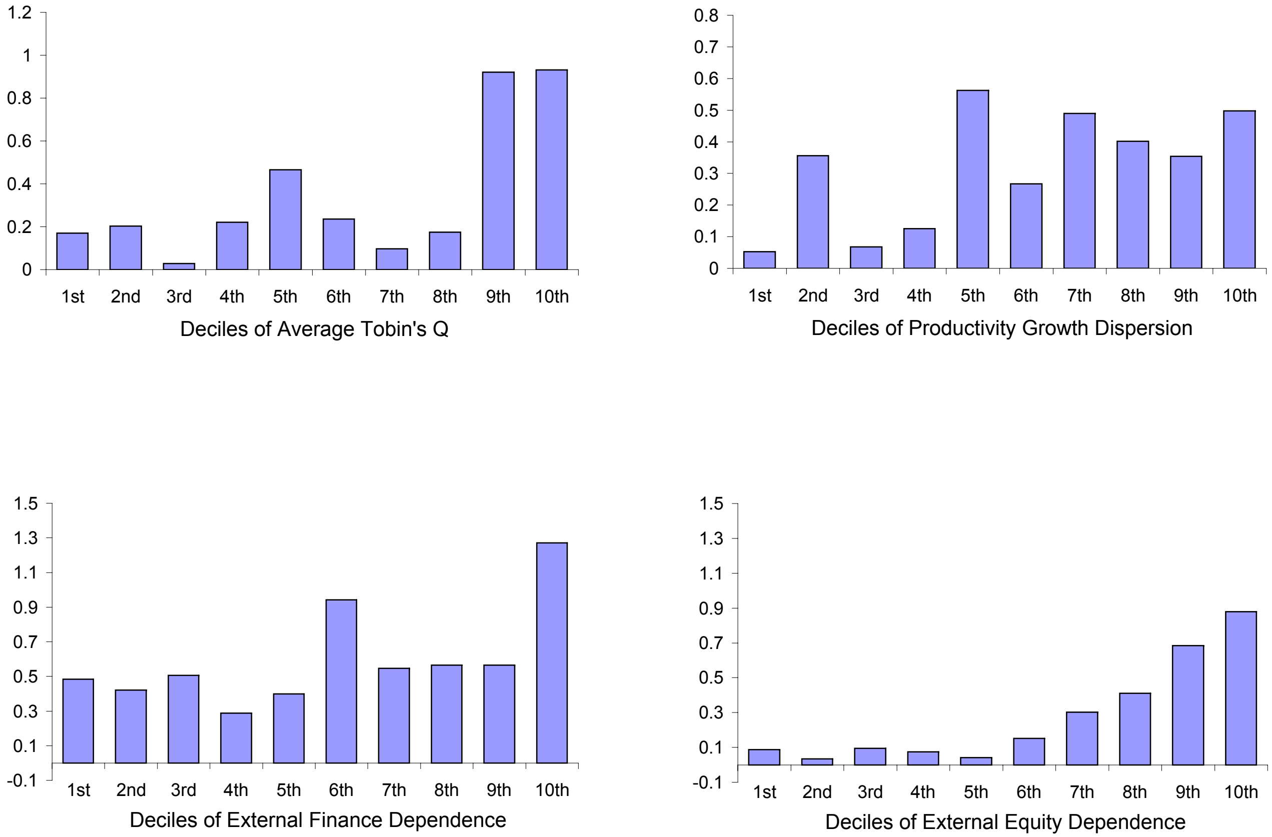


Notes: This figure describes the difference in the mean number of patents between standalone firms, firms that belong to medium-size business groups (groups with les than 50 affiliates), and firms that belong to large groups (groups with more than 50 affiliates) across different European countries from 1995 to 2004. Patents are matched from the European Patent Office. The sample includes only firms with at least one patent between 1979 and 2004.



Notes: This figure describes the difference in the mean number of citations per patent between standalone firms, firms that belong to medium-size business groups (groups with less than 50 affiliates), and firms that belong to large groups (groups with more than 50 affiliates) across different European countries from 1995 to 2004. Patents are matched from the European Patent Office. The sample includes only firms with at least one patent between 1979 and 2004.

Figure 3 - Difference in patenting across industry measures



Notes: This figure describes the difference in the mean number of patents between business group affiliates and standalone firms across deciles of various industry measures. All Variables (except Productivity Growth Dispersion) are computed at the three-digit SIC level based on Compustat firms in the period 1980-2004. Productivity Growth Dispersion (PGD) is computed from the complete Amadeus dataset in the period 1995-2004. External Finance Dependence is defined as the ratio between capital expenditures minus cash flow from operations and capital expenditures. External Equity Dependence is defined as the net amount of equity issued over capital expenditures. PGD is defined as the industry-country mean of the difference in labor productivity growth between the 90th and 10th percentiles. Average Tobin's Q is the industry average of the ration between market value and the book value of capital.

TABLE 1-
SUMMARY STATISTICS FOR INNOVATING FIRMS

PANEL A: FIRM-LEVEL CHARACTERISTICS

Variable	# firms	# Obs	Mean	Std. Dev.	Distribution 10^{th}	50^{th}	90^{th}
# patents EPO (annual)	12,389	60,827	0.44	2.62	0	0	1
Patents stock	12,389	60,827	3.99	18.39	0	1	6.8
# academic publications (annual)	1,094	5,717	0.82	4.49	0	0	2
Academic publicatiosn stock	1,094	5,717	3.52	18.45	0	0.61	5.93
Sales (`000)	12,389	60,827	147,918	1,136,312	555	16,303	202,525
Employess	11,228	48,995	538	3,037	7	109	982
Total Assets (`000)	9,800	54,785	148,413	891,783	543	10,957	188,945
Age	11,093	55,630	26	24	5	19	58
Sales Growth	10,982	48,803	0.09	0.34	-0.24	0.06	0.48
Capital (`000)	9,800	54,785	72,948	570,398	71	2,468	68,751
Capital/Employee (`000)	8,579	43,944	715	45,206	8	36	185
Sales/Employee (`000)	11,228	48,995	437	5,747	80	178	492
Cash Flow ('000)	9,489	44,083	17,563	161,702	63	1,050	18,629

PANEL B: BUSINESS GROUP CHARACTERISTICS

Variable	# firms	# Groups	Mean	Std. Dev.	Distribution 10^{st}	50^{th}	90^{th}
# of affiliates in a group	6,289	4,970	72	164	2	12	200
Sales (millions)	6,289	4,970	12,195	51,450	17	394	18,861
Employees	5,651	4,445	27,233	97,479	110	1,701	62,873
Cash flow (millions)	5,066	3,977	988	4,418	0.6	24	1,705
Patents stock	6,289	4,970	15	128	0.1	1	15
Industry concentration	6,181	4,868	0.26	0.27	0.03	0.17	0.5
Country concentration	6,287	4,970	0.46	0.39	0.05	0.33	1

Notes: These tables provide summary statistics both at the firm level and at the group level for our sample period: 1995-2004. Panel A includes all observations in the estimation sample firms (firms with at least one granted patent application in the period 1979-2004) and provides information on key firm characteristics. Patents data are taken from the European Patent Office (EPO). Academic publications include articles published in "hard" sciences journals by matching the name of the firm to the address field in the complete ISI Web of Science database (which includes about 25 million publications). Patent stock and academic publications stock are computed using the perpetual inventory method using a depreciation rate of 15 percent. Capital is fixed-assets and cash flow is defined as net income plus depreciation. Age is the number of years since the date of incorporation. Panel B provides information on key business group variables which are computed from the complete Amadeus database (about a million firms). Industry concentration is the Herfindahl-Hirschman Index (HHI) of concentration over the number of different three-digit industry SIC in which firms in the group operate. Country concentration is the Herfindahl-Hirschman Index (HHI) of concentration over the number of different countries in the group.

