WHAT HAPPENS WHEN FIRMS PATENT? NEW EVIDENCE FROM U.S. ECONOMIC CENSUS DATA

Natarajan Balasubramanian and Jagadeesh Sivadasan*

Abstract—We build a new concordance between the NBER Patent Data and U.S. Census microdata and use it to examine what happens when firms patent. We find strong evidence that increases in patent stock are associated with increases in firm size, scope, and skill and capital intensity. We find somewhat weaker evidence that changes in patenting are positively correlated with changes in total factor productivity. We also analyze first-time patentees and find similar effects following initial patent application. Together, these results suggest that patenting is indeed associated with real changes within firms, in particular with growth through increases in scope.

1. Introduction

PATENT statistics have been widely used in studies of firm innovation at least since Scherer (1965). The primary motivator for using these data has been, as Griliches (1990) put it, “the dream of getting hold of an output indicator of inventive activity.” In this context, Griliches poses two fundamental questions: What aspects of economic activity do patent statistics actually capture? and What would we like them to measure? To answer the first and, as Griliches argues, the potentially more relevant question, it is important to understand the changes that occur within firms when they patent. In this paper, we build and use a new concordance between the NBER Patent Data and firm data at the bureau to provide new evidence on this question.

To construct the concordance, we build on the pioneering efforts of Hall, Jaffe, and Trajtenberg (2001) and on the bureau’s extensive investments in developing and maintaining microlevel firm data. Using a detailed set of name-matching procedures, we develop an annual link between patent assignees in the NBER Patent Data (constructed by Hall et al., 2001) and firms in the Business Register (previously the Standard Statistical Establishment List) at the bureau. This new concordance covers about 48,000 assignees and about 121,000 assignee-years between 1975 and 1997, representing about two-thirds of all nonindividual, nonuniversity, and nongovernment assignees during this period.1

The comprehensiveness of the census data allows us to include private firms in our sample, and present a broader picture of patenting in the U.S. manufacturing sector than has been possible using publicly available data.2 Consistent with prior studies that used smaller samples, we find strong evidence for the highly skewed nature of patenting activity: just 5.5% of all manufacturing firms engage in patenting activity. These firms, however, have a disproportionate share of economic activity, accounting for about 59.3% of value added, 63.5% of capital stock, and about 52.2% of employment. We find significant heterogeneity across industries in the prevalence of patenting; this is broadly in line with Scherer’s (1983) and Bound et al.’s (1984) findings using data on a sample of large and listed U.S. corporations, respectively.

Turning to the main question of our study—What happens when firms patent—we find significant overall differences between patent-owning and non-patent-owning firms.3 We study the unconditional group means of four characteristics: size, factor intensity, productivity, and scope. We find that on average, patent-owning firms are much larger (by a factor of about 15 to 16 for output, value added and capital, and about 10 for employment), more skill intensive (by about 13%), and more capital intensive (by about 90%). Labor productivity of patenting firms is higher by about 52%, and total factor productivity (TFP) differences are about 17% (using a Solow residual definition). The gross number of products is higher by a factor of 2.75, and the net number of products by a factor of 2.11.

To understand if these overall differences simply reflect self-selection on preexisting differences, we investigate two aspects of patenting. First, we examine the within-firm elasticity of firm characteristics to changes in patent stock. We find that the elasticity of size with respect to patent stock is significant (0.156, 0.151, 0.173, and 0.139 log points for...

1 To facilitate the use of our new concordance by researchers, we have compiled a detailed user guide and technical documentation (Balasubramanian & Sivadasan, 2008). This technical note, to be made available to the Census and Bureau, includes a detailed discussion of the benefits and limitations of the new concordance, along with the advantages of the existing NBER-Compustat link.

2 Although the concordance covers all sectors, data access constraints limit the rest of our study to the manufacturing sector.

3 Throughout, we use terms such as patent-owning firms and patentees to refer to firms that are matched to one or more patent assignees. Although this is in line with studies in this literature, matching to an assignee does not always imply ownership. The patent data do not track reassignments or sale of patents by firms subsequent to the patents being granted, so that patent ownership is not tracked perfectly.

© 2011 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology
output, value added, capital stock, and employment, respectively). There is also evidence that skill and capital intensity respond positively to patent stock. The elasticity of productivity is smaller but still significant for both labor productivity and an OLS firm fixed effects (OLS-FE) estimate of total factor productivity. Both the gross and net number of products exhibit positive and significant elasticity with respect to patent stock.

We then use data on first-time patentees to examine changes associated with their first patent application. Using specifications that include both firm and industry-year fixed effects (which controls for industry-specific age trends), we find that firm size exhibits a significant increase after the first patent application. There is also evidence of an increase in skill and capital intensity, as well as in all productivity measures, following the first patenting event. The gross number of products shows a significant increase; results for net number of products suggest that firms drop some products around the switch. Using an event study approach, we attempt a finer exploration of the timing of these effects. These analyses suggest a significant increase in size and in the gross number of products coincident with the initial patent application. There is also some evidence of increases in factor intensity and labor productivity. Evidence for coincident increases in the TFP measures and in the net number of products is, however, weak.

Together our results strongly suggest that patenting is associated with real economic effects. These results add to findings from prior studies of patenting and innovation at the firm level (Griliches, 1981; Pakes, 1985; Austin, 1993; Sakakibara & Branstetter, 2001; Bloom & Van Reenen, 2002; Sampath & Ziedonis, 2004; Hall, Jaffe, & Trajtenberg, 2005), and support the use of patents as meaningful proxies of firm-level innovation. In particular, our results provide strong evidence that patenting is associated with firm growth through the introduction of new products. In addition to the results discussed above, our tests show that sales from new products do indeed account for all of the firm growth observed around the time of first patenting. We also find that the gross number of products increases with patenting even conditional on firm size measures. It may not seem very surprising that patents are associated with the introduction of new products, but this relationship has been underinvestigated in the literature, particularly using large samples. To our knowledge, this is the first such study using a comprehensive sample. Furthermore, this strong link between patenting and firm scope confirms the usefulness of patents as measures of new product introductions. The results also suggest that it may be important to consider firm scope (as in, for example, Klette & Kortum, 2004; Bernard, Redding, & Schott, 2006b; Nocke & Yeaple, 2006) in addition to firm productivity (as in, for example, Jovanovic, 1982; Hopenhayn, 1992; Ericson & Pakes, 1995; Melitz, 2003) as the source of heterogeneity across firms when modeling patenting behavior.

We undertook a number of other robustness checks of our results. We examined the baseline productivity results and found them to be robust to a variety of modifications of the baseline Cobb-Douglas two-input production function specification. Results of tests using a panel data GMM approach based on Arellano and Bond (1991) and Blundell and Bond (1998) to address potential endogeneity of patent stock and other inputs were qualitatively consistent with the baseline OLS-FE TFP results. The changes in factor coefficient estimates were in the same direction as observed by Griliches and Mairesse (1995), but the GMM models did not pass one or both of the specification tests.

Our baseline results were also robust to the exclusion of outliers, acquisitions, and highly diversified firms. We also examined if there were systematic differences between private and public firms by rerunning the baseline regressions separately for these two types of firms. We found the results to be similar for the size and scope variables. For the other variables, the effects were smaller and statistically insignificant in the sample of public firms, which is consistent with the significantly smaller sample.

Before we turn to the details, we must emphasize that patenting is not an exogenous event, so any changes associated with patenting should not be interpreted as being caused by patenting. Accordingly, we do not attempt to ascribe any causal interpretation to our results, but rather view them as consistent with patents capturing underlying innovative activity in the firm. Furthermore, we are agnostic about the drivers of innovation, external or internal to the firm. We did, however, attempt to rule out contemporaneous industry-wide demand shocks that affect both patenting and firm size. We included detailed four-digit industry-year effects in all of our specifications as controls for common four-digit industry-year demand shocks. As a robustness check, we used five-digit industry-year controls and found the correlation between size and patenting to be robust. As another check, we used lagged patent stocks as instruments for current patent stocks in a panel GMM approach similar to that discussed for productivity. Although the specifications failed one or both specification tests, the results were broadly consistent with the baseline results. We also checked and found that the size changes around the first patenting event were correlated with commonly used measures of patent quality. While these checks suggest that our results may not be driven by common industry-wide demand shocks that affect all firms, our results are consistent with latent demand being an important driver of innovation, with only successful innovators being able to tap into this demand (Schmookler, 1966).

The rest of the paper is organized as follows. Section II briefly discusses the key features and limitations of the new concordance. Section III discusses the data, variables, and empirical approach used in our study. Section IV provides descriptive statistics on the economic importance of patent-owning firms. Section V presents the overall differences in

---

4 We thank a referee for elucidating this argument and suggesting tests to address this.
means between patentees and nonpatentees on characteristics such as size, factor intensity, productivity, and scope. Section VI examines within-firm elasticity of various firm characteristics to patent stock. Section VII analyzes first-time patentees and is followed by a discussion of robustness checks in Section VIII. Section IX discusses our findings in the context of related literature and concludes.

II. The NBER Patent Data–Business Register Bridge

This section discusses some key features and limitations of the concordance between the NBER Patent Dataset and the Business Register (BR). Additional details are provided in a user guide and technical documentation intended to accompany the concordance (Balasubramanian & Sivadasan, 2008).

Our concordance has a number of novel features, all of which derive from the bureau’s extensive investments in the BR. The BR is based on information collected by the Internal Revenue Service and the bureau and covers all nonfarm employers in the United States. Thus, the concordance offers a link to establishments in all nonfarm sectors of the economy. Furthermore, it provides broader coverage of patent assignees than has been possible using publicly available data. Specifically, it extends the linkage of patent data to unlisted firms and facilitates an analysis of their patenting behavior; this has so far been difficult. The more important novelty arises from the dynamic ownership linkages developed and maintained by the bureau as part of the BR. The bureau collects ownership information on establishments in each quinquennial Economic Census. It also undertakes an annual company organization survey (COS), which includes all multieestablishment firms and a sample of single-unit firms to collect information on ownership and corporate structure. The new concordance leverages this information to provide dynamic updating of ownership linkages (albeit subject to some limitations, particularly in intercensus years). The current link between patent data and Compustat firm data available in the NBER Patent Data was created by manually searching through secondary sources of ownership structure. This involved enormous time and effort, and hence the ownership structure was limited to one year: 1989. Linking to the BR allows us to free-ride on the bureau’s work in developing, maintaining, and updating ownership links. Our concordance provides an assignee-year match, as opposed to a single assignee-firm match across all years. This year-by-year match permits tracking ownership changes among establishments. This appears to be important, as our checks show that patentee establishments are much more likely to change ownership than nonpatentee establishments are (see table 7 in Balasubramanian & Sivadasan, 2008).

