Competition and Innovation:
The Inverted-U Relationship Revisited

Aamir Rafique Hashmi
Department of Economics
National University of Singapore
February 15, 2011

Abstract

I re-examine the inverted-U relationship between competition and innovation (originally modeled and tested by Aghion et al. [2005]) by using data from publicly traded manufacturing firms in the US. I control for the possible endogeneity of competition by using various measures of foreign competition as instruments. I find a positive relationship between competition (as measured by the inverse of markups) and innovation (as measured by citation-weighted patents). The positive relationship is robust to many alternative assumptions and specifications. To reconcile the positive relationship in the US data with the inverted-U relationship that Aghion et al. [2005] find in the UK data, I modify their theoretical model and show that the modified model can explain both positive and inverted-U relationships. The key theoretical assumption is that the US manufacturing industries are technologically more neck-and-neck than their counterparts in the UK. There is some, though not strong, support for this assumption in the data.

JEL Classification Codes: L10; O30

*Email: aamir@nus.edu.sg; Phone: +65-6516-6547; Fax: +65-6775-2646; Webpage: aamirhashmi.com.
†I thank Diego Restuccia, Daniel Treffer and Johannes Van Biesebroeck for their guidance and Philippe Aghion, Michelle Alexopoulos, Nicholas Bloom, Rachel Griffith, Peter Howitt, Igor Livshits, Angelo Melino, Ivan Png, Carlos Rosell, Shouyong Shi, Xiaodong Zhu and audiences at the University of Toronto, the EEA meeting in Amsterdam, the FAEA meeting in Bali and the IO Reading Group meeting at the National University of Singapore for their comments. Special thanks are due to Rachel Griffith for sharing the data set and the STATA codes used by Aghion et al. [2005].
1 Introduction

Schumpeter [1950] averred that perfect competition was not the best market structure and ‘the large-scale establishment or unit of control’ was ‘the most powerful engine’ of progress. Since then, a large number of theoretical and empirical studies has explored the relationship between market structure and innovation. The Schumpeterian endogenous growth models pioneered by Aghion and Howitt [1992] formalize Schumpeter’s argument. Their original model (see Aghion and Howitt [1992]) predicts a negative monotone relationship between competition and innovation. The reason is that if innovation is driven by the expectation of higher profits then any increase in competition (that lowers profits) will reduce innovation. However, empirical works of Nickell [1996], Blundell et al. [1999], Carlin et al. 2004 and Okada 2005 find a positive relationship between competition and productivity (or innovation). In an important and influential paper, Aghion et al. [2005] (from here on, ABBGH) attempt to reconcile the Schumpeterian theory with this new evidence. They develop a simple model in the Schumpeterian tradition and derive an inverted-U relationship between competition and innovation. They test the inverted-U and other related predictions of their model by using data from publicly listed manufacturing firms in the UK and find strong empirical support for their theory.

In this paper, I replicate the empirical work of ABBGH by using a much richer dataset from publicly listed manufacturing firms in the US. I find strong evidence for a positive relationship between competition and innovation in the US data. To reconcile the evidence from the two datasets, I modify the theoretical model in ABBGH and show that the modified model can explain both the inverted-U and the positive relationships. The key assumption behind this result is that the average technology gap is lower in the country where the relationship is positive. There is some support in the data for this assumption.

The finding of a positive relationship between competition and innovation is in concordance with the other studies cited above. The policy implications of the positive relationship are profound and

---

1See Tirole [1988], chapter 10 sections 10.1 and 10.2, for a brief discussion of the theoretical models and Cohen and Levin [1989], sections 3.2 and 3.3, for a survey of the empirical studies of the Schumpeterian hypothesis. Kamien and Schwartz [1982] is a good survey of both theoretical and empirical literature up to the early 1980’s. For a more recent survey see Gilbert [2006].

2The idea of inverted-U relationship between competition and innovation goes at least as far back as Scherer [1967], who found an inverted-U relationship between market concentration and employment of scientists and engineers. In a recent paper, Makoyama [2003] also finds an inverted-U relationship between competition and growth where growth is driven by innovations.
different from those of the inverted-U. The first implication is that lower barriers to trade are better for innovation and growth. This is supported by other recent studies. For example, Bloom et al. [2011] study the effects of Chinese import competition on firm performance using a comprehensive panel dataset covering half a million firms in twelve European countries over the 1996-2007 period. They find that Chinese import competition led to increases in R&D, patenting, IT and TFP. Similarly, Gorodnichenko et al. [2010] study the effects of globalization on innovation using data from a unified survey of over 11,500 firms in 27 emerging markets. They find that greater pressure from foreign competition stimulates innovation.

The second implication is that the anti-trust policy should be very strict to promote competition and patent protection should be less generous to discourage market power. On the first point, the antitrust authorities, especially the European Competition Commission, already seem to be adopting a tougher stance in their anti-trust rulings. The Microsoft Case is a famous example. In February 2008 the Commission imposed a record penalty payment of 899 million euros on Microsoft for non-compliance with the Commission’s earlier rulings. On the second point, there is an ongoing debate in policy circles on how to best protect the intellectual property without discouraging further innovation. Boldrin and Levine [2008], for example, take an extreme position on the issue and argue that there should be no patent protection at all.

Clearly, the kind of broad evidence that this study provides is not sufficient to guide the complex industry-specific trade and anti-trust policies. Nonetheless, together with the evidence from a broad spectrum of studies with similar conclusions, the message of the present study is that more competition encourages creativity and innovation.

This paper is related to a large literature in Industrial Organization (IO) on the relationship between competition and innovation. There are at least three excellent and comprehensive surveys of this literature: Kamien and Schwartz [1982], Cohen and Levin [1989] and Gilbert [2006]. I contribute to this literature by providing further evidence of the positive role that competition plays in promoting innovation. Another contribution of the present study is to provide a coherent theoretical explanation for the apparently conflicting empirical results from the UK and the US datasets.

This paper, like ABBGH and so many other papers before that, is written in the old IO tradition of ‘structure, conduct and performance’. The objective is to find broad-brush evidence on the nature of the relationship between competition and innovation from cross-industry variation in the data.
This kind of work is complementary to the new empirical IO studies that focus on a single industry, estimate structural models and use counter-factual experiments to understand how variations in market structure affect innovative activity. Examples of the new IO studies on this issue include Goettler and Gordon [2009] and Hashmi and Van Biesebroeck [2010].