TABLE 2-

SELECTION INTO PATENTING: AFFILIATES VERSUS STANDALONES

PANEL A: GROUP AFFILIATES

	Innovating			Non-innovating		
Variable	# firms	Mean	50^{th}	# firms	Mean	50^{th}
# patents EPO (annual)	6,289	0.57	0	-	-	-
Sales (`000)	6,289	211,021	29,223	24,680	32,416	5,736
Employess	5,741	677	176	24,680	116	28
Total Assets	5,289	205,362	22,599	16,348	32,725	4,277
Age	5,910	29	22	22,699	19	14
Sales Growth	5,552	0.09	0.06	20,940	0.12	0.09
Capital (`000)	5,289	101,071	5,569	16,348	15,504	728
Capital/Employee (`000)	4,690	572	39	16,348	485	21
Sales/Employee (`000)	5,741	496	187	24,680	987	211
Cash Flow (`000)	5,133	18,056	1,429	15,862	3,215	271

PANEL B: STANDALONES

	Innovating			Non-innovating		
Variable	# firms	Mean	50^{th}	# firms	Mean	50^{th}
# patents EPO (annual)	6,100	0.31	0	-	-	-
Sales (`000)	6,100	82,860	6,723	45,268	7,426	2,724
Employess	5,487	393	44	45,268	32	15
Total Assets	4,511	81,642	3,942	36,809	6,242	1,730
Age	5,183	23	17	37,250	17	14
Sales Growth	5,430	0.09	0.06	35,373	0.12	0.09
Capital (`000)	4,511	39,975	853	36,808	2,524	316
Capital/Employee (`000)	3,889	887	33	36,808	141	21
Sales/Employee (`000)	5,487	375	165	45,268	496	206
Cash Flow (`000)	4,356	16,982	222	36,082	401	107

Notes: These tables provide summary statistics for group affiliates (Panel A) and standalones (Panel B) during our sample period: 1995-2004. Each panel divides the firms into innovating and non-innovating. Non-innovating firms are selected randomly as 10 percent of the complete sample of firms in each country that were not matched to the patent data in the EPO or USPTO, but that report employment and have annual sales of at least $1M. Patents data are taken from the European Patent Office (EPO). Capital is fixed-assets and cash flow is defined as net income plus depreciation. Age is the number of years since the date of incorporation.

TABLE 3-

SUMMARY STATISTICS FOR INDUSTRY VARIABLES

PANEL A: DISTRIBUTION OF VARIABLES

Variable	# industries	Mean	Std. Dev.	Distribution 10st	50th	90th
External Finance Dependence	178	1.02	0.71	0.34	0.87	1.88
External Equity Dependence	178	1.21	1.06	0.18	1.00	2.43
Investment Intensity	178	0.66	1.35	0.25	0.45	0.88
Productivity Growth Dispersion	178	0.39	0.16	0.25	0.35	0.64
Average Tobin's Q	178	2.10	1.14	1.04	1.76	3.85
Lerner Index	178	0.86	2.91	0.88	1.03	1.48
R&D/Sales	178	0.44	2.82	0	0.10	0.41

PANEL B: CORRELATIONS MATRIX

	External Finance Dependence	External Equity Dependence	Investment Intensity	Productivity Growth Dispersion	Average Tobin's Q	Lerner Index	R&D/Sales
External Finance Dependence	1.000						
External Equity Dependence	0.823	1.000					
Investment Intensity	0.073	0.196	1.000				
Productivity Growth Dispersion	0.128	0.224	0.218	1.000			
Average Tobin's Q	0.388	0.529	0.021	0.113	1.000		
Lerner Index	0.129	0.150	0.068	0.189	0.023	1.000	
R&D/Sales	0.204	0.179	-0.019	-0.049	0.118	0.011	1.000

Notes: These tables report the summary statistics of the key industry variables. All Variables (except Productivity Growth Dispersion) are computed at the three-digit SIC level based on Compustat firms in the period 1980-2004. Productivity Growth Dispersion (PGD) is computed from the complete Amadeus dataset in the period 1995-2004. External Finance Dependence is defined as the ratio between capital expenditures minus cash flow from operations and capital expenditures. External Equity Dependence is defined as the net amount of equity issued over capital expenditures. Investment Intensity is the ratio between capital expenditures and net property plant and equipment. PGD is defined as the industry-country mean of the difference in labour productivity growth between the 90th and 10th percentiles. Average Tobin's Q is the industry average of the ration between market value and the book value of capital. Panel A describes the distribution of the industry variables and Panel B shows the correlation matrix between the variables.

TABLE 4-
GROUP AFFILIATION AND INNOVATION (NEGATIVE BINOMIAL ESTIMATION)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	# Patents from the EPO					# Patents from the EPO and the USPTO				
Group affiliation:	*All*	*All*	*All*	*All*	*Only affiliates*	*All*	*All*	*All*	*All*	*Only affiliates*
Business Group Dummy	0.256***					0.295***				
	(0.049)					(0.046)				
log(Group's # of firms)		0.097***					0.107***			
		(0.016)					(0.014)			
Small Group Dummy (2≤# of firms≤3)			0.103	0.141***				0.095	0.102**	
			(0.064)	(0.050)				(0.061)	(0.047)	
Medium Group Dummy (4≤# of firms≤50)			0.260***	0.276***	0.163**			0.329***	0.254***	0.208***
			(0.058)	(0.042)	(0.067)			(0.053)	(0.040)	(0.063)
Large Group Dummy (# of firms>50)			0.418***	0.382***	0.303***			0.468***	0.329***	0.321***
			(0.081)	(0.055)	(0.087)			(0.075)	(0.051)	(0.081)
log(Sales)$_{t-1}$	0.349***	0.330***	0.338***	0.239***	0.345***	0.333***	0.314***	0.319***	0.244***	0.349***
	(0.016)	(0.016)	(0.016)	(0.012)	(0.020)	(0.014)	(0.015)	(0.015)	(0.011)	(0.019)
Pre-sample mean	0.018***	0.018***	0.018***	0.010***	0.015***	0.015***	0.015***	0.015***	0.009***	0.012***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
Patents$_{t-1}$				0.101***					0.107***	
				(0.011)					(0.012)	
Country dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Three-digit SIC dummies (178)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (9)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log Psedu Likelihood	-35,546.1	-35,498.5	-35,524.5	-34,300.6	-22,532.9	-49,445.6	-49,378.3	-49,406.8	-47,655.6	-31,764.1
Over-dispersion (Alpha)	3.184	3.155	3.168	2.298	2.298	3.731	3.696	3.711	2.717	2.717
	(0.124)	(0.118)	(0.119)	(0.075)	(0.075)	(0.127)	(0.121)	(0.124)	(0.076)	(0.076)
Observations	60,827	60,827	60,827	60,827	33,798	75,989	75,989	75,989	75,989	42,367
Number of firms	12,389	12,389	12,389	12,389	6,289	15,395	15,395	15,395	15,395	7,906