The concordance has a few important limitations that should be considered carefully by those using it. The quality of ownership linkages improves significantly in the census years, as more resources are devoted to the BR in these years, and because the results of the Economic Census provide more ownership-related information than the annual COS (Jarmin & Miranda, 2002). In our work here, we focus on the Census of Manufacturing and hence use the best-quality ownership linkages. The second limitation of the BR is that it does not cover nonemployers and does not uniformly cover zero-employee establishments of multiunit firms. Thus, patents owned by subsidiaries with no employees may not be captured by the concordance. A third important constraint is that the concordance is restricted to U.S. assignees that are not individuals or governments because the BR does not cover individuals or foreign firms without establishments in the United States. Hence, this concordance cannot be used to analyze patenting by these entities. Moreover, the BR does not enable a robust identification of U.S. establishments of foreign firms. Finally, the limitations of name-matching procedures imply that some assignees may not get matched or may have broken longitudinal links. For more details on these limitations and for a discussion of some other limitations, researchers should refer to section 7 of our technical note (Balasubramanian & Sivadasan, 2008).

Besides these limitations, researchers should also consider the distinct strengths of the existing Compustat–Patent Data link. These include the widespread availability of Compustat data, links to executive compensation (and potentially other external) data unavailable within the bureau, and coverage of foreign subsidiary information in consolidated financial statements in Compustat. However, researchers who have access to Compustat and also obtain access to the census data may be able to exploit strengths of both data sets by using the Compustat–BR Bridge file available at the census, along with our NBER Patent Data–BR Bridge.

We conclude this section with an overview of the coverage obtained in the new concordance (for details, refer to the user guide and technical documentation). Of the 2.92 million patents granted between 1963 and 1999, belonging to 175,115 assignees, about 40% belonging to about 57,600 assignees were matched at least once to a firm in the BR. Within a more relevant population of patents applied between 1975 and 1997 by U.S. assignees that were not individuals, universities, or government agencies, 90% of patents and 63.7% of assignees (or about 48,000) are matched at least

---

5 In this project, we used identifiers available in the BR. The Census Bureau has recently constructed a panel data set of establishments with improved longitudinal linkages, called the Longitudinal Business Database (LBD), which is described in detail in Jarmin and Miranda (2002). Our bridge file can be easily linked to the LBD using establishment identifiers common to the BR and LBD, enabling future researchers to exploit the strengths of the LBD. Also, the Census Bureau has recently put together the Integrated Longitudinal Business Database (ILBD), which combines the LBD with information on nonemployers. In future work, our methodology for matching assignees to the BR could be extended to cover the nonemployer firms in the ILBD.

6 We thank Manuel Trajtenberg for pointing out the benefits from constructing a dynamic link.

7 A number of studies have exploited the ownership linkages available at the Census Bureau. Those that have particularly relied on these ownership linkages include McGuckin and Nguyen (1995), Schoar (2003), Bertrand and Mullainathan (2003), Bernard and Jensen (2007), Maksimovic and Phillips (2002), and Hortacsu and Syverson (2008).
III. Data, Variables, and Empirical Approach

A. Data

In order to create the primary data set for our analyses, we link the NBER-BR Bridge to data from the five censuses in 1977, 1982, 1987, 1992, and 1997. The establishment data are aggregated to the firm level using the firm identifier (CFN or ALPHA) available in the BR. If a firm has establishments in multiple SIC-4 industries, it is assigned to the largest SIC-4 industry as measured by share of firm output. It is important to note that a firm could be matched to multiple assignees in a single year; hence, data from all assignees are used to compute patenting statistics for the firm. For instance, stock variables such as number of patents applied for in the current year are computed as the sum of the number of patents applied for by all assignees matched to the firm in that year. Indicator variables such as whether a firm has applied for a patent at any time through the current year are based on whether the firm is matched to any assignee until that point in time. Throughout, we use application year as the baseline.

We focus on the manufacturing sector due to data access limitations. We had access to detailed microdata only for the manufacturing sector, though we had name and address information for all sectors. Although this potentially limits the generalizability of our results to nonmanufacturing industries, it is important to note that the bulk of patenting activity occurs within the manufacturing sector (Scherer, 1983). Our checks also show that approximately 70% of patents that were applied in the five census years and matched to firms in the BR were by manufacturing firms. Not surprisingly, a number of earlier studies have focused on the manufacturing sector or industries within that sector (Hall & Ziedonis, 2001).

We undertake all our analyses using the quinquennial Censuses of Manufacturing (CMF) instead of using data from the Annual Surveys of Manufactures (ASM). The primary reasons for this decision are significantly greater coverage and better data quality in the CMF. More specifically, an important novelty in our work is the inclusion of nonlisted firms, which are generally small or medium sized. The universal coverage of all manufacturing establishments in the CMF allows us to better analyze these firms. In addition, the quality of ownership links is highest for the census years (Jarmin & Miranda, 2002), so that firm-level patent stock and patenting status measures are more accurately defined in these years. Finally, data on firm scope variables, specifically, output product codes for each establishment, are not available in the inter census years.

B. Variables

Appendix 1 provides detailed definitions of each of the variables used in this study. Here, we present an overview of the key variables.

We define two patenting variables: patent stock and a patenting status. The patent stock measure is a depreciated patent stock, defined as the number of depreciated patents, with each patent depreciated at an annual rate of 15% from the patent application year (following Hall et al., 2005). Patenting status is defined as an indicator variable: 1 from the application date of the first patent (conditional on later approval) and 0 before.

We define four sets of firm characteristics: size, factor intensity, productivity, and scope. Size is measured using four variables: log output, log value added, log capital stock, and log employment. Two aspects of factor intensity are examined: capital intensity, measured as log capital per employee, and skill intensity, measured as the ratio of white-collar to blue-collar workers.

Productivity in the baseline case is measured in three ways: labor productivity, defined as the log output per employee; Solow residual, defined using the four-digit industry-level factor share of costs; and the residual from an OLS-FEs regression of log value added on log capital and log employment. The first measure is influenced by changes in capital intensity. The latter two measures are alternative measures of TFP, and net out changes in inputs. It is important to remember two points when comparing the Solow measure and the OLS-FE measure. First, the Solow residual is a measure of real sales productivity because it is defined as change in real sales less changes in the factor-share-weighted materials, capital structures, capital equipment, and employment. Thus, the magnitudes of changes in the Solow real sales TFP residual will be lower than the changes in the value-added TFP residual in the OLS-FE specification by about 30%. Second, the two measures could differ because different weights on inputs could yield significantly different productivity residuals. Because the key assumption of competitive input and output markets underlying the Solow measure could be violated in the context of patenting, which grants temporary monopoly rights within product categories, we place

---

8 The NBER Patent Dataset has identifiers for individuals or governments but not for universities. We used a visual inspection of names to identify 430 universities.
greater emphasis on the OLS-FE measure relative to the Solow measure. Another advantage of the one-step OLS-FE approach is that it allows potential correlation between patent measures and inputs. Alternative measures of productivity, including a GMM approach to controlling for potential endogeneity of inputs and patent measures in the OLS-FE specification, are discussed in section VIII B.

Firm scope is measured in two ways: the logarithm of the gross number of products, defined as the sum of the number of products produced by the firm over its lifetime to date, and the logarithm of the net number of products, defined as the number of products reported by the firm in the current year. Each distinct seven-digit SIC product code of a firm is classified as a product. Note that the first measure of scope captures all new product additions; the second measure captures only increases in the number of offered products. Thus, if firms drop older products when they introduce new ones (as documented by Bernard, Redding, & Schott, 2006a), the second measure may show no change or could even decrease with new product additions. Alternative measures of product scope are examined in section VIII A.

C. Empirical Approach

Our empirical analysis can be divided into six parts. We begin in section IV with some basic statistics on the role of patentees in manufacturing. We then examine cross-sectional differences between patentees and nonpatentees in section V. To understand if these cross-sectional differences simply reflect selection, we analyze the changes within firms. In the third part of our analysis, in section VI, we use a fixed-effects regression framework to examine the elasticity of firm characteristics to changes in patent stock. To better understand the timing of the effects and control for contemporaneous changes within the industry, we then examine changes among first-time patentees in section VI. Specifically, in section VI A, we do a before-and-after and difference-in-differences regression analysis of changes in firm characteristics associated with a change in firm patenting status. In section VII B, we attempt to obtain a more finely grained look at the timing of changes using an event-study analysis of the first patenting event. Finally, in section VIII, we perform a number of robustness checks of our results. More details on the empirical methodology used in each of these steps are discussed in the relevant sections that follow.

IV. Basic Statistics on the Role of Patentees in Manufacturing

We begin by documenting some basic statistics about the importance and extent of patenting in U.S. manufacturing. Table 1 highlights the economic importance of patent-owning firms in the U.S. manufacturing sector. The first column provides the proportion of firms within an SIC-2 industry that own a patent at some point between 1963 and 1997. The most striking fact is that on average, only about 5.5% of all firms own patents. Furthermore, in no SIC-2 industry does this proportion exceed 20%. These figures are consistent with the stylized fact that a large fraction of firms report 0 or little R&D expenditure (Cohen & Klepper, 1992) and with the findings in Cohen, Nelson, and Walsh (2000) that patents are not viewed as very effective means of protecting profits from inventions in a majority of manufacturing industries. Another explanation could be that there are fewer
innovations in some industries during the period of our sample. Not surprisingly, industries vary in the fraction of patent-owning firms: only about 0.5% of all firms in Lumber and Wood Products (SIC 22) owned patents compared to a little over 17% in Instruments and Related Products (SIC 38).

The next three columns in table 1 present the share of patent-owning firms in industry value added, capital stock, and employment. Although they form a small fraction of firms by number, patent-owning firms appear to dominate economic activity. In twelve of the nineteen industries, the small number of patent-owning firms accounts for more than half the industry value added. On average, they account for more than 52% of all employment in manufacturing and over 63% of all capital stock. Furthermore, even in industries where the share of patent-owning firms is very small, these firms have a significant economic role (for example, patenting firms account for only 0.5% of all firms in lumber and wood products, but they contribute almost 30% of all value added). In industries such as transportation (SIC 37), chemicals (SIC 28), and instruments (SIC 38), almost 90% of industry value added can be attributed to the relatively few firms that own patents. Another interesting aspect of table 1 is that patent-owning firms generally account for a larger share of value added than of employment. This is true in almost every SIC-2 industry. Hence, on average, these firms seem to have much greater labor productivity than their nonpatentee counterparts.

It is well established that R&D activity is skewed, with a major share of activity being concentrated in large firms (Bound et al., 1984; Cohen & Klepper, 1996). The economic role of patenting firms in each size decile is presented in figure 1. As expected, the extent of patenting increases with size, but the increase is mostly concentrated in the largest size decile. Almost 80% of value added and a little less than 70% of employment in the tenth size decile pertain to patent-owning firms. In contrast, in the ninth size decile (the second largest-size decile), patent-owning firms account for less than 20% of value added and employment. Another point of note in figure 1 is that even in the largest size class, only about 20% of firms own patents.

Together, table 1 and figure 1 confirm that many of the facts previously documented in the literature using smaller data sets hold true in this more comprehensive census data set.