The rest of the paper is organized as follows. I provide a quick review of the theoretical model in ABBGH in Section 2. In Section 3 I outline my empirical strategy. I compare my empirical results with ABBGH’s in Section 4. I do a number of robustness checks in Section 5. In Section 6 I attempt to reconcile the apparently conflicting empirical evidence from the UK and the US datasets. I conclude in Section 7. Some details on the construction of empirical variables are in the appendix.

2 A Quick Review of the Model in ABBGH

In this section I provide a quick review of the theoretical model in ABBGH. The purpose is to set the stage for the empirical section. I postpone a formal exposition of the model until Section 6.

Consider an economy with a continuum of duopolies. There are two types of duopolies. The first type are with technologically equal firms. We shall call them Neck-and-Neck (NN) industries. The second type are with technologically unequal firms: one firm (the leader) is technologically ahead of the other (the laggard). We shall call them Leader-Laggard (LL) industries. Firms invest in innovation to improve their chances of getting ahead of their rivals.

As the level of product market competition increases in an industry, there are two opposing effects on the level of innovation by the firms. The first is the Schumpeterian effect: more competition reduces profits and hence there is less incentive to innovate. Due to the Schumpeterian effect, more competition reduces innovation. The second is the escape-competition effect: a firm needs to innovate to escape competition from the rival since profits from being a leader are higher than the profits from being a neck-and-neck firm. Due to the escape-competition effect, more competition increases innovation. The Schumpeterian effect was present in the original Schumpeterian growth model of Aghion and Howitt [1992]. The escape-competition effect is the innovation of ABBGH.

ABBGH show that the escape-competition effect dominates the Schumpeterian effect in NN industries (we shall see this formally in Section 6). Hence in these industries an increase in competition increases innovation. On the other hand, the Schumpeterian effect dominates the escape-
competition effect in the LL industries. Hence in these industries an increase in competition reduces innovation.

To derive the inverted-U relationship between competition and innovation, first consider a low level of competition. If competition is low, firms in NN industries innovate less and hence stay neck-and-neck. The firms in the LL industries innovate more and the laggards catch up with the leaders (leaders do not innovate by assumption). This changes the LL industries into NN industries. The overall effect is that when competition is low there are more NN industries. Since firms in these industries innovate less when competition is low, the overall level of innovation in the economy is low.

Next consider a high level of competition. If competition is high, firms in NN industries innovate more. This changes the NN industries into LL industries. On the other hand, the firms in LL industries innovate less and these industries remain LL. Overall when competition is high there are more LL industries. The firms in these industries innovate less and hence the level of innovation in the economy is low.

Finally consider an intermediate level of competition. In this situation firms in both LL and NN industries innovate. In the steady state both type of industries are present and the overall level of innovation is high.

To sum up, when competition is too low or too high the level of innovation is low. When competition is in the intermediate range the level of innovation is high. Combining these results, one gets an inverted-U relationship between product-market competition and innovation. This gives us the following testable prediction:

Prediction 1 There is an inverted-U relationship between product market competition and innovation. [c.f. Proposition 2, ABBGH p.715]

The above description of the model also implies that when competition is low, most of the industries are NN in the steady state. As competition increases the fraction of NN industries declines and that of LL industries increases. In other words, as competition increases, the average technology gap within industries increases. This gives us the second testable prediction:

Prediction 2 As competition increases, distribution of industries moves from NN to LL and average technology gap increases. [c.f. Proposition 4, ABBGH p.717]
Another prediction of the model is that in relatively more NN industries the escape-competition effect will be stronger and hence the peak of the inverted-U will be higher. We can state this as our third testable prediction.

**Prediction 3** In more neck-and-neck industries, the peak of inverted-U is higher and occurs at a higher level of competition. [c.f. Proposition 5, ABBGH p.717]

I test these predictions in Section 4 by using both the UK and the US datasets. But first I describe the US data and my empirical strategy in the next section.

## 3 Empirical Methodology

I have divided this section into three subsections. In Sub-section 3.1 I describe the US dataset and compare it with the UK dataset used by ABBGH. In Sub-section 3.2 I present the econometric model. In Sub-section 3.3 I talk about identification of the relationships of interest.

### 3.1 Data

I have compiled the dataset for this study from six sources. The accounting data to construct the measures of competition and productivity come from the Standard and Poor’s COMPUSTAT database. The patent and citation data to construct the measure of innovation come from NBER’s Patent Database. I use NBER’s Productivity database to construct industry-level price indices. I use CPI statistics from the US Department of Labor. The data on the measures of international competition that I use as instruments come from Peter Schott’s online international trade database. Finally, the dataset used by ABBGH was sent to me by Rachel Griffith.

There are three key empirical variables used in estimation: competition, innovation and technology gap. There are two other variables used as instruments. In Appendix A I describe the construction of each of these variables. Following ABBGH, I construct industry-level variables from the original firm-level data. Owing to the much richer US data (there are 7,789 publicly traded manufacturing firms in my sample compared to 311 in ABBGH’s sample), I am able to work at a more disaggregated level. All the empirical results on the US industries reported in Section 4 are based on data at the three-digit industry level. Due to the smaller number of firms, ABBGH worked with two-digit data. However, as I report in Section 5 my main results remain intact when I aggregate the US data to two-digit level.
I present a summary comparison of the US data with the UK data used by ABBGH in Table 1. All data are annual. The time coverage of the US dataset is from 1976 to 2001 (26 years). There are 116 three-digit industries with 2,756 industry-year observations. The UK dataset used by ABBGH covers the period from 1974 to 1994 (21 years). There are 17 two-digit industries with 354 industry-year observations.