Notes: This table reports the results of negative binomial regressions that examine the effect of business group affiliation on patents. The dependent variable in columns 1-5 is the firm-year number of patents granted by the European Patent Office. The dependent variable in column 6-10 includes also the number of patents by European firms granted by the United States Patents and Trademarks Office (USPTO). In order to avoid double counting of patents on the same invention (granted both in the EPO and the USPTO), we use the Triadic database published by the OECD and exclude USPTO patents that appear in the same family as an EPO patent (where a family is a set of patents that cover the same invention). Data is for the period 1995 to 2004. A firm belongs to a business group (i.e., business group dummy equals 1) if its ultimate owner controls at least one additional firm. The number of firms in a group includes all firms in the ownership data of Amadeus (even if these firms do not innovate or do not report accounting information). Following the "pre-sample mean scaling approach" of Blundell et al. (1999), our pre-sample fixed-effect is the number of patents a firm had from 1979 until the first year it appeared in our sample. A dummy is included for observations where the pre-sample mean is zero. Country, industry, and year fixed effects are included in all regressions. Standard errors (in brackets) are robust to arbitrary heteroskedasticity and allow for serial correlation through clustering by firm. * significant at 10%; ** significant at 5%; *** significant at 1%.

TABLE 5-
GROUP AFFILIATION AND THE QUALITY OF INNOVATION

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation:	Negative Binomial					
Dependent variable:	Citations-weighted # patents			Citations-weighted # publications		
Business Group Dummy	0.261***			0.574***		
	(0.075)			(0.203)		
log(Group's # of firms)		0.106***			0.194***	
		(0.021)			(0.043)	
Small Group Dummy (2≤# of firms≤3)			-0.059			0.379
			(0.105)			(0.261)
Medium Group Dummy (4≤# of firms≤50)			0.374***			0.467**
			(0.086)			(0.240)
Large Group Dummy (# of firms>50)			0.395***			0.964***
			(0.112)			(0.248)
log(Sales)$_{t-1}$	0.317***	0.298***	0.306***	0.209***	0.178***	0.192***
	(0.020)	(0.021)	(0.021)	(0.037)	(0.037)	(0.040)
Pre-sample mean	0.022***	0.022***	0.022***	1.924***	1.869***	1.887***
	(0.003)	(0.003)	(0.003)	(0.675)	(0.664)	(0.649)
Country dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes
Three-digit SIC dummies (178)	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (9)	Yes	Yes	Yes	Yes	Yes	Yes
Log Psedu Likelihood	-36,154.2	-36,115.6	-36,116.6	-5,808.2	-5,793.2	-5,796.5
Over-dispersion (Alpha)	6.310	6.274	6.279	19.455	19.186	19.260
	(0.240)	(0.236)	(0.236)	(1.907)	(1.879)	(1.877)
Observations	60,827	60,827	60,827	60,827	60,827	60,827
Number of firms	12,389	12,389	12,389	12,389	12,389	12,389

Notes: This table examines the effect of group affiliation on the quality of innovation using two dependent variables: citations-weighted number of EPO patents (columns 1-3), and citations-weighted number of publications (columns 4-6). Academic publications include articles published in "hard" sciences journals by matching the name of the firm to the address field in the complete ISI Web of Science database (which includes about 25 million publications). Data is for the period 1995 to 2004. A firm belongs to a business group (i.e., business group dummy equals 1) if its ultimate owner controls at least one additional firm. The number of firms in a group includes all firms in the ownership data of Amadeus (even if these firms do not innovate or do not report accounting information). Following the "pre-sample mean scaling approach" of Blundell et al. (1999), our pre-sample fixed-effect is the number of patents a firm had from 1979 until the first year it appeared in our sample. A dummy is included for observations where the pre-sample mean is zero. Country, industry, and year fixed effects are included in all regressions. Standard errors (in brackets) are robust to arbitrary heteroskedasticity and allow for serial correlation through clustering by firm. * significant at 10%; ** significant at 5%; *** significant at 1%.

TABLE 6-
GROUP AFFILIATION AND THE QUALITY OF INNOVATION

Dependent variable: log(Sales). OLS estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Group affiliation:	*Only affiliates*	*Only affiliates*	*Only standalone*	*Group's size ≥4*	*Group's size <4*	*Group's size >50*	*Group's size ≤50*	*Group's size >50*	*Group's size ≤50*
log(Group's # of firms)	0.086***								
	0.006								
log(Patents Stock)$_{t-1}$	0.056***	0.067***	0.053***	0.077***	0.044***	0.078***	0.054***	0.058***	0.062***
	(0.011)	(0.013)	(0.019)	(0.015)	(0.016)	(0.019)	(0.014)	(0.025)	(0.021)
Marginal effect of Patents Stock$_{t-1}$	2,466.9	3,175.9	1,540.9	3,835.9	1,217.4	5,841.4	1,495.6	4,343.6	1,717.1
log(Employees)$_{t-1}$	0.726***	0.733***	0.738***	0.712***	0.754***	0.639***	0.746***	0.303***	0.314***
	(0.014)	(0.019)	(0.017)	(0.023)	(0.015)	(0.036)	(0.013)	(0.054)	(0.038)
log(Capital)$_{t-1}$	0.206***	0.216***	0.179***	0.216***	0.186***	0.259***	0.187***	0.146***	0.053**
	(0.010)	(0.014)	(0.013)	(0.016)	(0.012)	(0.025)	(0.009)	(0.035)	(0.027)
Marginal effect of Capital$_{t-1}$	1,799.8	1,320.6	1,960.7	1,158.9	1,886.3	1257.92	1485.26	281.70	492.44
Country dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Three-digit SIC dummies (178)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (9)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed-effects	No	No	No	No	No	No	No	Yes	Yes
R^2	0.861	0.828	0.836	0.818	0.831	0.850	0.844	0.954	0.923
Observations	36,456	22,284	14,172	17,800	18,656	6,706	29,750	6,706	29,750
Number of firms	7890	4,358	3,532	3,421	4,469	1,240	6,650	1,240	6,650

Notes: This table examines the effect of group affiliation on the quality of innovation by analyzing the effect of patents stock on firm productivity. Patents stock is computed using the perpetual inventory method using a depreciation rate of 15 percent. The marginal return to patent stock is computed as the elasticity multiplied by sales over patent stock, evaluated at the mean. A firm belongs to a business group (i.e., business group dummy equals 1) if its ultimate owner controls at least one additional firm. The number of firms in a group includes all firms in the ownership data of Amadeus (even if these firms do not innovate or do not report accounting information). Capital is defined as fixed-assets. Country, industry, and year fixed effects are included in all regressions. Firm fixed-effects are included in columns 8-9. Standard errors (in brackets) are robust to arbitrary heteroskedasticity and allow for serial correlation through clustering by firm. * significant at 10%; ** significant at 5%; *** significant at 1%.