V. Overall Differences between Patentees and Nonpatentees

We now turn to an examination of the overall differences between patentees and nonpatentees. Table 2 presents the group means of the various firm characteristics for patentees and nonpatentees. Comparing the simple means of the size variables (columns 1 and 2), it is evident that, on average, patentees are much bigger: output is larger by a factor of 1.5, capital stock by a factor of 1.6, and employment by 9.8. These results demonstrate that the size differences between patenting and nonpatenting firms are substantial economically and statistically; they are also consistent with the findings in table 1 that suggest an important economic role for these firms. Patent-owning firms also exhibit higher skill intensity—the ratio white-collar to blue-collar workers for these firms is higher by about 13%—and significantly greater use of capital, almost 90 percent more capital per worker.

Patent-owning firms are also more productive than other firms. The mean (log) output per employee for patent-owning firms is higher by a factor of 1.5. The Solow TFP measure

12 These variables are logarithms of the original values, and hence the factor, say, for output is computed as $e^{2.709} = 15.01$.

13 In unreported results, we also examined differences in wage rates between patenting and nonpatenting firms. We find that patentees pay higher wages for both production and nonproduction workers. This is consistent with the use of higher-quality or skilled workers by patenting firms.
is also higher for patent-owning firms, although the magnitude of the difference (17%) is not as high as that for labor productivity.\textsuperscript{14} Finally, the gross and net number of products are higher for patentees by a factor of about 2.8 and 2.1, respectively. Translating the mean of the logs into levels, the patentees have about 5.1 additional products over their lifetime (in gross terms) and about 2.0 more products offered at any point in time (in net terms).\textsuperscript{15}

Taken together, the statistics in table 2 suggest that patent-owning firms are very different from their nonpatenting counterparts. They are much larger, tend to choose higher levels of skilled labor and capital, exhibit greater productivity, and have a larger number of product offerings than nonpatentees. The next sections go beyond these mean differences and attempt to understand if some of these differences are attributable to changes in patenting within firms.

VI. Patent Stock and Firm Characteristics

In this section, we examine if changes in patent stock are related to changes in firm characteristics. We use the following fixed-effects regression specification:

\[ y_{ijt} = \delta s_{ijt} + \eta_i + \mu_j + \epsilon_{ijt}, \tag{1} \]

where \( y_{ijt} \) is the characteristic (for example, log output) for firm \( i \) in period \( t \), \( s_{ijt} \) is the log of the depreciated patent stock, \( \eta_i \) refers to firm fixed effects, \( \mu_j \) are four-digit industry-year effects, and \( \epsilon_{ijt} \) is the residual error term. This specification is used to examine all variables except the one-step OLS-FE measure of productivity. Hence, for the other measures of productivity, labor productivity and the Solow TFP measure, productivity is measured, or derived, in a first step. In the second step, the productivity term is regressed on patent stock in specification 1.

For measuring the impact of changes in patent stock on the OLS-FE measure of productivity, we use the following single-step approach:

\[ v_{ijt} = \alpha k_{ijt} + \beta l_{ijt} + \delta s_{ijt} + \eta_i + \mu_j + \epsilon_{ijt}, \tag{2} \]

where \( v_{ijt} \) is the log value added, \( k_{ijt} \) is log of the capital stock, and \( l_{ijt} \) is log of employment. This specification is similar to those used in other studies of patenting (for example, Bloom & Van Reenen, 2002). This specification also controls for industry-specific age effects because age effects for firm \( i \) in industry \( j \) in year \( t \) can be decomposed as \( \theta_i \text{Age}_{it} = a_i + b_i t \), which are absorbed by the firms and industry-year fixed effects.\textsuperscript{16}

Because specifications (1) and (2) involve two sets of high-dimensional fixed effects (more than 2,000 industry-year effects and more than 5,000 firm fixed effects), it was practically infeasible to run these using standard regressions routines. To implement the specifications, we used the methodology and code proposed by Andrews, Shank, and Upward (2006).\textsuperscript{17} Standard errors were clustered at the industry level to allow arbitrary within-industry correlation of errors.\textsuperscript{18}

The results, summarized in table 3, indicate an economically and statistically significant elasticity of size with respect to patent stock. A 10% increase in log patent stock is associated with a 1.57% increase in output, a 1.52% increase in value added, a 1.73% increase in capital stock, and a 1.39% increase in employment.

There is a small but statistically significant effect on skill intensity; a 10% increase in patent stock is associated with an increase in the ratio of white-collar to blue-collar workers by 0.02%. Capital intensity is also positively correlated with patent stock, though the magnitude of the effect appears small; a 10% increase in patent stock is associated with a 0.37% increase in capital intensity.

Labor productivity increases by 1.76% for a 10% increase in log patent stock, and 1.73% in value added, a 1.39% increase in capital stock, and a 0.37% increase in capital intensity.

\textsuperscript{17} This methodology incorporates the grouping algorithm proposed by Abowd, Creecy, and Kramarz (2002). A detailed description of some of the procedures in the felsdvreg procedure proposed by Andrews et al. (2006) is available in Cornelissen (2008).

\textsuperscript{18} As a referee pointed out, for firms that switch industries, the industry clustering may not capture error correlations within firms. We checked and found our results robust to clustering at the firm level; standard errors were consistently slightly smaller when clustered at the firm level, but the qualitative conclusions (and significance levels) generally remained the same.
precisely equal to the coefficient in labor productivity. The TFP measured as the Solow residual shows a very small and insignificant elasticity with respect to patent stock. The one-step OLS-FE specification does show a significant increase; a 10% increase in patent stock increases this measure of productivity by 0.15%.

Finally, changes in the number of products produced by the firm are positively related to changes in patent stock. A 10% increase in patent stock is associated with a 0.35% increase in the gross number of products and a 0.40% increase in the net number of products.

To summarize, these analyses strongly suggest that patenting is associated with real changes within firms. The largest impact is on firm size, followed by firm scope. The impact appears to be smaller on productivity and factor intensity.

VII. Patenting Status and Firm Characteristics: Analysis of First-Time Patentees

The analysis suggests that some of the overall differences between patentees and nonpatentees documented in table 2 are indeed driven by within-firm changes associated with changes in the extent of patenting. However, the analysis in section VI excludes data on all nonpatentees (since the log of patent stock is undefined for these firms) and thus does not capture any changes associated with a switch in status from being a nonpatentee to a patentee. To analyze this transition effect, we examine a sample of first-time patentees who began as a nonpatentee and became a patentee during our sample period. As discussed earlier, we define the switch as happening in the year that the firm first applied for a patent, conditional on the application being subsequently approved.

We use a regression framework to analyze changes before and after patenting (using firm fixed effects) and examine the difference-in-differences effects relative to other firms in the industry (using firm and industry-year fixed effects). We supplement these analyses with an event-study approach that attempts a more finely-grained exploration of the timing of these effects. We define the first time a firm patents as an event and examine the timing of changes in firm characteristics relative to this event date.

A. Regression Analysis of First-Time Patenting

We begin with a before-and-after analysis of the change in firm characteristics that accompany the switch. We use the following specification:

\[ y_{ijt} = \gamma_{ba} D_{at} + \eta_i + \epsilon_{ijt}, \]  

(3)

where the variables are as defined in equation (1). The coefficient \( \gamma_{ba} \) provides a simple estimate of the changes that accompany the patenting event. For this analysis, the sample is restricted to firms that switch status during our sample period. As with equation (1), the OLS-FE results use a specification similar to (3), but with log value added as the dependent variable and labor and capital included on the right-hand side.

The results are presented in table 4, panel A. Column 1 shows results using all available data for the switchers. In column 2, we focus on the short-run effects by restricting the sample to five years before and after the switch year. The results suggest large before-and-after changes in all four size measures. Consistent with the results in table 3, the magnitudes of the increase in value added and output are similar (about 0.54 log points in column 2 for the five-year window around the switch). The increase in capital stock is larger than the increase in output (about 0.70 log points in column 2). In contrast, the jump in employment (about 0.44 log points in column 2) is smaller than the increase in output, and much smaller than the increase in capital. Consistent with these results, there is a significant increase in capital intensity after patenting (about 0.34 log points in column 2) and a small but significant increase in skill intensity (by about 3.8 percent).

Consistent with the larger increase in output relative to employment, we find an increase of 0.10 log points (see column 2) in labor productivity following the switch. The Solow residual measure of productivity shows no significant change around the switch, but the OLS-FE measure shows an increase in TFP of about 7.4% around the switch. As discussed in section III, differences in results between OLS-FE and Solow measures could be driven by the different capital and labor coefficients in the two methods. In particular, the results for size variables in Table 4, panel A do not unambiguously show that increases in sales are greater than the increases in all inputs. The similar magnitudes of changes in sales and value-added suggest that materials input changed roughly in proportion with sales. Capital input increased more than output, while employment increased less than output. Thus, whether TFP increased depends on how capital and employment are weighted in the estimation of TFP, with higher weighting of capital yielding lower TFP estimates. While the ratio of the overall average Solow factor share of labor to capital is similar to the ratio of the estimated OLS-FE coefficients, the discrepancy could be the result of higher weighting of capital in some industries for the Solow measure.

Both measures of scope show significant increases following the first application for patent, with the gross measure exhibiting a higher increase. Around the time of the switch, there is a 0.51 log point change in the gross number of products, compared to a 0.22 log point change in the net number of products. The mean log gross number of products in the
and the five years after the switch. All regressions include firm fixed effects. Robust standard errors clustered at the four-digit industry level are in parentheses.

The Pennsylvania Bureau of Labor Statistics (PBL) collects data on wage rates every year. In this data set, the white- to blue-collar wage ratio is already (log) transformed, and we do not transform it back to its original units. All other dependent variables, except the white- to blue-collar wage ratio, are logged.

For different industries, we control for common industry-specific shocks, as well as for industry-specific linear age trends, and obtain a difference-in-differences estimate of the change in firm characteristics by including industry-year and firm-fixed effects as well as for industry-specific linear age trends, and obtain a difference-in-differences estimate of the change in firm characteristics by including industry-year and firm-fixed effects as well as four-digit industry-year effects. Robust standard errors clustered at the firm level are in parentheses.

In any regression, the coefficient for the before-and-after effect could be the result of common industry-specific shocks, as well as for industry-specific linear age trends, and obtain a difference-in-differences estimate of the change in firm characteristics by including industry-year and firm-fixed effects as well as four-digit industry-year effects. Robust standard errors clustered at the firm level are in parentheses. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Panel B: The samples are restricted to those firms that enter without a patent and switch to being a patentee later in their life ("switchers"), plus the sample of nonpatentees (firms that did not apply successfully for a patent any time during the sample period). Additionally, in the samples in column 2, the data for switching firms are restricted to the five years before and the five years after the switch. All regressions include firm fixed effects as well as four-digit industry-year effects. Since we are interested in the changes relative to nonpatentees, the sample includes only switching firms and firms that never patent during the sample period. Additionally, in the samples in column 2, the data for switching firms are restricted to the five years before and the five years after the switch. All regressions include firm fixed effects as well as four-digit industry-year effects. Robust standard errors clustered at the firm level are in parentheses. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

As we noted below equation (2), this specification controls for industry-specific linear age trends, as these can be decomposed as \( \delta J_A e_{it} = a_i + b_j t \), which are absorbed by the firms and industry-year fixed effects. Since we are interested in the changes relative to nonpatentees, the sample includes only switching firms and firms that never patent during the sample period. The coefficient \( \gamma_{did} \) then captures the change in patentee characteristics relative to changes in nonpatentees in the same four-digit industry. As before, the OLS-FE results use specifications similar to equation (4), but with log value added as the dependent variable and labor and capital as the inputs. These specifications involved two high-dimensional fixed effects (more than 2,000 industry-year

This table reports coefficients on a dummy indicator variable that equals 1 if the firm has applied for any patents up to and including the observation year (conditional on subsequent approval), and 0 otherwise. All dependent variables, except the white- to blue-collar wage ratio, are logged.