First, consider the statistics on competition. Since competition is defined as one minus the Lerner’s index (see Appendix A for details), a lower level of competition implies higher markups. By this measure of competition, the US industries, on the average, are less competitive than their UK counterparts. The average value of competition variable is 0.76 for the US industries compared to 0.95 for the UK industries. However, this measure of competition is much more dispersed for the US industries. The standard deviation of competition variable is 0.11 for the US industries and 0.02 for the UK industries. The higher standard deviation in the US data is not because of the higher level of disaggregation. For example, the standard deviation of the competition variable for the two-digit US data (not reported in the table) is 0.10, which is still five times the corresponding number for the UK data.

Next consider the statistics on citation-weighted patents (CWPs). For this variable the summary statistics are similar for the two datasets. The average number of CWPs in the US data is slightly lower than that in the UK data (5.53 versus 6.66), though the standard deviation is slightly higher (9.98 versus 8.43). There are 13.6% (374 out of 2,756) observations in the US data with zero CWPs compared to 13.0% (46 out of 354) observations in the UK data.

Finally, consider the statistics on the technology gap. The average technology gap in the two datasets is the same (i.e. 0.49), though it is slightly more dispersed in the US data (the standard deviation in the US data is 0.20 versus 0.16 in the UK data).

3.2 Econometric Model

Since the purpose of empirical analysis in this paper is to replicate ABBGH’s results using the US data, I follow ABBGH closely in my empirical methodology except for one modification: instead of the Poisson regression model used by ABBGH, I use a Negative Binomial (NB) model. This modification is motivated by the well-known fact that patents, which I use as the measure of

innovative activity, generally do not satisfy the Poisson assumption of equal mean and variance. The empirical results below support this modification and the over-dispersion parameter is highly significant is almost all specifications. However, the different results that I obtain from the US data are not due to the use of the NB model. The general conclusions are the same with the Poisson model. I shall say more about this in Section [5]

To test Prediction 1, I proceed as follows. Let \( x_i \) be the vector of explanatory variables and let \( y_i \) denote the citation-weighted patents. Then the distribution of \( y_i \) is given by

\[
f(y_i|x_i, \nu_i) = \frac{e^{-\lambda_i \nu_i} (\lambda_i \nu_i)^{y_i}}{y_i!}, \tag{1}
\]

where \( \lambda_i = e^{x_i \beta} \) is the conditional mean and \( \nu_i \) is the error term. I assume \( \nu \) to follow a Gamma distribution with mean 1 and variance \( \alpha \) i.e.

\[
g(\nu) = \frac{\nu^{1-\alpha} e^{-\frac{\nu}{\alpha}}}{\alpha^{\frac{1}{\alpha}} \Gamma(\frac{1}{\alpha})}. \tag{2}
\]

With these assumptions on the distribution of \( \nu \), the conditional variance of \( y_i \) is given by \( \lambda_i(1 + \alpha \lambda_i) \). Note that if \( \alpha = 0 \), the NB model reduces to the standard Poisson model. The parameter \( \alpha \) is also called the over-dispersion parameter. If \( \alpha \) is statistically different from zero in the estimated model then the NB regression model provides a better approximation to the data than the Poisson regression model.

The specification of the conditional mean below is motivated by the hypothesis of an inverted-U (or more generally, a quadratic) relationship between competition and innovation. Denoting citation-weighted patents by \( y \), competition by \( c \) and the vector of other controls by \( z \), I model the log of conditional mean as:

\[
\ln(y_{jt}) = \alpha_0 + \alpha_1 c_{jt} + \alpha_2 c_{jt}^2 + \delta z + \epsilon_{jt}. \tag{3}
\]

Under the hypothesis of inverted-U, \( \alpha_1 \) is expected to be positive and \( \alpha_2 \) negative. I estimate the parameters of the model by using the Maximum Likelihood method.

To test Prediction 2, I regress average technology gap (defined as the average difference between the total factor productivity (TFP) of individual firms and the TFP of the industry leader) on competition.

---

5The unconditional mean and variance of citation-weighted patents are 5.5 and 99.6 in the US data and 6.7 and 71.1 in the UK data.
To test Prediction 3, I use the same negative binomial model as the one used to test Prediction 1 above. However, I modify the specification of conditional mean to allow for the interaction between competition and technology gap. The conditional mean is now defined as:

\[
\ln(y_{jt}) = \alpha_0 + \alpha_1 c_{jt} + \alpha_2 c_{jt}^2 + \beta_1 (d_m \cdot c_{jt}) + \beta_2 (d_m \cdot c_{jt}^2) + \delta z + \epsilon_{jt},
\]

(4)

where \(d_m\) is a dummy variable that takes the value 1 if the average technology gap in the industry is less than the median technology gap in the data, and 0 otherwise. If the inverted-U relationship in more neck-and-neck industries is steeper, as implied by Prediction 3, then we would expect \(\beta_1 > 0\) and \(\beta_2 < 0\).

To sum up, I use exactly the same econometric model as ABBGH do except for one difference: I use the NB model while ABBGH use the Poisson model. This modification is strongly supported by both the US and the UK datasets. More importantly, the different empirical results that I obtain for the US data do not depend on my use of the NB model.

3.3 Identification

The primary source of identification of the relationship between competition and innovation is variation in the level of competition across industries and over time. As the summary statistics in Table 1 suggest, there is a lot more variation in our measure of competition in the US data than in the UK data. This variation is helpful in identifying the relationship of interest. Indeed, as the empirical results below show, the parameter estimates are generally more precise with the US data. However, there are a few important identification issues that need to be addressed.

Firstly, if both innovation and competition are related to the business cycle or to some other variable that changes over time, then we might overestimate the relationship between innovation and competition. For example, suppose that there is no causal relationship between competition and innovation but both are pro-cyclical. In this case, we shall estimate a spurious positive relationship between the two variables, although their is no causal relationship between them by assumption. To overcome this problem, ABBGH use the year fixed effects and I follow suit.

Secondly, different industries have different propensities to innovate. Hence it is possible to see a lot of variation in innovative activity across industries that may not be due to variation in competition. Instead, as noted by ABBGH, it could be the result of the ‘other institutional features of the industry’. These may include technological opportunities and appropriability conditions in the industry. Once again following ABBGH, I use industry fixed effects to control for this problem.
Thirdly and most importantly, there is the issue of the endogeneity of competition. ABBGH call this problem the ‘major obstacle to empirical research in this area’. If successful innovations increase market power and hence reduce competition, the estimates will be biased towards finding a more negative (or less positive) relationship between competition and innovation. ABBGH control for the endogeneity of competition by using various policy variables as instruments for competition. The policy variables are correlated with competition but are likely to be uncorrelated with the error term.