TABLE 7-
GROUP AFFILIATION AND INNOVATION: CROSS-SECTIONAL INCLUDING NON-INNOVATING FIRMS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	# Patents from the EPO (Negative Binomial)				Dummy for innovating (Probit)			
Group affiliation:	*All*	*All*	*All*	*Only affiliates*	*All*	*All*	*All*	*Only affiliates*
Business Group Dummy	0.512***				0.306***			
	(0.113)				(0.020)			
log(Group's # of firms)			0.170***	0.103**			0.107***	0.055***
			(0.030)	(0.035)			(0.007)	(0.009)
Small Group Dummy (2≤# of firms≤3)		0.182				0.157***		
		(0.146)				(0.027)		
Medium Group Dummy (4≤# of firms≤50)		0.679***				0.356***		
		(0.131)				(0.024)		
Large Group Dummy (# of firms>50)		0.714***				0.549***		
		(0.166)				(0.038)		
$\log(\text{Sales})_{t-1}$	0.674***	0.649***	0.637***	0.702***	0.078***	0.063***	0.065***	0.181***
	(0.041)	(0.044)	(0.044)	(0.049)	(0.007)	(0.007)	(0.007)	(0.010)
Country dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Three-digit SIC dummies (178)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log Psedu Likelihood	-3,477.9	-3,471.1	-3,469.9	-2,368.4	-14,851.3	-14,789.3	-14,803.9	-6875.5
Over-dispersion (Alpha)	6.835	6.794	6.822	5.981	-	-	-	-
	(0.783)	(0.779)	(0.785)	(0.692)				
Observations	55,597	55,597	55,597	21,760	55,597	55,597	55,597	21,760

Notes: This table reports the results of negative binomial (columns 1-4) and Probit (columns 5-8) regressions that examine the effect of business group affiliation on patents for a large sample including also non-innovating firms. Non-innovating firms are selected randomly as 10 percent of the complete sample of firms in each country that were not matched to the patent data in the EPO or USPTO, but that report employment and have annual sales of at least $1M. The estimation sample is cross-sectional for the year 2003. The dependent variable in columns 1-4 is the firm-year number of patents granted by the European Patent Office. The dependent variable in columns 5-8 is a dummy that receives the value of 1 for innovating firm and zero otherwise. A firm belongs to a business group (i.e., business group dummy equals 1) if its ultimate owner controls at least one additional firm. The number of firms in a group includes all firms in the ownership data of Amadeus (even if these firms do not innovate or do not report accounting information). Country and industry fixed effects are included in all regressions. Standard errors (in brackets) are robust to arbitrary heteroskedasticity and allow for serial correlation through clustering by ultimate owner. * significant at 10%; ** significant at 5%; *** significant at 1%.

TABLE 8-

BUSINESS GROUPS AND INNOVATION: THE EFFECTS OF INDUSTRY EXTERNAL DEPENDENCE AND INVESTMENT INTENSITY

DEPENDENT VARIABLE: NUMBER OF PATENTS (FIRM-YEAR)	External Finance Dependence			External Equity Dependence			Investment Intensiveness		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dummy for business group	0.075	0.325	0.199	0.126*	-0.326	-0.924	0.223***	0.228	0.244
	(0.091)	(0.941)	(0.956)	(0.073)	(0.982)	(1.063)	(0.006)	(0.961)	(1.031)
Business Group Dummy × External Dependence	0.161***	0.139***	0.151***	0.118***	0.136***	0.162***	0.127**	0.127**	0.103**
	(0.060)	(0.051)	(0.057)	(0.043)	(0.042)	(0.046)	(0.065)	(0.066)	(0.056)
Contros for other industry characteristics									
Business Group Dummy × Industry R&D/Sales		-0.009	-0.019		-0.009	-0.023**		0.003	-0.008
		(0.012)	(0.022)		(0.012)	(0.011)		(0.012)	(0.012)
Business Group Dummy × Industry Lerner Index		-0.194	-0.165		0.464	0.977		0.002	
		(0.968)	(0.986)		(0.999)	(1.084)		(0.985)	-0.068 (1.061)
Contros for other firm characteristics									
$\log(\text{Sales})_{t-1}$ × External Dependence	-0.017	-0.016*	-0.019	-0.017**	-0.017**	0.002	-0.024**	-0.025***	0.001
	(0.011)	(0.009)	(0.022)	(0.008)	(0.008)	(0.012)	(0.008)	(0.009)	(0.021)
$\log(\text{Cash Flow})_{t-1}$ × External Dependence			0.072**			-0.014			-0.010
			(0.029)			(0.015)			(0.011)
$\log(\text{Capital})_{t-1}$ × External Dependence			-0.009			-0.013			-0.011
			(0.017)			(0.015)			(0.014)
log(Sales)t-1	0.355***	0.357***	0.122***	0.268***	0.268***	0.060**	0.024**	0.258***	0.063**
	(0.019)	(0.019)	(0.036)	(0.019)	(0.019)	(0.028)	(0.009)	(0.015)	(0.027)
log(Cash Flow)t-1			0.072**			0.129***			0.128***
			(0.029)			(0.028)			(0.018)
log(Capital)t-1			0.069**			0.075***			0.065***
			(0.029)			(0.027)			(0.019)
Pre-sample fixed-effect	0.017***	0.009***	0.004***	0.009***	0.009***	0.003***	0.009***	0.009***	0.004***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Patents_{t-1}	0.102***	0.104***	0.118***	0.104***	0.104***	0.113***	0.103***	0.103***	0.112***
	(0.010)	(0.011)	(0.013)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.012)
Country dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Three-digit SIC dummies (110)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (9)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log Psedu Likelihood	-29,018.1	-27,477.9	-14,819.0	-28,023.6	-27,474.4	-14,927.0	-28,027.9	-27,478.4	-14,934.7
Over-dispersion (Alpha)	3.043	2.203	1.323	2.185	2.202	1.305	2.185	2.200	1.314
	(0.139)	(0.086)	(0.066)	(0.085)	(0.087)	(0.065)	(0.085)	(0.086)	(0.065)
Observations	60,827	60,827	29,982	60,827	60,827	29,982	60,827	60,827	29,982
Number of firms	12,389	12,389	7,315	12,389	12,389	7,315	12,389	12,389	7,315

Notes: This table reports the results of negative binomial regressions that examine the effect of external dependence and investment intensity on patents. The dependent variable is the firm-year number of patents granted by the European Patent Office for the period 1995 to 2004. A firm belongs to a business group if its ultimate owner controls at least one additional firm. External Finance Dependence, External Equity Dependence, and Investment Intensity are computed as the average three-digit SIC level for the period 1980-2004 based on Compustat firms. External Finance Dependence (columns 1-3) is defined as the ratio between capital expenditures minus cash flow from operations and capital expenditures. External Equity Dependence (columns 4-6) is defined as the net amount of equity issued over capital expenditures. Investment Intensity (columns 7-9) is the ratio between capital expenditures and net property plant and equipment. Lerner index is defined as one minus the average industry profit margin (i.e., a higher Lerner index means stronger competition), based on the complete sample of Amadeus firms. Pre-sample fixed-effect is the number of patents a firm had from 1979 until the first year it appeared in our sample. A dummy is included for observations where the pre-sample mean is zero. Cash flow is defined as net income plus depreciation and capital is defined as fixed assets. Country, industry, and year fixed effects are included in all regressions. Standard errors (in brackets) are robust to arbitrary heteroskedasticity and allow for serial correlation through clustering by firm. * significant at 10%; ** significant at 5%; *** significant at 1%.