As we noted below equation (2), this specification controls for industry-specific linear age trends, as these can be decomposed as \( \delta J_A e_{it} = a_i + b_j t \), which are absorbed by the firms and industry-year fixed effects. Since we are interested in the changes relative to nonpatentees, the sample includes only switching firms and firms that never patent during the sample period. Additionally, in the samples in column 2, the data for switching firms are restricted to the five years before and the five years after the switch. All regressions include firm fixed effects as well as four-digit industry-year effects. Robust standard errors clustered at the firm level are in parentheses. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Panel A: The samples are restricted to those firms that enter without a patent and switch to being a patentee later in their life ("switchers"). Additionally, the samples in column 2 are restricted to the five years before and the five years after the switch. All regressions include firm fixed effects. Robust standard errors clustered at the four-digit industry level are in parentheses. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 4.—Patenting and Firm Characteristics: Before-and-After Effects

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Before-and-After Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Output</td>
<td>0.77094***</td>
<td>0.54074***</td>
</tr>
<tr>
<td>Value added</td>
<td>0.78273***</td>
<td>0.54607***</td>
</tr>
<tr>
<td>Capital stock</td>
<td>1.04167***</td>
<td>0.70520***</td>
</tr>
<tr>
<td>Employment</td>
<td>0.57554***</td>
<td>0.43893***</td>
</tr>
<tr>
<td><strong>Factor intensity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skill intensity (white- to blue-collar worker ratio)</td>
<td>0.05057***</td>
<td>0.03844***</td>
</tr>
<tr>
<td>Capital intensity (capital stock per worker)</td>
<td>0.56193***</td>
<td>0.34257***</td>
</tr>
<tr>
<td><strong>Productivity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor productivity</td>
<td>0.19561***</td>
<td>0.10181***</td>
</tr>
<tr>
<td>TFP (Solow residual)</td>
<td>0.000565 (0.01907)</td>
<td>0.00039 (0.01469)</td>
</tr>
<tr>
<td>OLS-FE (one-step)</td>
<td>0.11319***</td>
<td>0.07398***</td>
</tr>
<tr>
<td><strong>Scope</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross number of products</td>
<td>0.78727***</td>
<td>0.51138***</td>
</tr>
<tr>
<td>Net number of products</td>
<td>0.33172***</td>
<td>0.21792***</td>
</tr>
<tr>
<td>Number of observations</td>
<td>31,336</td>
<td>15,382</td>
</tr>
<tr>
<td><strong>B. Differences in Differences Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Output</td>
<td>0.64800***</td>
<td>0.35360***</td>
</tr>
<tr>
<td>Value added</td>
<td>0.45141***</td>
<td>0.36646***</td>
</tr>
<tr>
<td>Capital stock</td>
<td>0.45525***</td>
<td>0.37409***</td>
</tr>
<tr>
<td>Employment</td>
<td>0.40201***</td>
<td>0.33900***</td>
</tr>
<tr>
<td><strong>Factor intensity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skill intensity (white- to blue-collar worker ratio)</td>
<td>0.01895***</td>
<td>0.02187***</td>
</tr>
<tr>
<td>Capital intensity (capital stock per worker)</td>
<td>0.09346***</td>
<td>0.08205***</td>
</tr>
<tr>
<td><strong>Productivity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor productivity</td>
<td>0.05679***</td>
<td>0.03460***</td>
</tr>
<tr>
<td>TFP (Solow residual)</td>
<td>0.02640***</td>
<td>0.02236***</td>
</tr>
<tr>
<td>OLS-FE (one-step)</td>
<td>0.07066***</td>
<td>0.04652***</td>
</tr>
<tr>
<td><strong>Scope</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross number of products</td>
<td>0.09553***</td>
<td>0.13320***</td>
</tr>
<tr>
<td>Net number of products</td>
<td>0.0115 (0.00960)</td>
<td>0.04123***</td>
</tr>
<tr>
<td>Number of observations</td>
<td>1,333,963</td>
<td>1,318,009</td>
</tr>
</tbody>
</table>

This sample is about 1.7; hence, a 0.5 log point change (say, from 1.4 to 1.9) represents an addition of approximately 2.7 new products (change from 4.0 to 6.7 products). Similarly, the mean net number of products is about 0.95; a change of 0.21 translates to an addition of about 0.6 products in level terms (assuming a change from 0.9 to 1.1 in log terms). These larger changes in gross number of products imply that subsequent to the transition, patenting firms drop a number of products when they add new products.22

An important limitation of specification 3 is that part of the before-and-after effect could be the result of common industry-specific shocks. Also, some of the growth in size could be explained by age effects. Suppose all surviving firms grow; then older firms could be expected to be larger than younger firms. This survivor-to-age effect could be different for different industries. We control for common industry shocks, as well as for industry-specific linear age trends, and obtain a difference-in-differences estimate of the change in firm characteristics by including industry-year and firm-fixed effects as in the following specification:

\[
y_{ijt} = \gamma_{did}D_{it} + \eta_i + \mu_j + \epsilon_{ijt}.
\] (4)

We thank a referee for the suggestion that led us to examine both the gross and net number of products. As the referee noted, the importance of adding and dropping products in the manufacturing sector is documented in Bernard et al. (2006a).
effects and more than 250,000 firm effects) and were estimated using the felsdvreg procedure developed by Andrews et al. (2006). Unlike with equations (1) and (2), the very high number of firm effects made it computationally infeasible to estimate standard errors clustered by industry. Hence, we were constrained to clustering standard errors at the firm level for these regressions.23

The results, presented in Table 4, panel B confirm the findings for panel A. All measures of output increase significantly after patenting, including in the window around the switch. The estimated increases in output, value added, and capital stock are comparable (about 0.45 log points in column 1 and 0.37 log points in column 2), while the increase in employment is slightly smaller, but still large and significant (about 0.40 log points in column 1 and 0.34 log points in column 2).

Both measures of factor intensity also show significant increases. The skill intensity measure shows an increase of 1.9 percentage points after patenting and a slightly higher 2.2 percentage points around the switch. Consistent with the larger increase in capital relative to employment, the capital intensity measure increases subsequent to patenting (about 0.09 log points in column 1 and a slightly lower 0.08 log points around the switch).

Also consistent with the greater increase in output relative to employment, labor productivity increases postpatenting with a slightly lower increase around the switch (0.035 log points) relative to the overall change (0.057 log points). The changes in output, value added, capital, and employment suggest that output increased almost in proportion to materials and capital but greater than employment. Consistent with this, we find a significant increase in the Solow measure of real sales productivity by about 2.6% overall and about 2.2% around the switch. The OLS-FE productivity measure also shows an increase of about 7% overall and about 4.85% around the switch.

The gross number of products shows a significant increase subsequent to patenting. This increase is larger around the switch (column 2), suggesting that industry peers make up some of the difference in the subsequent years. The pattern is similar for net number of products: the overall change (column 1) is insignificant, while there is a significant increase around the switch (column 2). Also, the larger changes in the gross number of products (relative to net number of products) suggest that patenting firms drop products in the postpatenting period.

Taken together, the results in this section bolster the findings in section VI. Similar to the within-firm changes associated with increases in patent stock, we observe significant changes associated with a change in patenting status. As before, the strongest effects are on firm size, with the other variables exhibiting smaller changes.

23 As an approximate check, we adjusted the standard errors in panel B of table 4 by the ratio of SEs clustered by firm to SEs clustered by industry (the ratio being obtained from panel A, table 4). All the significance levels remained unchanged, except for column 1 for the Solow residual measure, where the significance dropped to the 10% level.

24 Note that although the data we use are from the five-yearly censuses, the year of patenting is known accurately, as long as the first year of patenting was after 1963.

25 Supplementary materials referenced throughout the article are available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00058.

B. Timing of Effects: Event Study Analysis of First-Time Patenting

The analyses in the previous section indicate significant changes in firm characteristics subsequent to a change in patenting status. In this section, we attempt a more finely-grained investigation of the timing of changes associated with a change in patenting status using an event-study approach.

In particular, for each firm that switches status from non-patentee to patentee, we define a variable, index, as the difference between the current year and the first year the assignee applies for a patent.24 Thus, index takes negative integer values for the years before the firm first applies for a patent, zero in the first year of applying for a patent, and positive integer values subsequent to its first patent application year. Because the number of observations for extreme values of index was small, the variable was censored at −15 on the lower end and at +15 on the upper end. This variable was subsequently used in OLS regressions of the following form:

\[ y_{ijt} = \sum_{k=-15}^{15} \beta_k D_{kit} + \epsilon_{it}, \]  

where \( y_{it} \) is the variable of interest for firm \( i \) in industry \( j \) in period \( t \), dummy \( D_{kit} = 1 \) if index = \( k \) for firm \( i \) in year \( t \). We then plot the \( \beta_k \) coefficients against index to obtain a picture of how the dependent variable changes (relative to industry peers) before and after patenting. To focus on the effects of the patenting event over the short to medium horizon, we restrict attention to seven years before and seven years after the first patent application (index values ranging from −7 to +7).

We first undertake a simple before-and-after analysis, including in the sample only patentee firms. The results from this analysis are presented in figure A1 in the supplementary appendix.25 This figure suggests a significant increase in the size variables coincident with the patent application date. There also appears to be an increase in capital and skill intensity around the patent application date, as well as an increase in the labor productivity measure. Consistent with the earlier results, there is a significant increase in scope around the patenting event.

As discussed in section VIIA, the before-and-after analysis could be potentially affected by industry-specific shocks, common to both patentees and nonpatentees. Another concern is that of survivor bias. The NBER patent database includes only applicants that successfully applied for a patent. Although we do not specifically select on survival, the successful patentees could be those that are more likely to survive the two- or three-year gap between the application date and...
grant of the patent. Further, because we restrict attention to switching firms with observations before and after the patenting event, this means that the postpatenting (index $\geq 0$) coefficients in the before-and-after analysis are estimated based on firms that survive between the pre- and postpatenting periods.

In section VII A, we discussed how inclusion of firm and industry-year effects controls for age-to-survivor bias by controlling for industry-specific age trends. A further concern could be that the age and survivor effects are more complex. In particular, age effects could be nonlinear (for example, survivors grow faster when they are younger), or the growth of the patenteer firms is correlated with their initial size (for example, surviving firms that are smaller at a younger age grow faster than firms that are born larger). These concerns could bias our results if patent applicants are generally younger or inordinately smaller at a younger age.