To address the issue of the endogeneity of competition, I use two measures of foreign competition as instruments for competition. The first measure is the tariff rates. If tariffs on import substitutes of the products of a certain industry decline, it is likely to reduce the markups in the industry and increase competition. The second measure is the freight rates. The idea here is that if freights are lower for some products (perhaps because they are produced in nearby countries) then the local producers of those products will face more competition compared to the producers of the products whose import substitutes have to be shipped from far off places and incur higher transportation costs. Similarly, if the freights decline over time due to improvements in shipment technology, the competition from foreign products will increase. I show in Section 4 below that both these measures of foreign competition are correlated with my measure of competition. This satisfies the first (and verifiable) requirement of good instruments that they should be correlated with the endogenous variable. To satisfy the second (but unverifiable) requirement of good instruments, we need them to be uncorrelated with the error term. To this end, I argue that these measures of foreign competition are driven by exogenous factors like international trade agreements (in the case of tariffs) and improvements in transportation and shipment technology (in the case of freights) and hence are unlikely to be directly related to industry level innovative activity.

To implement the instrumental variable strategy, I follow ABBGH and add a control function to the specifications of the conditional mean in (3) and (4). The control function consists of residuals from the regression of competition on instruments and the time and industry dummies (see equations (5) and (6) in ABBGH (p. 710) and the related discussion there).

\[ I also tried a third instrument: import penetration. The idea was that the higher the import penetration in an industry, the greater will be the level of foreign competition. However, this instrument was not correlated with my measure of product market competition and hence did not qualify as a strong instrument. \]
4 Empirical Results

4.1 Prediction 1

Before I compare the empirical results on Prediction 1 from the US data to those from the UK data, I replicate the results in Table I of ABBGH using the NB model. The purpose is to show that ABBGH’s empirical results on Prediction 1 are very similar whether one uses the Poisson model or the NB model. I report these results in Table 3. For easy reference, I have reproduced an exact image of Table I in ABBGH as Table 2. The main conclusion about the inverted-U relationship between competition and innovation remains unchanged when NB model is used (compare column (4) in Table 2 with column (4) in Table 3). The use of the NB model is supported by the highly significant over-dispersion parameter in three of the four specifications. The only case in which the over-dispersion parameter is not significant is when I use 5-year averages of the data. This is not surprising because five-year averages smooth out a lot of dispersion in the data. I conclude from the comparison of Tables 2 and 3 that ABBGH’s results on Prediction 1 are robust to the use of the NB model. Hence, there is strong evidence of an inverted-U relationship between competition and innovation in the UK data regardless of the regression model used.

I report the results on Prediction 1 from the US data in Table 4. There is clear evidence of a negative relationship between competition and innovation in the first three specifications. However, in these specifications I do not control for the possible endogeneity of competition. It is possible that the negative relationship that I estimate is the result of the fact that more innovation leads to higher profit margins and hence lowers competition. In the fourth specification, when I use the two measures of foreign competition as instruments for competition, this possibility appears to be supported by the data. The relationship between competition and innovation is now positive. The two instruments are highly significant individually as well as jointly ($F$-Stat = 30.1) in the reduced form regression of competition on the instruments and the time and industry dummies. It means that the instruments are strongly correlated with competition. The coefficient on the control function in the regression is also highly significant and has a negative sign.

The change in the relationship between competition and innovation from negative to positive when endogeneity of competition is taken care of is quite remarkable. It suggests that endogeneity of competition is indeed a serious problem in the US data. When ABBGH control for endogeneity of competition using various policy instruments, the estimated coefficients on competition and competition squared barely change (compare columns 2 and 4 in Table 2). It suggests that endogeneity
of competition is not a serious problem in the UK data that ABBGH use.

Another notable feature of Table 4 is that the over-dispersion parameter is highly significant in all four specifications. This provides strong support for the use of the NB model.

For a visual comparison of the results from ABBGH’s preferred specification (column (4) in Table 3) to those from mine (column (4) in Table 4), I plot the estimated citation-weighted patents from the two specifications against competition in Figures 1(a) and 1(b). I also add the scatter plot of the raw data in both cases. Figure 1(a) looks very similar to Figure I in ABBGH (p.706), although the former is based on a NB regression and the latter on a Poisson regression, and shows an inverted-U relationship between competition and innovation. Figure 1(b), which is based on the US data, shows a positive relationship.

4.2 Prediction 2

Prediction 2 suggests that as the level of competition increases in an industry, the average technology gap within the industry should increase. To test this prediction, I run a simple linear regression of technology gap on competition. The estimates from the ABBGH data are reported in columns (1) and (2) of Table 6. I shall only comment on results in column (2) because they are based on a regression that controls for the fixed time as well as industry effects and hence is my preferred specification. The coefficient on competition is 0.94 and statistically significant. This estimate supports Prediction 2: as product market competition increases, the average technology gap also increases.

The estimates from the US data are reported in columns (1) and (2) of Table 7. Once again my preferred specification is the one in column (2). The coefficient on competition is just -0.01 and is statistically indistinguishable from zero. Hence, there is no evidence of a positive (or negative) relationship between competition and technology gap in the US data. In other words, the US data does not support Prediction 2.

I plot the estimated technology gap from the two datasets in Figures 1(c) and 1(d).

4.3 Prediction 3

Prediction 3 suggests that the inverted-U relationship should be steeper in more neck-and-neck industries. To test this prediction I interact competition and competition squared variables with a dummy variable that takes the value 1 if the technology gap in the industry is less than the
median technology gap in the data (see [4]). All the following results in this subsection are based on the same specification of the conditional mean.

First I estimate the NB model using ABBGH’s dataset and compare the results with the original results in ABBGH based on the Poisson model. For easy reference, I have reproduced an exact image of Table III (p.719) from ABBGH as Table 5. The results from the NB model using ABBGH’s dataset are in columns (3)-(6) of Table 6. Let us just compare column (4) in Tables 5 and 6. The results are very similar and provide clear support for Prediction 3. The inverted-U relationship is steeper in more neck-and-neck (i.e. low-gap) industries.