TABLE 9-

BUSINESS GROUPS AND INNOVATION: THE EFFECTS OF INDUSTRY ASYMETRIC INFORMATION

DEPENDENT VARIABLE: NUMBER OF PATENTS (FIRM-YEAR)						
	Productivity Growth Dispersion			Average Tobin's Q		
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy for business group	0.008	-0.250	-0.585	0.097	-0.199	-0.505
	(0.118)	(1.249)	(1.260)	(0.095)	(2.000)	(2.134)
Business Group Dummy × Industry Asymetric Information	0.788***	0.850***	0.932**	0.087**	0.097***	0.091**
	(0.296)	(0.379)	(0.408)	(0.037)	(0.036)	(0.041)
Contros for other industry characteristics						
Business Group Dummy × Industry R&D/Sales		0.004	-0.006		-0.005	-0.015
		(0.011)	(0.012)		(0.012)	(0.012)
Business Group Dummy × Industry Lerner Index		0.239	0.475		0.289	0.512
		(1.196)	(0.120)		(2.031)	(2.172)
Contros for other firm characteristics						
$\log(\text{Sales})_{t-1}$ × Industry Asymetric Information	-0.307***	-0.306***	-0.107	-0.022***	-0.021***	-0.011
	(0.071)	(0.071)	(0.135)	(0.007)	(0.007)	(0.170)
$\log(\text{Cash Flow})_{t-1}$ × Industry Asymetric Information			-0.037			-0.003
			(0.116)			(0.014)
$\log(\text{Capital})_{t-1}$ × Industry Asymetric Information			0.037			-0.001
			(0.116)			(0.013)
log(Sales)t-1	0.372***	0.372***	0.139**	0.322***	0.321***	0.149***
	(0.031)	(0.031)	(0.062)	(0.025)	(0.025)	(0.050)
log(Cash Flow)t-1			0.176***			0.103***
			(0.047)			(0.038)
log(Capital)t-1			0.044			0.056
			(0.049)			(0.036)
Pre-sample fixed-effect	0.009***	0.009***	0.004***	0.008***	0.008***	0.003**
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.014)
Patents_{t-1}	0.103***	0.102***	0.118***	0.103***	0.103***	0.120***
	(0.011)	(0.011)	(0.013)	(0.011)	(0.011)	(0.013)
Country dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes
Three-digit SIC dummies (110)	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (9)	Yes	Yes	Yes	Yes	Yes	Yes
Log Psedu Likelihood	-27,458.2	-27,457.9	-14,812.7	-23,444.0	-23,303.9	-12,686.9
Over-dispersion (Alpha)	2.189	2.189	1.321	2.028	2.042	1.277
	(0.087)	(0.087)	(0.065)	(0.082)	(0.083)	(0.068)
Observations	60,827	60,827	29,982	60,827	60,827	29,982
Number of firms	12,389	12,389	7,315	12,389	12,389	7,315

Notes: This table reports the results of negative binomial regressions that examine the effect of asymmetric information on patents. The dependent variable is the firm-year number of patents granted by the European Patent Office for the period 1995 to 2004. A firm belongs to a business group if its ultimate owner controls at least one additional firm. Our proxies for asymmetric information are Productivity Growth Dispersion (columns 1-3) and average Tobin's Q (columns 4-6). Productivity Growth Dispersion is defined as the industry difference between the 90 and 10 percentiles of productivity growth averaged over the period 1995-2004, based on the complete sample of Amadeus firms. Average Tobin's Q is the industry average of the ratio between the value of the firm and the book value of its tangible assets, based on US Compustat firms. Lerner index is defined as one minus the average industry profit margin (i.e., a higher Lerner index means stronger competition), based on the complete sample of Amadeus firms. Pre-sample fixed-effect is the number of patents a firm had from 1979 until the first year it appeared in our sample. A dummy is included for observations where the pre-sample mean is zero. Cash flow is defined as net income plus depreciation and capital is defined as fixed assets. Country, industry, and year fixed effects are included in all regressions. Standard errors (in brackets) are robust to arbitrary heteroskedasticity and allow for serial correlation through clustering by firm. * significant at 10%; ** significant at 5%; *** significant at 1%.

TABLE 10-
THE EFFECT OF LIQUIDITY

Dependent variable: # Patents from the EPO. Negative Binomial estimation							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Group affiliation/size:	*All*	*Only affiliates*	*Only standalone*	*Group's size ≥4*	*Group's size <4*	*Group's size ≥4*	*Group's size <4*
Innovation:	*Innovating*	*Innovating*	*Innovating*	*Innovating*	*Innovating*	*Also non-innovating*	*Also non-innovating*
log(Cash Flow)$_{t-1}$	0.021***	0.017*	0.045***	0.019	0.115***	0.124	0.254***
	(0.008)	(0.009)	(0.016)	(0.027)	(0.023)	(0.082)	(0.084)
log(Capital)$_{t-1}$	0.078***	0.076***	0.081***	0.069***	0.096***	0.309***	0.403***
	(0.015)	(0.019)	(0.025)	(0.026)	(0.023)	(0.072)	(0.083)
Business Group Dummy	0.219***						
	(0.046)						
log(Sales)$_{t-1}$	0.128***	0.126***	0.149***	0.150***	0.059**	0.267***	0.031
	(0.020)	(0.026)	(0.032)	(0.035)	(0.027)	(0.106)	(0.129)
Pre-sample mean	0.004***	0.004***	0.002	0.003***	0.003		
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)		
Patents$_{t-1}$	0.117***	0.105***	0.124***	0.102***	0.133***		
	(0.012)	(0.014)	(0.022)	(0.014)	(0.024)		
Country dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Three-digit SIC dummies (178)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (9)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log Psedu Likelihood	-15,102.5	-10,208.4	-4,985.1	-8,290.2	-6,749.7	-950.2	-750.9
Over-dispersion (Alpha)	1.310	1.294	3.287	1.294	1.123	3.717	3.330
	(0.065)	(0.073)	(0.329)	(0.081)	(0.096)	(0.555)	(0.958)
Observations	30,138	17,661	12,825	13,818	16,766	5,486	19,930
Number of firms	7,308	4,119	3,295	3,235	4,255	5,486	19,930

Notes: This table reports the results of negative binomial regressions that examine the effect of liquidity on innovation. The dependent variable is the firm-year number of patents granted by the European Patent Office. Data is for the period 1995 to 2004. Non-innovating firms are selected randomly as 10 percent of the complete sample of firms in each country that were not matched to the patent data in the EPO or USPTO, but that report employment and have annual sales of at least $1M. Cash flow is defined as net income plus depreciation and capital is defined as fixed assets. A firm belongs to a business group (i.e., business group dummy equals 1) if its ultimate owner controls at least one additional firm. Following the “pre-sample mean scaling approach” of Blundell et al. (1999), our pre-sample fixed-effect (columns 1-5) is the number of patents a firm had from 1979 until the first year it appeared in our sample. A dummy is included for observations where the pre-sample mean is zero. Country, industry, and year fixed effects are included in all regressions. Standard errors (in brackets) are robust to arbitrary heteroskedasticity and allow for serial correlation through clustering by firm. * significant at 10%; ** significant at 5%; *** significant at 1%.

A. Appendix

A.1. Determining business group affiliation

This section details the construction and output of our newly developed algorithm. The purpose of the algorithm is to determine the structure of European business groups based on the Amadeus ownership database. The algorithm consists of two parts: a control-chain generator that constructs the ownership and control links between different European firms, and a name matching procedure that groups together firms controlled by the same ultimate owner.

A.1.1. Control chain generator

Objective The Amadeus ownership database includes detailed information on the percentage of ownership between European corporate shareholders and their European subsidiaries. The data span virtually all European countries (including Eastern Europe). We develop an ownership algorithm that constructs the internal structure of business groups based on these inter-company ownership links. The main benefits of the algorithm are: (i) it constructs the ownership chains without relying on the (often missing) information on whether an ownership link is direct or indirect,[14] (ii) it completes missing ownership links by transitivity, (iii) it identifies cross-holdings, and (iv) it handles complex ownership structures. These features allow us to develop robust measures of business group characteristics (such as group size).