In order to carefully control for these concerns, we undertake a difference-in-differences analysis relative to a matched nearest neighbor. We assigned each switching firm and a matched nonpatentee control firm to a cell and reran specification 5 with cell-year fixed effects. The matched control firm, chosen from among nonpatentees in the same SIC four-digit industry and of the same age as the switching firm, was the one closest in employment size to the switching firm in the closest data year prior to patenting. The underlying sample consists of approximately 8,000 switchers and an equal number of nonpatentee control firms.

Points on the graph are coefficients on a dummy indicator variable, which equals 1 if the firm applied (successfully) for a patent, interacted with an index indicating the number of years from patenting. Specifications include cell-year fixed effects, where each cell consists of exactly one switcher (a firm that switches status from nonpatentee to patentee), and one control firm, which is a nonpatentee firm in the same four-digit industry and same age as the switcher, and nearest in employment size to the switcher firm in the closest data year prior to patenting. The underlying sample consists of approximately 8,000 switchers and an equal number of nonpatentee control firms.

In section VII A, we discussed how inclusion of firm and industry-year effects controls for age-to-survivor bias by controlling for industry-specific age trends. A further concern could be that the age and survivor effects are more complex. In particular, age effects could be nonlinear (for example, survivors grow faster when they are younger), or the growth of the patenteer firms is correlated with their initial size (for example, surviving firms that are smaller at a younger age grow faster than firms that are born larger). These concerns could bias our results if patent applicants are generally younger or inordinately smaller at a younger age.

In order to carefully control for these concerns, we undertake a difference-in-differences analysis relative to a matched nearest neighbor. We assigned each switching firm and a matched nonpatentee control firm to a cell and reran specification 5 with cell-year fixed effects. The matched control firm, chosen from among nonpatentees in the same SIC four-digit industry and of the same age as the switching firm, was the one closest in employment size to the switching firm in the closest data year prior to patenting. The underlying sample consists of approximately 8,000 switchers and an equal number of nonpatentee control firms.

Points on the graph are coefficients on a dummy indicator variable, which equals 1 if the firm applied (successfully) for a patent, interacted with an index indicating the number of years from patenting. Specifications include cell-year fixed effects, where each cell consists of exactly one switcher (a firm that switches status from nonpatentee to patentee), and one control firm, which is a nonpatentee firm in the same four-digit industry and same age as the switcher, and nearest in employment size to the switcher firm in the closest data year prior to patenting. The underlying sample consists of approximately 8,000 switchers and an equal number of nonpatentee control firms.

26 A note with more detail (including simulation results and computer code) about how the matching procedure with cell-year effects addresses potential survivor bias is available on request from the authors. We also undertook an event-study analysis controlling for SIC four-digit industry-specific age-quintile fixed effects (supplementary appendix figure A2). These figures confirm that the results from the before-and-after analysis of figure A1 are not driven by (potentially industry-specific) age-related effects. However, by matching on preapplication size, the matched-control procedure in figure 2 controls also for potential size-dependent growth rates among survivors.
The results, presented in figure 2, are qualitatively similar to the before-and-after results in figure A1. In particular, the size measures show a significant increase in year 0—the year in which the firms first apply for a patent. The magnitude of the increase is similar for sales, value added, and capital, but somewhat smaller for labor, consistent with the results in previous sections.

The results for skill intensity are similar to the patterns for size—there appears to be an increase in the year of patenting. The pattern for capital intensity is less stark with no systematic effect around the patent application year. The productivity graphs are presented in the lower left of figure 2. There is a jump in labor productivity in the year of patenting, but it is not persistent. There appears to be no systematic change in either of the TFP measures, though there is some evidence of an increase in year 0.

The scope figures show a stark jump in the log gross number of products, but no systematic increase in the net number of products. This is broadly consistent with the results in table 4, panel B, and suggests that patenting firms retire some products around the time new products are introduced.27

Overall, the magnitudes of size changes in figure 2, as well as qualitative results for skill intensity and gross number of products, are similar to those in column 2 of panel B in table 4. In contrast to column 2, the effects for total factor productivity and capital intensity do not appear significant in figure 2. The main source of differences between figure 2 and panel B is that by including cell-year fixed effects, in Figure 2 we allow age effects to vary arbitrarily across age and initial size; the specification in Table 4, panel B, on the other hand, controls for linear industry-specific age effects (through the simultaneous inclusion of firm fixed effects and industry-year effects). Consistent with this explanation, simple before-and-after graphs (presented as figure A1 in a supplementary appendix) obtained by running specification 5 without any fixed effects, show effects that are qualitatively and quantitatively similar to the results in table 4, panel A.

### VIII. Robustness Checks

We performed a number of checks to examine the robustness of our results. We begin by discussing the results on firm scope in section VIII A, and then turn to the robustness of productivity results in section VIII B. Section VIII C discusses a number of checks to address potential endogeneity from demand shocks. Finally, section VIII D includes a discussion of other robustness checks.

#### A. Firm Scope (New Products) Analysis

An unsurprising but novel finding of our study is evidence of the introduction of new products around the time of patenting. In this section, we explore this finding in greater detail.

To understand the relationship between patenting and new product introduction, we focus on the changes around the patenting event among first-time patentees. We analyze the contribution of products introduced by these patentees after their first patent application, both before and after (using specification 3), as well as relative to peers in the industry (using specification 4). Results of our analysis are presented in the first three rows of table 5.

We focus our discussion here on the short-term effects in columns 2 and 4. The before-and-after specifications in column 2 indicate an increase in sales of about $6.3 million around the switch. Interestingly, sales of old, prepatenting products decline by about $1.4 million on average. Thus, all of the increase in total sales comes from a $7.7 million increase in new postpatenting products, which more than offsets the decline in sales from older products. The magnitude of the overall increase in sales is smaller in the difference-in-differences estimates in column 4 (about $4.4 million), but the estimated decline in old products’ sales is higher (about $2.2 million). As before, the increase in new product sales offsets the decrease in old product sales and accounts for all of the change in total sales.

<table>
<thead>
<tr>
<th>Table 5.—Analysis of Scope Results: New Product Sales and Product Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>Switchees Overall</td>
</tr>
<tr>
<td>Total product sales (in $000s)</td>
</tr>
<tr>
<td>Old (prepatenting) product sales (in $000s)</td>
</tr>
<tr>
<td>New (postpatenting) product sales (in $000s)</td>
</tr>
<tr>
<td>Gross number of products (count)</td>
</tr>
<tr>
<td>Net number of products (count)</td>
</tr>
<tr>
<td>Number of observations</td>
</tr>
</tbody>
</table>

27 Graphs with confidence intervals are presented in graphs A3a, A3b, A3c, and A3d of the supplementary appendix, available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00058, and confirm the qualitative conclusions reached above.
As another check, we used the simple count of the gross and net number of products instead of their logarithms. These results are presented in the fourth and fifth rows of table 5 and are qualitatively similar to the baseline results.

We also examined the sensitivity of the scope variable results in tables 3 and 4 to the inclusion of size and factor intensity controls.28 Note, however, that this analysis is largely exploratory. The impact of controlling for size could be ambiguous. On the one hand, patent-related innovations could result in a higher number of products per employee or per unit revenue. On the other hand, patent-related innovations could have a larger impact on size than on scope. For instance, the invention of one blockbuster product, like the light bulb, may dramatically increase revenue and employment in the innovating firm, so that conditional on revenue or employment, the number of products could actually decline after patenting. This effect is particularly likely to be true relative to nonpatenting firms that do not introduce such blockbuster products but bring a number of less successful products to the market.

These results are presented in supplementary appendix table A1. We find that the elasticity of gross number of products to patent stock is lower in magnitude, but positive and significant even after conditioning on size and factor intensity variables. The before-and-after and the difference-in-differences effects of first-time patenting show the same pattern with the additional controls: somewhat lower in magnitude, but positive and significant. In most cases, the elasticity of the net number of products with respect to patent stock is not statistically significant. While the before-and-after effect is positive and significant, the difference-in-differences results are negative and significant once we add size controls. This is consistent with patenting firms introducing new products that generate more per-product revenue than the retired products did.

Together, these results support the baseline results and are consistent with the hypothesis that around the time of patenting, firms introduce new products and drop older ones from their product portfolio.

Table 6.—Robustness of OLS-FE (One-Step) Productivity Estimates

<table>
<thead>
<tr>
<th>Coefficient on patent stock (table 4)</th>
<th>Alternative Capital Measure</th>
<th>Split Blue- and White-Collar Labor</th>
<th>Order 2 Translog</th>
<th>Real Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01516*** (0.00413)</td>
<td>0.01923*** (0.00557)</td>
<td>0.01262*** (0.00355)</td>
<td>0.01844*** (0.00351)</td>
<td>0.00834*** (0.00195)</td>
</tr>
<tr>
<td>0.11319*** (0.01994)</td>
<td>0.11138*** (0.01523)</td>
<td>0.11453*** (0.01133)</td>
<td>0.08543*** (0.01055)</td>
<td>0.06945*** (0.00624)</td>
</tr>
<tr>
<td>0.07006*** (0.00774)</td>
<td>0.07071*** (0.01157)</td>
<td>0.05494*** (0.00792)</td>
<td>0.1027*** (0.00757)</td>
<td>0.03947*** (0.00448)</td>
</tr>
</tbody>
</table>

This table reports coefficients on depreciated patent stock or patent dummy variables. The sample in row 1 is all patentees (patent stock is undefined for nonpatentees). The sample in row 2 is restricted to firms that enter without a patent and switch to being a patentee later in their life (switchers). Row 3 includes the sample of switchers in row 2 plus the sample of nonpatentees (firms that did not own or apply successfully for a patent any time during the sample period). All regressions include firm fixed effects. Regressions in rows 1 and 3 also include four-digit industry-year effects. Robust standard errors clustered at four-digit industry level (firm level in row 2) are in parentheses. * Significant at 10%. ** Significant at 5%. *** Significant at 1%.

B. Productivity Analysis

Alternative OLS-fixed effects specifications. In this section, we examine the robustness of our results regarding total factor productivity. We first examine alternative specifications of the one-step OLS-FE approach. The results are presented in table 6, with the baseline results from tables 3 and 4 presented in column 1 for comparison.

In column 2, we present results using an alternative definition of capital based on a perpetual-inventory method.29 In column 3, we allow coefficients to differ by blue-collar and white-collar labor by splitting employment into these two components. In column 4, we use a second-order translog specification (in addition to linear terms for capital and employment) and include quadratic terms as well as their cross-product. In column 5, we use log real sales as the dependent variable regressed on log real materials, log capital stock, and log employment.