The results in columns (3) and (4) of Table 6 are based on the specifications that do not control for the endogeneity of competition. In columns (5) and (6) I report the results after controlling for the endogeneity of competition. My preferred specification is the one in column (6). A comparison between columns (4) and (6) suggests that the estimates are very similar after we control for the endogeneity of competition. So the broad conclusion that emerges from the above estimates based on the ABBGH’s dataset is that, as suggested by Prediction 3, the inverted-U relationship is steeper for more neck-and-neck industries. I depict this in Figure 1(e), where the black line represents the results for all industries and the red (or grey, if viewed in greyscale) line depicts the results for more neck-and-neck industries. The estimated inverted-U is higher for more neck-and-neck industries.

In Table 7 columns (3)-(6) I report estimation results based on the US data to test Prediction 3. I shall only comment on the results in column (6) because they are based on my preferred specification. Here again we see that after controlling for the fixed industry effects and the endogeneity of competition, the relationship between competition and innovation is positive. Moreover, the relationship is steeper for the more neck-and-neck industries. This result partially supports Prediction 3. Although we do not see the inverted-U relationship in the US data, we do see that the positive relationship is steeper for more neck-and-neck industries. I plot this result in Figure 1(f), where again the red (or grey, if viewed in greyscale) line represents the more neck-and-neck industries.

4.4 Summary of Empirical Results

The UK data used by ABBGH support all three theoretical predictions reported in Section 2. The US data used in this study do not fully support any of the three predictions, though they partially support Prediction 3. Prediction 1 suggests an inverted-U relationship between competition and innovation but the US data imply a positive relationship. Prediction 2 suggests a positive relation-
ship between competition and technology gap but there is no statistically significant relationship between the two in the US data. Prediction 3 suggests that the inverted-U should be steeper in more neck-and-neck industries. The estimates from the US data suggest that although there is no inverted-U relationship between competition and innovation, the positive relationship that does exist is indeed steeper for the more neck-and-neck industries. Figure 1 summarizes all these results graphically.

The two key differences between the UK and the US datasets that emerge from the above analysis are the following.

1. The UK dataset suggests an inverted-U relationship between competition and innovation while the US dataset suggests a positive relationship.

2. There is a positive relationship between competition and technology gap in the UK data. There is no such relationship in the US data.

Do the different results from the US data invalidate the theoretical model in ABBGH? Or is it possible to tweak the model a bit so that it can explain the apparently conflicting results from the two datasets? I try to answer these questions in Section 6. However, before doing so I test the robustness of the empirical results from the US data in the next section.

5 Robustness Analysis

In this section, I explore whether my empirical results are robust to a number of alternative empirical assumptions. I have divided this section into three subsections, one for each prediction.

5.1 Prediction 1

The main conclusion about Prediction 1 from the US data is that there is a positive relationship between competition and innovation. The empirical results of my preferred specification to test this prediction are in column (4) of Table 4. I now explore whether this conclusion is robust to some alternative empirical assumptions. The alternative empirical assumptions that I explore include the use of 2-digit data, the use of Poisson regression model, the use of an alternative definition of competition, the use of R&D expenditure as the measure of innovation and the use of a flexible

---

7The benchmark measure of competition is one minus the average of the firm-level Lerner’s index within the industry (see Appendix). If a firm makes a loss, the firm-level Leaner’s index could be negative. If there are many
spline to approximate the relationship. The results of this exercise are reported in Table 8 and Figure 2. Column (1) of Table 8 is the same as column (4) of Table 1. This is the benchmark case.

In column (2) of Table 8, I report the results from 2-digit US data. This is interesting because ABBGH also use 2-digit level data. The numbers in column (2) show that the positive relationship is still present and equally strong in the 2-digit level data. In terms of the qualitative comparison, the only significant difference is that when 2-digit data are used, the over-dispersion parameter is not significant. This is not surprising because 2-digit data gloss over a lot of variation that is present at the 3-digit level.

In column (3) of Table 8, I report the results from a Poisson model. Once again the positive relationship is highly significant.

In column (4) of Table 8, I report the results based on an alternative measure of competition. This is important because while computing the benchmark measure of competition I restrict the firm-level Lerner’s index to be non-negative. This restriction lumps together the zero-profit firms and the loss-making firms. However, if the loss-making firms innovate much less than the zero-profit firms, the restriction may mask lower innovation at very high levels of competition and hence may hide the negatively sloping part of the inverted-U. However, when I use the alternative measure of competition, the results barely change and the positive relationship remains strong and statistically significant.

In column (5) of Table 8, I report the results based on R&D expenditures as the measure of innovation. Now the coefficient on competition is negative and on competition squared is positive. Hence the relationship appears to be U-shaped. However, as I show in Figure 2(a), for most of the observations (around 95%) the relationship is positive. The negative part of the U-shaped curve is primarily driven by five outliers. These are the five industry-year observations that show very high level of R&D expenditure at very low levels of competition. Hence, my broad conclusion from this result is still that there is a positive relationship between competition and innovation. Here, it is relevant to point out that ABBGH also test the robustness of their results by using R&D as an alternative measure of innovation. In this case too, they find an inverted-U relationship, though it is not statistically significant.

loss making firms in the industry, this could result in our measure of competition being greater than one, although theoretically it cannot be greater than one. To avoid this problem, I restrict firm-level Lerner’s indices to be non-negative when computing the benchmark measure of competition. The alternative measure of competition relaxes this constraint and allows the Lerner’s index to be negative.
Finally, I approximate the relationship between competition and innovation by using a flexible spline. The fitted spline is shown in Figure 2(b). Here again, the evidence for the positive relationship is quite clear.\footnote{I also tried some other robustness checks. I do not report their results to save on space. These included: 1) use of a cubic, instead of a quadratic, model; 2) use of simple patent count as the measure of competition; and 3) restricting the time coverage of the sample to 1976-1994 period to make it similar to the coverage of ABBGH’s sample (1974-1994). The positive relationship between competition and innovation is robust to all these alternative assumptions.}

To sum up, the positive relationship between competition and innovation that we find in the US data is robust to the alternative empirical assumptions that I have explored in this subsection.