Input We include all ownership links from the Amadeus ownership database that represent a control relation. For this, we make the following assumptions: for private subsidiaries, a shareholder exerts control if its direct percentage of ownership is larger than 50. For public firms, the percentage of direct ownership has to be larger than 20 to represent a control relation (since ownership is typically less concentrated in public firms than in private firms).[15]

There are 843,390 direct ownership links that satisfy our control assumptions, where 406,379 shareholders control 843,124 subsidiaries. The average percentage of direct ownership is 94.6 with a median of 100 (77 percent of the ownership links represent a wholly-owned relation). There are 2,484 public subsidiaries. For these subsidiaries, the average percentage of direct ownership is 53.6 with a median of 49 (only 1 percent of these links represent a wholly-owned relation).

Description of the algorithm The algorithm follows three steps: (i) completes missing ownership links, (ii) generates lists of all subsidiaries and parents for each

[14] Indirect ownership links are very common in our data. Suppose, for example, that firm A owns 60% of the shares of B and that B owns 60% of the shares of C. In this case, firm A has a *direct ownership* of 60% in B and an *indirect ownership* of 36% in C.

[15] Similar assumption was made by La Porta et al. (1999) and Faccio and Lang (2002).

company, and (iii) constructs the ownership chains bottom-up.[16] To illustrate our methodology, it would be useful to consider the following example. Suppose Figure A.1 correctly describes the ownership structure of a business group. The ultimate owner (for example, a family) at the apex of the group controls 7 public and private firms. Amadeus provides detailed data on direct ownership links. Thus, our raw data include the links $A \rightarrow D$, $B \rightarrow F$, $C \rightarrow G$, and $D \rightarrow E$. Note that the percentage of ownership for the link $C \rightarrow G$ has to be larger than 20 (because firm G is public), where for the percentage of ownership for all other links has to be larger than 50 (because the other subsidiaries are private). Because there is no information about indirect ownership links, the link $A \rightarrow E$ is missing from the raw data. The first step of the algorithm is to complete missing links. As we observe the ownership relations $A \rightarrow D$ and $D \rightarrow E$, our algorithm infers the ownership relation $A \rightarrow E$. Note that at this stage of the algorithm we still do not know whether the ownership relation is direct or indirect (and if it is indirect, how many layers separate firm E from firm A). The second step of the algorithm is to construct two lists for each firm: shareholders and subsidiaries. This step saves valuable running time, which is especially important when dealing with large scale ownership data. The following table is generated:

Firm	Shareholder	Subsidiary
A	-	D, E
B	-	F
C	-	G
D	A	E
E	A, D	-
F	B	-
G	C	-

Note that from step 1, we already know that firm A is a shareholder of firm E. Also, because we assume the ultimate owner is a family, firms A, B, and C have no corporate (European) shareholder. The third and final step of the algorithm is to construct the structure of the group based on the above ownership relations. Because of the missing links problem, our algorithm does not assume that an ownership relation is direct; the only input the algorithm receives is the existence of the ownership relation. We start with a firm that has no subsidiaries from the list generated in step 2. We illustrate the procedure for firm E, which is the most interesting in this example. Firm E is placed at the bottom of the ownership chain. Next, we move to the shareholder list of firm E. It includes firms A and D. Starting arbitrary with A, place A above E. Proceeding to firm D, there are three possibilities for its location: (i) D is above E and above A; (ii) D is above E, but below A; (iii) D is above E, but not below neither above A (different ownership chain). For (i) to be the right structure, D has to appear in the shareholder list of firm A. From step 2, we rule this out. For (ii) to be the right

[16] Unlike business groups in East Asia (such as the Japanese keiretsu), most European business groups are organized in pyramids (Figure A.1). This means that interlocking shareholdings are not common and, therefore, ownership chains can be constructed bottom-up.

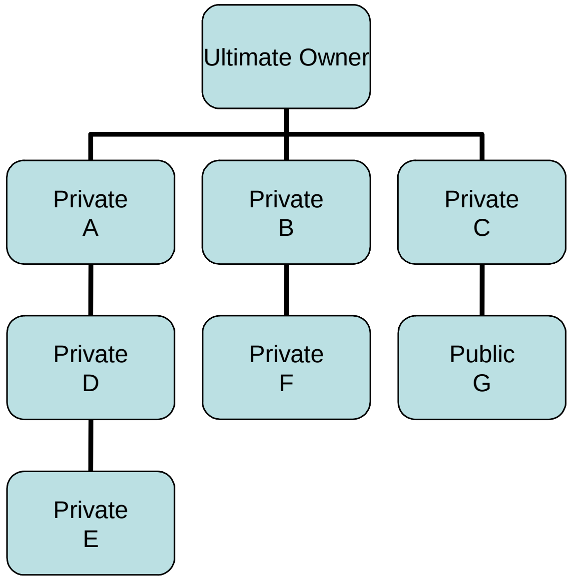


Figure A.1: Example of a business group.

structure, D has to appear on the subsidiary list of firm A. From step 2, this holds. Finally, for (iii) to be the right structure, A cannot appear on either the shareholder or subsidiary lists of firm D. From step 2, this is ruled out. At the end of this procedure, we have determined for each ownership chain the highest shareholder firm - we call this firm the leading shareholder.

Our algorithm fails in the case of cross-holdings. A cross-holding is an ownership structure where a shareholder is also a subsidiary of its own subsidiary. For example, suppose we also observe the ownership link $E \rightarrow A$. Our ordering procedure will not work because there is no starting point: no firm is placed at the bottom of the business group and, therefore, the leading shareholder cannot be determined.[17] Yet, we observe only few cases of cross-holdings in the data (0.5 percent of the ownership links are associated with at least one cross-holding).

Output We start with 843,390 direct ownership links that satisfy our control assumptions. After the algorithm completes all the missing links, we end up with 1,642,379 ownership links - almost double the ownership links we started with. Based on the complete set of ownership links, our algorithm extracts 769,725 ownership chains. Only 4,141 ownership chains are associated with a cross-holding. We drop these chains because for them, we cannot determine the exact structure of the ownership chain. The average ownership chain includes 3.8 firms with a median of 3 firms (25 percent of the

[17] A less 'severe' case of cross-holding is where we observe $E \rightarrow D$. In this case, our algorithm constructs two ownership chains: $A \rightarrow D \rightarrow E$ and $A \rightarrow D \rightarrow E$, where both correctly characterize the ownership structure. The leading shareholder is firm A in both cases, which allows us to correctly group firms into business groups.

ownership chains have 5 firms and more, where the maximum number of firms in a chain is 16). 330,098 firms are located at the top of the chain (leading shareholders). On average, a leading shareholder owns 87 percent of the firm it controls (directly or indirectly), where the median is 100 percent and the minimum is 0.1 percent.[18]

A.1.2. Ultimate owner name matching

The second part of our algorithm groups all firms across ownership chains based on the name of the ultimate owner of each leading shareholder (i.e., the firm at the top of each control chain). The raw data from the Amadeus ownership database contain strings describing the ultimate-owner names. The names vary in their patterns, for example: "JOHN SMITH", "JOHN F SMITH", "SMITH AND SONS", "THE SMITH FAMILY", "PROF. VAR DER SMITH", etc.