28 We thank a referee for suggesting this check.

29 This uses data from both the CMF and ASM and is computed as follows. Separate stocks are computed for buildings (or structures) and machinery. Real capital stock (\( k_t \)) in any given year, for example, for machinery, is computed as \( k_t = (1 - d)k_{t-1} + I_t - R_t \), where \( d \) is an industry-year specific depreciation rate for machinery, \( I_t \) is the capital investment in machinery (deflated by an industry-year specific investment deflator for the year \( t - 1 \)), and \( R_t \) is the capitalized value of capital equipment rentals. (We thank John Haltiwanger and Hyowook Chiang for providing us the depreciation rates and rental prices.) Capital investment in machinery also includes investments in all used equipment capital (regardless of machinery versus buildings) since the relevant variables are not well populated. The capitalized value of capital equipment rentals is obtained by dividing the rental expenditure by industry-year specific equipment rental prices. This computation is done separately for buildings and machinery. Since establishments are not necessarily observed in their first year of operation, following Bahk and Gort (1993), capital stock in the first year (initial capital stock) is defined to be the book value of the assets at the end of the year deflated by an industry-year specific capital equipment deflator. This is defined separately for buildings and machinery where the decomposition is available. In case such a decomposition of year-end asset values is not available (1973, 1988–1991, and 1993–1997), all initial capital stock is assigned to machinery. If an establishment is not observed every year, following Olley and Pakes (1996), gross investment is imputed linearly (\( \hat{I}_t = 0.5I_t + I_{t-1}(k-1) \)), where \( \hat{I}_t \) is the imputed investment for period \( t = k + 1 \) to \( t = 1 \) and \( I_t \) is the actual observed investment in period \( t \). We did not use this method for the baseline because most firms in our sample were observed only during the census years, and thus required significant imputations.
As seen in table 6, the baseline results are robust to these alternative specifications. The coefficient tends to be somewhat higher in the more flexible translog specification. As expected (see footnote 10), the magnitude of the estimates is smaller but still significant for the real sales specification in column 5.

Finally, in unreported results, we allowed coefficients on inputs to vary by SIC two-digit industry, and the baseline results remained robust.

**Arellano-Blundell-Bond GMM estimators.** Next, we attempted to control for endogeneity of inputs and patent stock in the baseline specification (2) corresponding to the OLS-FE results in table 3. We used a GMM approach based on Arellano and Bond (1991) and Blundell and Bond (1998). In contrast to industry-year effects used in the baseline table 3, for reasons of computational feasibility, we used common-year effects.30

The results are presented in table 7. Column 1 reports the baseline (table 3) OLS-FE results. In the “difference” GMM approach proposed by Arellano and Bond (1991), fixed effects are eliminated by first differencing; lagged levels on the right-hand side variables are used as instruments for the first differenced variable. In our application, because all of the right-hand side variables are considered endogenous, lags of order 2 or more are potentially valid instruments. In column 2, we use order 2 lags as instruments for first differences: \(X_{i,t-2} \) is used to instrument the first difference \(X_{it} - X_{i,t-1} \), where \(X \) denotes the vector of endogenous variables (capital, employment, and patent stock). In column 3, lags of order 2 and 3 are used. Under the “system” GMM approach proposed by Blundell and Bond (1998), the moments in the “difference GMM” approach are augmented by another set of moment conditions (which are valid if a stationarity assumption for initial conditions holds). In particular, equations in levels are also estimated, and the endogenous variables in levels are instrumented using suitably lagged first differences. In column 4, first differences are instrumented with order 2 lags of the levels and levels are instrumented with order 2 lag of the first differences. In column 5, first differences are instrumented with order 2 and order 3 lags of the levels and levels are instrumented with order 2 and order 3 lags of the first differences. \(m2 \) is a test for second-order serial correlation proposed by Arellano and Bond (1991). In columns 2–5, a two-step GMM estimator is used. Robust Windmeijer (2005) corrected standard errors are reported in parentheses. In column 1, standard errors are clustered at the four-digit industry level.

In column 1 reports OLS fixed effects estimates (corresponding to results in table 4) for comparison; it includes firm fixed effects and four-digit industry-year effects. Columns 2–5 include year dummies (unreported).

In “difference” GMM (columns 2 and 3), first differences of capital, labor, and patent stock are instrumented using suitable lags of the same variables in levels. In column 2, a single lag of order 2 is used; in column 3, lags of order 2 and order 3 are used. System GMM extends the set of moment conditions used in difference GMM. In particular, equations in levels are also estimated, using suitably lagged first differences as instruments. In column 4, first differences are instrumented with order 2 lag of the levels and levels are instrumented with order 2 lag of the first differences. In column 5, first differences are instrumented with order 2 and order 3 lags of the levels and levels are instrumented with order 2 and order 3 lags of the first differences.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>OLS-FE</td>
<td>Difference</td>
<td>Difference</td>
<td>System</td>
<td>System</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GMM (Lag 2)</td>
<td>GMM (Lags 2, 3)</td>
<td>GMM (Lag 2)</td>
<td>GMM (Lags 2, 3)</td>
</tr>
<tr>
<td>Capital</td>
<td>0.16769***</td>
<td>2.08765**</td>
<td>0.81726***</td>
<td>0.57061***</td>
<td>0.57060***</td>
</tr>
<tr>
<td></td>
<td>(0.00934)</td>
<td>(1.00861)</td>
<td>(0.13406)</td>
<td>(0.01778)</td>
<td>(0.01771)</td>
</tr>
<tr>
<td>Employment</td>
<td>0.77485***</td>
<td>1.25678**</td>
<td>0.57624***</td>
<td>0.31613***</td>
<td>0.31477***</td>
</tr>
<tr>
<td></td>
<td>(0.01173)</td>
<td>(0.54302)</td>
<td>(0.09359)</td>
<td>(0.02905)</td>
<td>(0.02902)</td>
</tr>
<tr>
<td>Patent stock</td>
<td>0.01516***</td>
<td>1.06069*</td>
<td>0.19208*</td>
<td>0.07324***</td>
<td>0.07396***</td>
</tr>
<tr>
<td></td>
<td>(0.00413)</td>
<td>(0.70884)</td>
<td>(0.09894)</td>
<td>(0.00710)</td>
<td>(0.00705)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>66,423</td>
<td>66,423</td>
<td>66,423</td>
<td>66,423</td>
<td>66,423</td>
</tr>
<tr>
<td>Hansen p-value</td>
<td>NA</td>
<td>NA</td>
<td>0.0003</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>m2 p-value</td>
<td>0.187</td>
<td>0.937</td>
<td>0.689</td>
<td>0.689</td>
<td>0.685</td>
</tr>
</tbody>
</table>

30 We also checked some specifications for robustness to allowing for different coefficients by SIC two-digit industry, and found the results generally consistent with those presented here.

As seen in table 6, the baseline results are robust to these alternative specifications. The coefficient tends to be somewhat higher in the more flexible translog specification. As expected (see footnote 10), the magnitude of the estimates is smaller but still significant for the real sales specification in column 5.

Finally, in unreported results, we allowed coefficients on inputs to vary by SIC two-digit industry, and the baseline results remained robust.

**Arellano-Blundell-Bond GMM estimators.** Next, we attempted to control for endogeneity of inputs and patent stock in the baseline specification (2) corresponding to the OLS-FE results in table 3. We used a GMM approach based on Arellano and Bond (1991) and Blundell and Bond (1998). In contrast to industry-year effects used in the baseline table 3, for reasons of computational feasibility, we used common-year effects.30

The results are presented in table 7. Column 1 reports the baseline (table 3) OLS-FE results. In the “difference” GMM approach proposed by Arellano and Bond (1991), fixed effects are eliminated by first differencing; lagged levels on the right-hand side variables are used as instruments for the first differenced variable. In our application, because all of the right-hand side variables are considered endogenous, lags of order 2 or more are potentially valid instruments. In column 2, we use order 2 lags as instruments for first differences: \(X_{i,t-2} \) is used to instrument the first difference \(X_{it} - X_{i,t-1} \), where \(X \) denotes the vector of endogenous variables (capital, employment, and patent stock). In column 3, lags of order 2 and 3 are used. Under the “system” GMM approach proposed by Blundell and Bond (1998), the moments in the “difference GMM” approach are augmented by another set of moment conditions (which are valid if a stationarity assumption for initial conditions holds). In particular, equations in levels are also estimated, and the endogenous variables in levels are instrumented using suitably lagged first differences. In column 4, first differences are instrumented with order 2 lags of the levels and levels are instrumented with order 2 lag of the first differences. In column 5, first differences are instrumented with order 2 and order 3 lags of the levels and levels are instrumented with order 2 and order 3 lags of the first differences. \(m2 \) is a test for second-order serial correlation proposed by Arellano and Bond (1991). In columns 2–5, a two-step GMM estimator is used. Robust Windmeijer (2005) corrected standard errors are reported in parentheses. In column 1, standard errors are clustered at the four-digit industry level. *Significant at 10%. **Significant at 5%. ***Significant at 1%.
to the GMM estimates. Although the specification tests fail overall, the larger magnitudes of the GMM estimate of capital and patent stock coefficients suggest that the link between patenting and productivity observed in the OLS-FE specifications is not driven by endogeneity of input variables or patent stocks.

Other checks. One source of increase in measured productivity for patentees could be an increase in markups following the introduction of new products. Recall that our TFP measures subtract a weighted average of inputs from deflated real value added. Because the deflator is common across firms in an industry, firms that have a higher price relative to the industry would appear to have higher measured productivity (Klette & Griliches, 1996). To check this, we examined a proxy for markup using the net margin drawn from Aghion et al. (2005). We define the net margin as

\[
\text{Margin} = \frac{\text{Sales} - \text{materials cost} - \text{labour cost} - \text{financial cost}}{\text{Sales}},
\]

where financial cost is measured as the product of cost of capital and total assets. We also checked robustness to using value added in the denominator; any differences between the sales and value-added measure would indicate if there were potential savings or losses on material costs. Results in the supplementary appendix table A2 present a mixed picture. The elasticity of margin on sales with respect to patent stock is surprisingly negative (in column 1), suggesting a decline in margins with an increase in patent stock. However, the before-and-after effects estimated in column 2 for first-time patentees is positive and significant. The difference-in-differences effects (column 3) are also positive but small and statistically insignificant. With the measure of margin on value added, all effects are positive and significant (columns 4, 5, and 6). Thus, it appears that there may be some, though not conclusive, evidence that higher markups associated with patenting may be driving the observed increases in productivity.

C. Endogeneity from Demand Shocks

We are careful not to ascribe a causal interpretation to our results. We interpret them as consistent with patents capturing innovative activity within the firm. This interpretation could be biased if industry-specific demand shocks cause both the increase in firm size as well as increased patenting activity. Then the observed association between size and patenting may not be a reflection of innovative activity in the firm but rather of the demand shock. Our baseline specifications in table 3 and panel B of table 4 include SIC four-digit industry-year effects so that they control for demand shocks at a detailed level. In this section, we undertake a series of additional tests to check if demand pull explains a significant part of the observed correlation between firm size and patenting.

Before turning to these tests, note that we do not identify the underlying causes of innovation, and hence do not rule out a role for demand (latent or explicit) in spurring innovation. The demand-pull channel for innovation was first proposed by Schmookler (1966) and is part of some models of endogenous technological change (for example, Acemoglu & Linn 2004). It is plausible that an increase in latent or explicit demand for certain products spurs innovation by some firms (that also patent) and that these innovating firms grow (relative to noninnovating industry peers) by capturing share through the introduction of new or better products. In this case, the link between patenting and size increases is still meaningful; although they are triggered by the demand shock, the correlation between size changes and changes in patenting truly reflects innovation in the patenting firms. Thus, what we try to rule out below is the effect of demand shocks that affect nonpatenting firms as well, not omitted idiosyncratic ability to exploit demand shocks (which we see as synonymous with innovative ability).