5.2 Prediction 2

The main conclusion about Prediction 2 from the US data is that there is no clear relationship between competition and average technology gap within the industry. In this subsection, I explore the robustness of this result to a set of alternative empirical assumptions. The results of this exercise are in Table \[9\] and Figure \[9\] In all the cases, I control for the fixed industry and year effects. Column (1) in Table \[9\] is the same as column (2) in Table \[7\]. This is the benchmark case.

The benchmark measure of technology gap is computed by averaging the TFP gap between the leader and each firm within the industry. In column (2) of Table \[9\], I report the results based on an alternative definition of technology gap. This time, the measure of technology gap is based on labor productivity (LP) differences between the leader and individual firms. The conclusion is unaffected: there is no clear relationship between competition and technology gap in the US data. Also note that the constant term in these regressions is highly significant. This implies that the technology gap is more or less constant and independent of the level of competition.

In column (3) of Table \[9\], I report the results based on the alternative measure of competition.\footnote{Please see Section 5.1 for a description of the alternative measure of competition.} Once again the conclusion remains the same and we see no clear relationship between competition and the technology gap.

In column (4) of Table \[9\], I report the results based on 2-digit level data. Here, the coefficient on competition is positive and although it is not significant at 10% level, it is close to be significant (the $p$-value is around 0.11). This statistically weak positive relationship is quantitatively weak too: a one standard deviation increase in competition leads to a mere 0.09 standard deviation increase.
in the technology gap. I plot this relationship in Figure 3. In fact, if one drops the three outliers with technology gap less than 0.2, the relationship becomes almost flat.

To sum up, the conclusion about the absence of a clear relationship between competition and technology gap in the US data is robust to the alternative empirical assumptions that I have tried in this subsection.

5.3 Prediction 3

The conclusion about Prediction 3 from the US data is that the positive relationship between competition and innovation is steeper in the case of more neck-and-neck industries. In this subsection I explore the robustness of this conclusion to a set of alternative empirical assumptions. Since it is easier to compare the slopes of non-linear curves visually, I use graphs to present the results of this robustness exercise. The results are in Figure 4. The benchmark case, which I do not reproduce in Figure 4 is in Figure 1. In all four panels of Figure 4 the black line represents the results for all industries and the red (or grey, if viewed in greyscale) line depicts the results for more neck-and-neck industries.

In Figure 4(a), I report the results from a Poisson model. The positive relationship is steeper for more neck-and-neck industries. In Figure 4(b), I report the results based on the alternative definition of competition. Once again, the positive relationship is steeper for more neck-and-neck industries. In Figure 4(c), I report the results based on R&D as the measure of innovation. Here the two curves almost coincide. However, a closer look reveals that the positively sloping part of the red (or grey) curve is slightly steeper than the corresponding part of the black curve, though, admittedly, the difference is very small. In Figure 4(d), I report the results based on 2-digit level data. The two curves, again, look very similar but like the previous case, on closer examination, on can see that the red (grey) line is slightly steeper.

To sum up, the conclusion that the positive relationship between competition and innovation in the US data is steeper for more neck-and-neck industries is robust to the alternative empirical assumptions that I have explored in this subsection.

The general conclusion from the robustness analysis is that the empirical results reported in Section 4 still hold when we change the empirical model to Poisson, change the dataset to 2-digit level, modify the measure of competition or use R&D expenditures as the measure of competition. This conclusion is reassuring and we can have some confidence that our empirical results are robust.
and not driven by any particular empirical assumption. In the next section, I try to reconcile the different empirical results that ABBGH and I obtain from the UK and the US datasets.

6 Reconciling the Evidence

To reconcile the apparently different empirical results from the US and the UK datasets I proceed as follows. First, I argue that a partial equilibrium industry model is a better theoretical counterpart to the data generating process than the general equilibrium model used by ABBGH. Next, I present the partial equilibrium industry model, which builds on the simple duopoly model in ABBGH. Finally, I show that the partial equilibrium model provides a simple and intuitive way to reconcile the evidence from the two datasets.

6.1 Why a Partial Equilibrium Model?

The basic building block of the theoretical model in ABBGH is a simple duopoly model. They then assume a steady-state and derive the general equilibrium of the model. They define the steady state as an invariant distribution of industries by technology gap. There are at least three problems with their general equilibrium framework that make it less suitable for explaining the empirical results that we get from a panel of industry-level data.

First, ABBGH assume that the economy is always in the steady state. It implies that when the level of competition changes exogenously from one period to the next, the economy instantly adjusts to the new invariant distribution. Note that in ABBGH, for each level of competition there is a corresponding invariant distribution of industries by technology gap.

Second, to derive the invariant distribution, one needs to assume that the level of competition is the same across industries. However, the main identifying assumption in the econometric model used by ABBGH and myself is that the level of product market competition varies exogenously over time and across industries. Hence the assumption of the same level of competition across industries imposes an unnecessary restriction on the theoretical model.

Third, the inverted-U relationship that ABBGH derive from their theoretical model is at the aggregate economy level and not at the industry level (see ABBGH, Proposition 2, p. 715. Especially note the words “aggregate innovation rate”). However, what we empirically test is the

\[10\) The general equilibrium model in ABBGH can also deliver a positive relationship between competition and
relationship between competition and innovation at the industry level. In their model, the industry level relationship is either positive or negative (see ABBGH, Proposition 1, p. 714).

I would like to argue that because of these three problems with the general equilibrium model, the partial equilibrium version of ABBGH’s model is a better theoretical counterpart to the data generating process. The reason is that the outcomes related to the variables of interest (in this case: competition, research intensity and technology gap) from the partial equilibrium model map directly into the industry-level empirical variables that we use in estimation.

However, the partial equilibrium model in ABBGH in its current form can either deliver a positive or a negative relationship between competition and innovation. This is mainly because of the assumption that the maximum technology gap between the leader and the laggard can only be one step. When I relax this assumption, the partial equilibrium model generates richer dynamics and, depending on the average degree of neck-and-neckness in the industry, can generate a positive, a negative or an inverted-U relationship.