The name matching process deals with three main issues. First, ultimate owner names are not standardized, i.e., the same name can be spelled differently across subsidiaries. Second, common names may lead to 'over-grouping'. Third, for wealthy families, we frequently observe that different members control different leading shareholders. Thus, we have to determine whether to group firms at the family level or at the individual level.

We deal with theses issues as follows. First, we develop a name standardization procedure that harmonizes the different string patterns in our data. For example, "PROF. JOHN SMITH AND SONS" becomes "JOHN SMITH". Second, we search for publicly available information on each of our largest 500 business groups. When we cannot verify from public sources (such as Forbes and The Economist) that a given family is indeed wealthy, we check for name commonality. We compute the frequency of the appearance of the name in the ultimate owner population. In case this frequency is higher than the median frequency, we assume the common name problem and do not include that ultimate owner in our sample. Third, we group firms to business groups at the family level (for example, De Rotchild family, Nasi-Agnelli family). As a robustness test, we check the sensitivity of our findings to grouping at the individual level and find that our results are robust to this alternative grouping.

This name matching procedure leaves us with 1,736,034 standardize family names which we then match to the leading shareholders identified by the ownership algorithm (369,800 names are actually matched to leading shareholders). Finally, we add firms without subsidiaries that are controlled by the same ultimate owner.

A.1.3. Summary

We identify 581,108 business groups. The average business group has 33 firms with a median of 4 (the largest business group has 1882 firms). About 35 percent of business

[18]When the ultimate owner is a European firm, we can actually check the output of our algorithm with that of Amadeus. In 96 percent of the cases where Amadeus reports that an ultimate owner is a widely-held European firm, it is also the leading shareholder found by our algorithm. We attribute the 4 percent difference in ultimate owner assignments to the different control definitions assumed by Amadeus.

groups (211,308) are widely-held (the leading shareholder is a widely-held European firm), where the remaining business groups have a dominant shareholder (family, non-European firm, or state). Amadeus indicates whether an ultimate owner is one of the following types: an individual, a financial company, an industrial company or the state. For business groups where the ultimate owner is a European firm, we determine whether this ultimate owner is a financial or industrial firms based on the industry location of the apex firm. The number of firms in a business group varies substantially across ultimate owner types. For family ultimate owners, the average business group has 27 firms. This number rises to 97 firms when the ultimate owner is a financial institution, to 46 firms when the ultimate owner is a widely-held corporation and to 521 when the ultimate owner is the state.

A.2. Matching patent data

A.2.1. European Patent Office (EPO)

The matching between EPO patent applicants and Amadeus firms has been a collaborative project with the Institute for Fiscal Studies (IFS) and the Centre for Economic Performance (CEP).[19] This section is a brief summary of the matching procedure described in the CEP/IFS AmaPat document and is included here for completeness.

Our main information source on patents is the April 2004 publication of the PATSTAT database, which is the standard source for European patent data. This database contains all bibliographic data (including citations) on all European patent applications and granted patents, from the beginning of the EPO system in 1979 to 2004.

We match the name of each EPO applicant listed on the patent document to the full name of a firm listed in Amadeus (about 8 million names). Since we are interested only in matching patent applicants to firms, we exclude applicant names that fall into the following categories: government agencies, universities, and individuals. We identify government agencies and universities by searching for a set of identifying strings in their name. We identify individuals as patents where the assignee and the inventor name strings are identical.

The matching procedure follows two main steps. (i) Standardizing names of patent applicants. This involves replacing commonly used strings which symbolize the same thing, for example "Ltd." and "Limited" in the UK.[20] We remove spaces between characters and transform all letters to capital letters. As an example, the name "British Nuclear Fuels Public Limited Company" becomes "BRITISHNUCLEARFUELSPLC". (ii) Name matching: match the standard names of the patent applicants with Amadeus firms. If there is no match, then try to match to the old firm name available in Amadeus. We need to confront a number of issues. First, in any given year, the Amadeus database excludes the names of firms that have not filed financial reports for

---

[19] We extend our gratitude to the tremendous work done by Rachel Griffith and the IFS team, especially Gareth Macartney in developing and implementing the patent matching. More information about the matching is available at: "AmaPat: Accounting, Ownership and Patents for European Firms" (CEP/IFS AmaPat document).

[20] The complete list of strings is available in the CEP/IFS AmaPat document.

four consecutive years (e.g. M&A, default). We deal with this issue in several ways. First, we use information from historical versions of the Amadeus database (1995-2003) on names and name changes. Second, even though Amadeus contains a unique firm identifier (BVD ID number), there are cases in which firms with identical names have different BVD numbers. In these cases, we use other variables for identification, for example: address (ZIP code), Date of incorporation (whether consistent with the patent application date), and more. Finally, we manually match most of the remaining corporate patents to the list of Amadeus firms.

A.2.2. United States Patents and Trademarks Office (USPTO)

The procedure described above matches European firms to patents registered with the EPO. Yet, some European firms register patents only with the USPTO, without applying to the EPO. In order to identify the European firms that only apply to the USPTO, we match the complete set of Amadeus firms to the name of the patent applicants from the USPTO. The most updated patent database for the USPTO is the 2002 version of the NBER patents and citations data archive.[21] Because this database covers patent information only up to 2002 and our accounting data go up to 2004, we updated the patent data file by extracting all information about patents granted between 2002 and 2004 directly from the USPTO website.[22] Having updated the USPTO patent database, we follow the matching procedure described above to create the matched USPTO patent data for the Amadeus firms.

Firms can apply for patents for the same invention with both the EPO and the USPTO. Patents protecting the same invention across different organizations are called a patent family. To avoid double counting of inventions, information on patent families is needed. We collect this information from the OECD Triadic database on patent families.[23] Having identified inventions that belong to the same family, we exclude patents granted by the USPTO that belong to the same family of patents granted by the EPO.

A.3. Matching academic publications

The largest database on academic publications is the ISI Web of Knowledge (WoK) by Thomson. This includes millions of records on publications in academic journals. The data is divided to three main categories based on the publication type: hard sciences, social sciences, and arts and humanities. Because we are interested in capturing investment in scientific research, we focus only on the hard sciences section of WoK. This section includes publication records over the period 1970-2004. The address field on each record indicates the affiliation of the authors of the publication. This affiliation is typically either a research institution or a firm. We use the name appearing

[21] http://elsa.berkeley.edu/~bhhall/bhdata.html

[22] http://patft.uspto.gov/netahtml/PTO/srchnum.htm

[23] This includes patents that are registered in all three main patents offices: the EPO, JPO, and USPTO.

in this field and match it to the complete list of Amadeus firms. We follow the same matching procedure as described above for the EPO and USPTO patent matching. Articles may have more than one author (the median number of authors per article is 2). In this case, the address field would include multiple affiliations. We assign an academic publication to a specific firm if the name of this firm appears at least once in the address field of the article. This procedure means that a single article can be assigned to more than one firm, but a firm cannot be assigned more than once to the same article. For each article, we also extract information on the number of times it was cited, the journal in which it was published, and the year of publication. Information about the importance of journals is taken from the Journal Citations Report index (JCR). Finally, European research institutions can be incorporated, thus, they appear in Amadeus as potential firms to be matched. To screen out such firms, we follow two steps. First, as for patent matching, we drop Amadeus names that include strings that are associated with research institutions (such as, UNIVERSITY, RESEARCH, INSTITUTION, etc.). Second, we manually examine the websites of firms that have a large number of publications but appear as small firms in terms of their sales and number of patents. For these firms, we check whether their primary activity is research. In case the primary activity is research, we exclude them from our matched sample.