31 Two points relating to the instrument matrix are noteworthy. First, following the approach proposed in Holtz-Eakin, Newey, and Rosen (1988), missing observations are substituted with zeros so that there is no loss of data from missing observations. Second, we collapse the instrument matrix as discussed in Roodman (2006), as this significantly improved the computational speed with little effect on the estimated coefficients. Thus, while the full instrument matrix has T-2 columns even when a single lag is used under difference GMM, once the instrument matrix is collapsed, the number of columns in the instrument matrix equals the number of lags used in the difference GMM specification and equals the number of lags + 1 in the system GMM specification. Thus, the Hansen-Sargan overidentification test is applicable for the single lag difference GMM specification in column 2 (which is just identified), but applicable for the other GMM columns 3, 4, and 5. In Blundell and Bond (2000), the discussion of the production function assumes errors are AR1 in the annual data, and hence they run specifications using a dynamic (common-factor) specification. Because we use quinquennial data, we assumed zero first-order serial correlation in errors (excluding fixed effects) and used the Cobb-Douglas specification directly. Note that in all specifications for the one-step OLS-FE (one-step) in table 7, the error terms pass the serial correlation test. While this test may be invalid given the failure of the overidentification tests, the results are not inconsistent with our assumption of no serial correlation in the error term (net of the fixed effect) in the quinquennial data.

32 Following Aghion et al. (2005), the cost of capital is assumed to be 0.085 for all firms and time periods. Checks with a range of numbers around 0.085 yielded similar results.

33 We also attempted to directly examine physical productivity using data on physical quantities produced, which is available in the census data. However, in order to regress physical quantity output on inputs, we needed to restrict attention to firms producing a single product (as input breakdowns for each product are not available). This limits observations to predominantly homogeneous product industries (see Foster, Haltiwanger, & Syverson, 2008). For these observations, the number of patentees was extremely limited, and hence we were unable to disclose these results. Even using a sample of firms with at least 80% of output in one product code, the sample size was very small. Using an OLS-fixed-effects specification as in equation (2) on this sample, we found a positive but insignificant effect on physical productivity (results not disclosed).

34 We thank a referee for pointing out this interpretation and suggesting many of the checks we use in this section.

35 For example, lagged patent stock used as instruments in section VIII C could be correlated with an omitted ability to exploit demand shocks. Also, even detailed industry and location-year effects (in section VIII C) would not control for this idiosyncratic ability to exploit demand shocks.
GMM analysis of baseline effects. A popular approach to controlling for the effects of contemporaneous demand shocks is to use the panel GMM (IV) methodology by Arellano and Bond (1991) and Blundell and Bond (1998). m2 is a test for second-order serial correlation proposed by Arellano and Bond.

Link to patent quality. If the observed increases in size are indeed linked to patenting, it is reasonable to expect a correlation between the magnitude of size increases and the quality of the patents. In order to test this, we adopt the same specification as in equation (3), but add an interaction between the patent ownership dummy and one of two measures of the quality of the first patents filed by the firm. These quality measures use information on future citations received by those patents. Assuming future citations are uncorrelated with current demand shocks, a positive coefficient on the interaction term would suggest that observed changes are driven by real innovation captured by the patent quality measure rather than being artifacts of unobserved demand shocks.

We considered two measures of quality: log total number of forward citations to the patents and average number of citations per patent. The results are presented in table 9. For all four measures of size, the interaction of the patent ownership dummy and the quality variable is significantly positive for both measures of patent quality.37 Thus, size changes associated with initial patenting are correlated with the quality of patents filed. These results are consistent with the interpretation that size changes accompanying patenting are related to innovation at the firm level. It should be noted, however, that this analysis is not conclusive evidence against a role for demand-pull innovation. It is possible that larger demand shocks could lead to bigger or higher-quality innovations, so that future citations are related to current unobserved demand shocks. The results are also consistent with findings in the literature that citation-weighted patent counts are informative about patent quality (Hall et al., 2005).

Controlling for more specific demand shocks. Our baseline specifications for table 3 and panel B of table 4 include four-digit industry-year effects. We performed additional analyses that attempted to control for more disaggregated demand shocks.

First, we disaggregated log output for each firm to the SIC five-digit industry level and then regressed this on the patent stock for the firm (analogous to table 3). Results are in row 1 of supplementary appendix table A5. Although the coefficient is positive and significant, it is considerably attenuated compared to that in table 3. The reason for the attenuation could be the inconsistency between the measures on the two sides of the specification. The output measure on the left-hand side inputs were also instrumented, is discussed in detail in section VIII B. The elasticity of gross number of products is positive and significant in the difference specifications and negative and significant in the system GMM specifications. However, both of the specification tests are strongly rejected for system GMM. The results are reversed with the net number of products; coefficients are positive for system and negative for difference GMM.

36 Results for other variables are presented in the supplementary appendix table A3. For the other variables, too, we find that specifications fail one or both of the specification tests. In brief, we find the following. Skill intensity elasticity is positive but insignificant in the difference specifications, but positive and significant (and considerably higher than in table 3) in the system specifications. Capital intensity elasticity is positive in all specifications but seems implausibly higher than the baseline (table 3) results in the difference specifications. Labor productivity elasticity is implausibly higher than baseline effects in the difference specifications and about three times the baseline effects in the system GMM specifications. The Solow residual is negative and insignificant in the difference specifications, but positive and significant under system GMM. The OLS-FE measure, where
is disaggregated to the five-digit level. However, it is not possible to meaningfully disaggregate the firm patent stock by five-digit industry, so the right-hand side is total firm patent stock. Then, if the firm expands in one SIC five-digit sector following patenting but not in the other SIC five-digit sectors, we would observe a more attenuated elasticity estimate. 38

One potential weakness with using SIC four-digit industry-year effects is that for diversified firms, the SIC four-digit industry-year effects may not capture shocks to other subsectors important to the firm. To address this problem, we restricted our sample to firms where the share of the predominant four-digit sector is at least 80% of the total output. The results are presented in rows 2 and 3 of supplementary appendix table A5. We find the baseline effects are quite robust to restricting attention to this sample. 39

Finally, we rekindled the analysis used for results in panel B of table 4b allowing for region (defined as the Standard Metropolitan Statistical Area)-SIC 2 digit industry-year effects. 40 The results in the last row of table A5 are similar to, though somewhat larger than, the baseline.

To summarize, although none of the tests in this section provides conclusive evidence, taken together they suggest that contemporaneous demand shocks may not be driving all of the baseline results.

D. Other Robustness Checks

In this section, we discuss additional robustness checks we performed. The results are presented in the supplementary appendix.

38 As a concrete example, for a firm operating in four sectors, a 10% increase in one sector’s output following a 10% increase in patent stock would result in an estimate of 0.25, even if the expanding sector had a disproportionate share of total output. If some SIC five-digit sectors contract while others expand within a firm, the estimated overall effect could be close to zero, even in cases where there is significant overall output growth.

39 Results from analysis of all variables for the sample of specialized firms are presented in table A6. Qualitatively, all the baseline results are robust to this check. Foster et al. (2008) use a similar restriction to define specialized firms.

40 More detailed industry-region effects were computationally infeasible. Also, we explored specifications using seven-digit product code year effects, but found it computationally infeasible to include these effects along with firm fixed effects.

Measurement error from acquisitions. Because all our analysis is at the firm rather than establishment level, it is possible that the results for first-time patentees (see section VII) are affected by acquisition of a patent-owning firm by a non-patentee. This could account for the increase in firm size and firm scope contemporaneous with the patenting event.

To check this, we excluded firms that appeared to acquire any firm in a five-year period before, and including, the year they switched status, and we repeated the analysis in table 4. These results are presented in supplementary appendix table A7. As a much stricter test, the sample was restricted to firms that remained single-establishment firms throughout the sample period, thus avoiding all potential acquisition-related measurement errors. These results are presented in supplementary appendix table A8. 41 We find that the qualitative conclusions as well as the magnitude of the effects are similar to our baseline, suggesting that our results in section VII are not biased by measurement errors induced by acquisitions.

Comparison of results for public and private firms. An important novelty of our paper is the coverage of private firms, which are typically smaller than publicly listed firms. Thus, it is interesting to compare results using separate subsamples of public and private firms. We defined a firm to be public in any given year if it was matched to a Compustat firm in the Compustat-BR Bridge. The results corresponding to tables 3 and 4, for the sample of public (listed) firms are presented in supplementary appendix table A9, and for private (unlisted) firms in table A10. Not surprisingly, the number of observations is much smaller in the public firm sample, especially for the subsample of switching firms (corresponding to specifications in panel A of table 4). 42 Results in column 1 suggest that elasticity of size and firm scope variables with respect to patent stock is in fact somewhat higher for the sample of

41 In an (unreported) test, we dropped switching firms that have more than five patents in the year they switch status to patentee. While a large number of initial patents could arise because some firms do apply for more than five patents on their first try, this could also arise from a nonpatentee firm acquiring a patenting firm that had built up a portfolio of patents over time. We found our results robust to this test.

42 Switching status is much more likely for smaller or younger firms that are less likely to be listed. Also, as Kortum and Lerner (2000) suggest, firms may patent at the venture capital stage, prior to listing.
public firms. However, the factor intensity and productivity
results are not significant in this sample.

The before-and-after results corresponding to panel A of
table 4 suggest big increases in size around the patenting
event, with the highest increases for capital, output, value
added, and employment (in that order). The skill intensity
results are not significant, but capital intensity shows a big
increase. We find large increases in labor productivity and in
scope measures (both gross and net number of products) but
no increase in the Solow TFP residual measure. The OLS-FE
measure is significant overall but not around the switch.

The difference-in-difference results (corresponding to
panel B of table 4) are qualitatively similar to the baseline
results for size and scope variables. However, except for size
changes for switchers overall, none of the other effects are
significant. The results for private firms in table A10 are qual-
itatively and quantitatively much closer to the baseline, which
is not surprising, as the overall baseline sample is dominated
by the private firms. The biggest differences between the pri-
ivate and public sample are in the difference-in-differences
estimates around the switch, which are larger in magnitude
and consistently higher in statistical significance in the private
firm sample.

Overall the differences between the results using the full
sample and those using only the sample of public firms are
consistent with the smaller sample size of public firms. For
the firm size and scope variables, results using the sample of
public firms agree with the full sample results, but the effects
are generally less significant. For the other variables, we do
not see any significant changes in the sample of public firms,
whereas we find statistically significant changes with the full
sample.

These results suggest that using only a sample of public
firms would suffice to examine a range of questions; thus, our
results with the full sample are largely consistent with earlier
work undertaken using mostly public firm data (Bound et al.,
1984; Scherer, 1983). A sample of private firms, however,
would be particularly useful to study the growth or other
effects of innovation on small or young firms.

Other checks. Here we discuss a number of additional
(unreported) checks of the results that we undertook.

We checked if the results were being driven by a few indus-
tries, or a few years, by redoing the analysis excluding one
SIC two-digit industry, and one year, at a time. The results
were very similar to the baseline specifications. To test for the
influence of outliers, we reran specifications in tables 3 and 4,
excluding the top and bottom 0.5 percentile of the dependent
variable. The results were virtually identical for table 3 and
panel B of table 4, and qualitatively very similar for panel
A of table 4 (some magnitudes were slightly lower, but the
signs and significance levels remained the same).