In what follows, I present a slightly modified version of the duopoly model in ABBGH. I show that the equilibrium of the modified model and the simulations based on it can help us understand most of the empirical findings in this paper. The key theoretical assumption is that the industries in the US are technologically more neck-and-neck than those in the UK.

6.2 The Partial Equilibrium Industry Model

The model is similar to the one in the working paper version of ABBGH (see Aghion et al. 2002) except for two significant differences. First, I focus on the partial equilibrium of an industry and do not use the assumption of invariant distribution to derive the general equilibrium. Second, I do not restrict the maximum technology gap between the leader and the laggard to one step. I now describe the model in some detail.

aggregate economy-wide innovation. The basic condition for this result is that the help parameter (which is fixed in their model) should be large enough. So much so, that even when the level of competition is very high, the laggards are able to innovate solely because of spillovers from the leaders, although their own R&D effort may be zero. This leads to a higher degree of neck-and-neckness in the steady state and generates a positive relationship between competition and aggregate innovation.
6.2.1 Economic Environment

Time is discrete. Consider a duopoly and a mass 1 of infinitely lived consumers with preferences
\[ U = \sum_{t=0}^{\infty} \beta^t \ln Q_t, \]
where
\[ Q_t = \left( q_{it}^\alpha + q_{-it}^\alpha \right)^{1/\alpha}, \quad \alpha \in (0, 1] \]

where subscripts \( i \) and \( -i \) denote the two firms in the industry. Each firm produces one good.

The parameter \( \alpha \) measures the degree of substitutability between the two goods. If \( \alpha = 1 \), the goods are perfect substitutes. In addition to this, \( \alpha \) plays another important role in the model: it is a measure of exogenous competition. If \( \alpha \) is higher, the goods are closer substitutes and this implies higher competition between the firms. The representative consumer allocates her income between the two goods to maximize utility.

Each firm produces according to a constant returns to scale technology using labor as the only input and taking the wage rate as given. The supply of labor is assumed to be perfectly elastic\footnote{This assumption is made for convenience only. We only need that the firms take the wage rate as given. The assumption of perfectly elastic labor makes the wage rate constant over time.}
\[ q_i = A_i L_i, \]
where \( A_i = \gamma^{k_i} \). The parameter \( \gamma \) represents the size of innovation and \( k_i \) is the technology level of firm \( i \). If firm \( i \) is the leader and firm \( -i \) is the laggard then the technology gap in the industry is given by
\[ n = k_i - k_{-i}. \]

In this model, the technology gap is the only state variable. The marginal cost of production for firm \( i \) is given by
\[ c_i = w \gamma^{-k_i}, \]
where \( w \) is the wage rate. Each firm chooses its price to maximize profits, taking the price of its rival as given (Bertrand Equilibrium). The problem of the firm is:
\[ \max_{p_i | p_{-i}} \pi_i = (p_i - c_i) q_i (p_i, p_{-i}). \]

Innovation depends on research intensity \( x \in [0, \infty) \). If \( x \) units of labor are devoted to R&D, the probability of a successful innovation is given by \( \Pr (x \left( n \right), n) \in [0, 1] \). Let \( V \left( n \right) \) denote the...
value of the firm that is \( n \) steps ahead of its rival (\( n \) could be negative). The Bellman equation for the firm is given by:

\[
V(n) = \max_{x(n)\mid x(-n)} \left\{ \pi(n) - wx(n) + \beta \left[ \Pr(x(n), n) [1 - \Pr(x(-n), -n)] V(n + 1) + \right. \right.
\]
\[
\left. \left[1 - \Pr(x(n), n)] \Pr(x(-n), -n) V(n - 1) + \Pr(x(n), n) \Pr(x(-n), -n) V(n) + \right. \right.
\]
\[
\left. [1 - \Pr(x(n), n)] [1 - \Pr(x(-n), -n)] V(n) \right\}.
\]

(5)

The firm chooses its research intensity \( x(n) \) by taking the research intensity of its rival \( x(-n) \) as given. The period profit function is given by \( \pi(n) \). Since there are two firms and two possible outcomes of innovation for each, in total there are four possibilities: (1) If the firm succeeds in innovating and its rival does not, an event that occurs with probability \( \left[ \Pr(x(n), n) [1 - \Pr(x(-n), -n)] \right] \), the technology gap will increase from \( n \) to \( n + 1 \); (2) If the firm does not succeed and its rival does, an event that occurs with probability \( [1 - \Pr(x(n), n)] \Pr(x(-n), -n) \), the technology gap will decrease from \( n \) to \( n - 1 \); (3) If both firms succeed, an event that occurs with probability \( \Pr(x(n), n) \Pr(x(-n), -n) \), the technology gap will remain unchanged at \( n \); and (4) If both firms do not succeed, an event that occurs with probability \( [1 - \Pr(x(n), n)] [1 - \Pr(x(-n), -n)] \), the technology gap will again remain unchanged at \( n \).

I define \( \Pr(x(n), n) \) as

\[
\Pr(x(n), n) = \left[1 - e^{-ax}\right] + \max\left[0, 1 - e^{\eta(n-\hat{n})}\right],
\]

(6)

where \( a > 0 \) and \( \eta > 0 \) are parameters, \( x \) is the research intensity and \( \hat{n} \leq 0 \) is a technology gap threshold explained below. There are two components of the right-hand side of (6). The first component, \( [1 - e^{-ax}] \), is the probability of a successful innovation based on a firm’s own research effort. Parameter \( a \) measures the productivity of the R&D effort. The second component, \( \max\left[0, 1 - e^{\eta(n-\hat{n})}\right] \), is the additional probability of success that a laggard enjoys because of knowledge spillovers from the leader. This is similar to the help parameter \( h \) in ABBGH. The difference is that I make it a function of technology gap: the further behind the laggard is, the easier it is for him to move one step ahead by immitating the leader.\(^{12}\) This assumption is important to avoid divergence between the firms in forward simulations. Without this assumption, the technology gap in the industry continues to increase until it is so high that the innovation incentives for both the

\(^{12}\)Goettler and Gordon [2009] make a similar assumption in their study of competition and innovation in the microchip industry.