A.4. Accounting database

The accounting information is taken from Amadeus. The database contains financial information on about 8 million firms from 34 countries, including all the European Union countries and Eastern Europe. The accounts of each firm are followed for up to ten years. The information source for Amadeus is about 50 country vendors (generally the office of register of Companies). The main advantage of Amadeus over other data sources is its coverage of small and medium size firms.

The accounting database includes items from the balance sheet (22 items) and income statement (22 items). No information is available from the changes in cash flow report (i.e., investment data is not available). The accounting data is harmonized by BvD to enhance comparison across countries. This comparison becomes easier over time due to the improvement in the European Union harmonization is accounting standards. In addition to accounting data items, Amadeus provides a description of firms including their product market activity. The main descriptive items are legal form (public versus private), date of incorporation, types of accounts (consolidated versus unconsolidated), country, US SIC and NAIC for the product market activity of the firm (primary and non-primary). The industry location information includes up to eight different six-digit NAIC codes per firm (note that the sales of the firm are not broken-up across the different product markets).

An important feature of the data is the criteria for dropping firms from the sample over time. As long as a firm continues to file its financial statements, it continues to appear in Amadeus. In case a firm becomes inactive, it stops filing its financial statement (alternatively, a firm can be late in filing its financial statement). This firm will be kept in the sample for four extra years since the last year financial statements were reported (thus, in the fifth year the firm will be removed from the sample). For example, a firm that becomes inactive and stops filing its reports in 1995 (i.e., 1994 is the last year when a financial statement was reported) will remain in the database until 1998 (including) and in 1999, it will be dropped from the sample (all observations of the specific firm will be taken out from the Amadeus database in the 1999 update). In order to mitigate the problem of losing dead firms, we purchased old Amadeus disks that allow tracking firms that exit the sample in previous years. For example, the firm that exits in 1995 will appear in the 1998 Amadeus disk, but not in the 1999 disk. By using both 1998 and 1999 disks, we mitigate the selection bias of dropping inactive firms after 4 years of missing data.

A.5. Constructing industry variables

We construct the industry measures used in the econometric specifications, using data from Compustat and Amadeus. The following variables are based on Compustat. They are a weighted average over the period 1980-2004 and are computed at the three-digit US SIC level. External Finance Dependence - this variable is defined as the ratio of Capital Expenditures (Compustat #128) minus Cash Flow (#110) to Capital Expenditures. When #110 is missing, Cash Flow is defined as the sum of the following Compustat items: #123, #125, #126, #106, #213, and #217. External Equity Dependence - this variable is defined as the ratio of the net amount of equity issued (#108 minus #115) to capital expenditures. Investment Intensity - this variable is defined as the ratio of capital expenditures to net property, plant, and equipment (#8). Tobin's Q - this variable is defined as the ratio of firm value to the book value of capital. Firm value is the sum of the values of common stock, preferred stock, and total debt net of current assets (#11 + #10 + #9 - #4). Book value of capital includes net plant, property and equipment, inventories and intangibles other than R&D (#3 + #8 + #33). Tobin's Q was set to 0.1 for values below 0.1 and at 20 for values above 20. R&D Intensity - this variable is defined as the ratio of R&D expenditures (#46) to sales (#12).

The following variables are based on Amadeus and are a weighted average over the period 1995-2004: Productivity Growth Dispersion - this variable is defined as the difference of the three-year average productivity growth rate between the 90th percentile and the 10th percentile for each three-digit US SIC. Productivity is defined as the weighted average of the ratio between sales and number of employees. Lerner index - this variable is defined as the industry median of the firm-weighted average of one minus the ratio of profits to sales. Both variables are computed based on UK data.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

832	Espen R. Moen Åsa Rosén	Incentives in Competitive Search Equilibrium
831	Eran Yashiv	U.S. Labor Market Dynamics Revisited
830	Laura Alfaro Andrew Charlton	Growth and the Quality of Foreign Direct Investment: Is All FDI Equal?
829	Richard Layard David Clark Martin Knapp Guy Mayraz	Cost-Benefit Analysis of Psychological Therapy
828	Emma Tominey	Maternal Smoking During Pregnancy and Early Child Outcomes
827	Christos Genakos Tommaso Valletti	Testing the "Waterbed" Effect in Mobile Telephony
826	Luis Garicano Paul Heaton	Information Technology, Organization, and Productivity in the Public Sector: Evidence from Police Departments
825	Laura Alfaro Andrew Charlton	Intra-Industry Foreign Direct Investment
824	Miklós Koren Silvana Tenreyro	Technological Diversification
823	Régis Barnichon	The Shimer Puzzle and the Correct Identification of Productivity Shocks
822	Carlos Thomas	Search Frictions, Real Rigidities and Inflation Dynamics
821	Giulia Faggio Kjell G. Salvanes John Van Reenen	The Evolution of Inequality in Productivity and Wages: Panel Data Evidence
820	Luis Araujo Emanuel Ornelas	Trust-Based Trade
819	Régis Barnichon	Productivity, Aggregate Demand and Unemployment Fluctuations
818	Rachel Griffith Sokbae Lee John Van Reenen	Is Distance Dying at Last? Falling Home Bias in Fixed Effects Models of Patent Citations
817	Tsutomu Miyagawa YoungGak Kim	Measuring Organization Capital in Japan: An Empirical Assessment Using Firm-Level Data

816	Andreas Georgiadis Alan Manning	Spend It Like Beckham? Inequality and Redistribution in the UK, 1983-2004
815	Guy Michaels Xiaojia Zhi	Freedom Fries
814	Ralph Ossa	Trade Liberalization, Outsourcing, and Firm Productivity
813	Karsten Kohn Alexander C. Lembcke	Wage Distributions by Bargaining Regime: Linked Employer-Employee Data Evidence from Germany
812	Yannis M. Ioannides Henry G. Overman Esteban Rossi-Hansberg Kurt Schmidheiny	The Effect of Information and Communication Technologies on Urban Structure
811	Guy Michaels	The Division of Labor, Coordination, and the Demand for Information Processing
810	Maia Güell Jose V. Rodriguez Mora Chris Telmer	Intergenerational Mobility and the Informative Content of Surnames
809	Stephen J. Redding Daniel M. Sturm Nikolaus Wolf	History and Industry Location: Evidence from German Airports
808	Anna Lipinska	The Maastricht Convergence Criteria and Optimal Monetary Policy for the EMU Accession Countries
807	Eran Yashiv	The Beveridge Curve
806	Ghazala Azmat Alan Manning John Van Reenen	Privatization, Entry Regulation and the Decline of Labor's Share of GDP: A Cross-Country Analysis of the Network Industries
805	Henry G. Overman Patricia Rice Anthony J. Venables	Economic Linkages Across Space
804	Benjamin Aleman-Castilla	The Returns to Temporary Migration to the United States: Evidence from the Mexican Urban Employment Survey
803	Eran Yashiv	Labor Search and Matching in Macroeconomics
802	Nicholas Oulton	Jeremy Greenwood and Per Krusell, "Growth Accounting with Investment-Specific Technological Progress: A Discussion of Two Approaches" A Rejoinder

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7955 7595 Email info@cep.lse.ac.uk
Web site http://cep.lse.ac.uk