Furthermore, for the sample of switchers, we tested if the
distribution of changes in some of the key variables (output,
gross number of products, and Solow measure of produc-
tivity) around the first patenting event was systematically
different from the distribution of changes at other points. We
used two-sample Wilcoxon rank-sum and two-sample equal-
ity of median tests. Consistent with the results in panel A of
table 4, the tests confirmed that the distribution of size jumps
around initial patenting stochastically dominates the distri-
bution of size jumps at other points. The same results were
obtained for the gross number of products, but not with the
Solow productivity measure. These tests provide additional
confirmation that a few outliers are not driving the average
changes in these variables.

The literature on innovations makes a distinction between
product innovations and process innovations. Thus, it would
be interesting to see if patents associated with process and
product innovations have different types of effects. Unfortu-
nately, there is no information in the NBER Patent Data
that helps directly classify patents into these two categories.
Using data on the titles of the patents and a keyword search
procedure, we made a rough classification of patents into
process and product patents. Specifically, we defined a pro-
cess patent dummy variable as equal to 1 for patent filings
where the term process (or various synonyms for process)
appeared and synonyms for product did not appear.43 With
this approach, only about 10% of the patents were classified
as process patents. We found that almost all of the effects in
table 3 remain unchanged for the sample of product patents,
but were smaller in magnitude and statistically insignificant
for process patents. Though crude, this check is consistent
with the finding in the literature that patents are used more to
protect product innovations (Levin et al., 1987).

We also tried to exclude subsidiaries of foreign firms.
We were unable to do so because of data issues. We were
informed by the bureau that reliable data on whether an estab-
ishment is a subsidiary of a foreign company are not available
at this time.

In addition to these statistical tests, a visual inspection
of a small sample of single-establishment switching firms
was performed to check if these firms show an increase in
employment in the year of patenting. A majority of these
firms exhibited an increase, with a small minority displaying
substantial increases. Though not conclusive, this test is also
consistent with the results.

IX. Discussion and Conclusion

In this paper, we build a new concordance between the
NBER Patent Data and U.S. Census micro data and use it
to characterize patenting in U.S. manufacturing and exam-
line what changes happen within firms that patent. In line

43 The synonyms we used for process were method, methodology, tech-
nique, and procedure. The synonyms for product were apparatus, equip-
ment, machine, device, tool, instrument, gadget, and appliance. We also
excluded patent filings with phrases “and process,” “and method,” “and tech-
nique,” and “and procedure” from the definition of process patent dummy,
as visual inspections suggested that these were often found in apparent new
product-related patents.
with prior research (Scherer, 1983; Bound et al., 1984), we find considerable heterogeneity across industries in patenting activity. We find that patenting is concentrated among large firms, and although patenting firms are a small fraction of all firms, they account for a large proportion of economic activity. Consistent with related work in the literature (Hall et al., 2005; Bloom & Van Reenen, 2002), our analysis strongly suggests that patenting is associated with real, large, and statistically significant changes within firms. Our most robust conclusion is that patenting is associated with a significant increase in firm size. We find strong evidence that this growth is associated with the introduction of new products. We also find somewhat weaker evidence that patenting is positively correlated with firm-level productivity increases, as well as with increases in capital and skill intensity.

Griliches (1990) proposed that, ideally, we would like patent statistics “to measure and better understand the economic processes that lead to the reduction in the cost of producing existing products and the development of new products and services” (p. 1669). Our results provide strong evidence that patents capture innovation through the introduction of new products. We find a strong, positive elasticity of gross number of products with respect to patent stock and observe a considerable increase in the gross number of products for first-time patentees after their first patent application. The results also suggest that relative to nonpatentees, first-time patentees generate significantly more sales from new products introduced after their first patent application. Although we find that productivity is positively associated with patenting in many of our specifications, our productivity measures use deflated output and can be affected by changes in markup as well as reductions in input use. Hence, with the current data, we are unable to distinguish between these two causes of productivity increase and cannot conclusively answer if patents also capture pure efficiency-enhancing innovations. Taken together with the starker timing effects observed in section VII B for the gross number of products and size variables (compared to productivity), the results tentatively suggest a stronger correlation between patenting and changes in scope than with improvements in productivity. This is also consistent with the findings in Levin et al. (1987), who find that patents are viewed as more effective mechanisms for appropriating value from product than from process innovations and that secrecy was more effective for process innovations.

The finding of an important role for new product introductions in the growth of patenting firms also suggests that features of a recent class of models that explain firm heterogeneity based on differences in firm scope (such as Klette & Kortum, 2004; Bernard et al., 2006b; Nocke & Yeaple, 2006) may be useful in modeling patenting at the firm level. Our results are broadly consistent with the model of innovation presented in Klette and Kortum (2004). In this model, a firm is defined by the portfolio of goods it produces, and it innovates by extending its product line. Innovations are not directly linked to productivity; hence, assuming patents represent successful innovations, the scope (but not necessarily the productivity) of a firm increases coincident with patenting. In this model, the size of the innovative steps is positively related to measured productivity and R&D intensity. Since the arrival rate of innovations (and, hence, patenting) is a positive function of R&D intensity, ex ante differences in productivity between patenting and nonpatenting firms would be consistent with the model.44

A few caveats should be borne in mind while interpreting our results. It is worth reiterating that patenting is an endogenous decision: underlying innovation shocks affect both the dependent variables and the decision to patent. Hence, our results must not be given a causal interpretation. Rather, we interpret our results as suggesting that patents are meaningful proxies for real innovative activity within firms. Another limitation of our work is the lack of specific information on innovation-related investments by firms prior to patenting.45 Without good information on the costs, we cannot assess if the increases in size and scope that we document reflect an increase in welfare or even an increase in the value of the firm. Finally, our work does not reveal the underlying drivers of innovation within firms. Although we undertake tests to rule out potential bias from common industry demand shocks, we cannot rule out a role for (latent) demand in driving innovation.

Our analysis in this paper can be extended in a number of ways. An interesting extension would be to examine intertemporal trends in the impact of patents, especially in the context of regulatory changes (e.g., the establishment of the court of appeals in 1982 or the Hatch-Waxman Act of 1984), the large increase in patent filing in recent years (Kortum & Lerner, 1998), and recent concerns regarding patent thickets (Shapiro, 2001) and patent trolling (Barker, 2005). Examining cross-industry variations in the impact of patenting on size and scope is another interesting extension, as the importance (and prevalence) of patenting varies considerably across industries. More important, we hope that our concordance will significantly extend the number of questions that researchers will be able to analyze using the NBER Patent Data, particularly about innovation by small, young, and unlisted firms. We have addressed only the first few questions here.

44 Additional assumptions that not all innovations are patented and that patenting involves a fixed cost but provides some benefits in the form of extra protection for markups may be needed to explain ex ante size and productivity differences between patentees and nonpatentees and why all firms do not patent.
45 It is likely that for most firms, a significant part of labor and capital resources committed to innovation is captured in the operating expenses reported in the census data. However, we do not have a breakdown of the specific costs incurred by the firms in activities that led to the innovation that was patented. Moreover, data access constraints prevented us from using R&D data. For researchers interested in the R&D data, Kerr and Fu (2006) provide a concordance between census data and the NSF R&D Survey.
REFERENCES


—“Multiproduct Firms and Trade Liberalization,” Tuck School of Business, Dartmouth, working paper (December 2006b).


APPENDIX 1

Definitions of Variables

Size Measures

*Output:* For any year before 1996, real output is defined as the sum of shipments (deflated using the SIC 4-digit industry shipment deflator in the CES-NBER database developed by Bartelsman, Becker, & Gray, 2000) and the difference between year-beginning and year-ending deflated work in process and deflated finished goods inventories (the year-beginning inventory is deflated using the previous year’s industry shipment deflator and year-end inventory is deflated using the current year deflator). Following the LRD documentation, output for industries 2032, 2033, 2035, 2037, 2038, 2085, 2091, 2111, 2121, 2131, and 2141 was defined to be the sum of deflated shipments and the difference between year-beginning and year-ending deflated work in process inventories. Similarly, for industry 3731, output is defined to be the deflated shipments. For years including and after 1996, due to the unavailability of inventory data, output is simply defined to be the deflated shipments.

*Value added:* Real value added is defined as the difference between real output and real materials. Real materials are defined as the sum of deflated cost of material purchases, external contract work, fuel, and electricity. For the years before 1997, fuel and electricity are deflated using the energy deflator and the others using a materials deflator. Starting in 1997, a single materials deflator is used.

*Capital stock:* This study uses the reported year-end assets deflated by an industry-year specific capital equipment deflator developed by CES researchers John Haltiwanger and Hywook Chiang. This deflator was based on BEA estimates, and Davis, Haltiwanger, and Schuh (1996). An alternative definition of capital stock is discussed in section VIII B.

*Employment:* Measured as the total number of employees reported by each firm.

Factor Intensity Measures

*Capital intensity:* Log capital stock per worker (capital stock divided by employment)

*Skill intensity:* The ratio of white-collar to blue-collar workers (number of nonproduction workers divided by the number of production workers)

Productivity Measures

*Labor productivity:* Log output divided by employment

*TFP (Solow residual):* Defined as $\text{TFP}_\text{it} = y_{it} - \alpha_m n_{it} - \alpha_s s_{it} - \alpha_l l_{it}$, where $y_{it}$ is the log of output of firm $i$ in year $t$, $m$ is log real materials (see the definition of value added above), $s$ is the log of real depreciated capital structure stock, $e$ is the log of real depreciated capital equipment stock, and $l$ is employment. The elasticities $\alpha_m$, $\alpha_s$, and $\alpha_l$ are defined equal to the material share, capital share, and labor share of total costs in the four-digit SIC (1987) industry $j$ to which firm $i$ belongs. The factor share measures, as well as the deflators and depreciation rates used to construct the firm-specific real depreciated capital structure, and equipment stocks were taken from the NBER productivity database (Bartelsman et al., 2000).

*OLS-FE productivity measure:* The residual from an OLS firm-fixed-effects regression of log real value added on log real depreciated capital stock and log employment. Definitions of real value added, real depreciated capital stock, and employment are as given above.

Product Scope Measures

*Gross number of products:* The total number of distinct seven-digit product codes reported by the firm over its lifetime up to and including the current year

*Net number of products:* The number of products reported by the firm in the current year

Other Key Variables

*Depreciated patent stock:* The total of depreciated patents, with each patent depreciated at an annual rate of 15% (following Hall et al., 2005).

*Industry:* The study adopts 1987 four-digit SIC codes as the basis of industry definition. For most observations before 1997, the “ind” variable provided in the ASM and CM files is used to define the establishment’s primary industry of operation. However, some of the observations wrongly use the 1972 codes and are corrected per Chiang (2004) who bases these corrections on Davis et al. (1996). The magnitude and impact of these corrections are very small, though. In addition, establishments with nonexistent SIC codes (e.g., 3079) are dropped.