21
firms are zero. With this assumption, the technology gap remains low because it is easier for the laggard to catch up. Parameter \( \eta \) is the help parameter. Note that if \( n \geq 0 \), i.e. the firm is either neck-and-neck with its rival or the leader, then the help component is zero. That means the help is only available to the laggards. However, if \( \hat{n} < 0 \), the help is only available to the laggards whose gap from the leader is more than \(-\hat{n}\) steps. To see this more clearly, assume that \( \hat{n} = -3 \). Then the help component is positive only if \( n \leq -4 \). The idea is that the spillovers are only helpful for the laggards who are some distance behind the leader. This assumption helps in keeping the innovation incentives of neck-and-neck firms high. The reason is that if a laggard can benefit from the leader even if the former is just one step behind, it will discourage innovation by the neck-and-neck firms because if they are successful in their innovation, their rival will immediately benefit from the spillovers. We need to keep the innovation incentives of neck-and-neck firms high because it is crucial for reconciling the evidence from the US and the UK datasets. I comment more on this point below.

6.2.2 Equilibrium

Given the parameters of the model \((\alpha, a, \beta, \eta, \gamma, \hat{n}, w, N)\), where \( N < \infty \) is an arbitrarily large positive number that represents the maximum allowable technology gap, the industry equilibrium consists of relative price and quantity sequences \( \left\{ \frac{p_i}{q_i}(n), \frac{q_i}{q_{i-1}}(n) \right\}_{n=-N}^{N} \), a profit sequence \( \{\pi(n)\}_{n=-N}^{N} \), research intensities \( \{x(n)\}_{n=-N}^{N} \) and firm values \( \{V(n)\}_{n=-N}^{N} \) such that:

1. the relative quantity \( \frac{q_i}{q_{i-1}}(n) \) maximizes the representative consumer’s utility, given the relative price \( \frac{p_i}{q_{i-1}}(n) \);

2. the relative price \( \frac{p_i}{q_{i-1}}(n) \) maximizes firms’ profits when the firms take the relative quantity \( \frac{q_i}{q_{i-1}}(n) \) as given and the equilibrium profits are given by \( \{\pi(n)\}_{n=-N}^{N} \).

3. Each firm chooses its research intensity to maximize its value, taking the research intensity of its rival as given. The solution to this problem generates equilibrium strategies \( \{x(n)\}_{n=-N}^{N} \) and the corresponding equilibrium values \( \{V(n)\}_{n=-N}^{N} \).
6.2.3 Solving the Model

Normalizing the total expenditure on the two goods to one, we can solve the representative consumer’s problem to derive the demand facing firm \( i \). The demand is given by:

\[
q_i = \frac{\frac{1}{p_i}}{\frac{1}{p_i} + \frac{1}{p_{-i}}}.
\]

The first-order condition for firm \( i \) is:

\[
(p_i - c_i) \frac{\partial q_i}{\partial p_i} + q_i = 0.
\]

After some algebra, the first-order condition can be written as

\[
\frac{p_i}{p_{-i}} = \gamma^{-n} \phi \left( \frac{p_i}{p_{-i}} \right),
\]

where \( \gamma^{-n} = c_i / c_{-i} \) and \( \phi \left( \frac{p_i}{p_{-i}} \right) \) is

\[
\phi \left( \frac{p_i}{p_{-i}} \right) = \frac{1 - \alpha + \left( \frac{p_i}{p_{-i}} \right) \frac{1}{1 - n}}{1 - \alpha + \left( \frac{p_i}{p_{-i}} \right) \frac{1}{1 - n}}.
\]

Equation (7) implicitly defines equilibrium price ratio \( \frac{p_i}{p_{-i}} (n) \). Given the equilibrium price ratio, the equilibrium profit is

\[
\pi (n) = \frac{(1 - \alpha) R_i (n)}{1 - \alpha R_i (n)},
\]

where \( R_i (n) \) is the share of firm \( i \) in total industry revenue and is defined as

\[
R_i (n) = \frac{1}{1 + \left[ \frac{p_i}{p_{-i}} (n) \right] \frac{1}{1 - n}}.
\]

In order to solve for equilibrium strategies \( \{ x (n) \}_n^{\text{N}} \) and the corresponding equilibrium values \( \{ V (n) \}_n^{\text{N}} \) we need to solve the dynamic program in (5). The specific form of \( \Pr (x (n), n) \) assumed in (6) greatly simplifies the problem because given the value function, one can solve explicitly for the equilibrium strategies. To see this, note that the first order condition for \( x (n) \), after some simplification, is given by:

\[
\frac{\partial \Pr (x (n), n)}{\partial x (n)} = \frac{w}{\beta b (n)},
\]

23
where
\[
\begin{align*}
b(n) &= \left[1 - \Pr(x(-n),-n)\right] [V(n+1) - V(n)] + \Pr(x(-n),-n) [V(n) - V(n-1)].
\end{align*}
\]

Using (6), we get the following closed form solution for \(x(n)\)
\[
x(n) = \ln (a \beta b(n)/w)/a.
\]

This policy function is based on an arbitrary value function. To find the equilibrium value function we start from an arbitrary value function and iterate on it using the Bellman equation. Once we have the equilibrium value function, we can use the last equation to find the corresponding equilibrium policy function.

There is no closed-form solution for the relative price ratio or the equilibrium value and policy functions. Hence we need to solve the model numerically. Given the simple structure of the model, the numerical solution is computationally very light and can be found in just a few seconds.

6.3 Using the Partial Equilibrium Model to Reconcile the Evidence

We need to specify the parameter values before we can find the numerical solution to the model and use it to examine the relationships of interest. There are eight parameters in the model.

1. \(\alpha\): This is the substitutability parameter and also serves as the exogenous measure of competition. It can take values between 0 and 1. I solve the model for various values of \(\alpha\) in this range.

2. \(a\): This is R&D productivity parameter. I set \(a = 5\).

3. \(\beta\): This is the discount factor. I assume the time period to be one year and set \(\beta = 0.95\).

4. \(\eta\): This is the help parameter. I set \(\eta = 0.2\).

5. \(\gamma\): This is the innovation size. I assume \(\gamma = 1.1\). It means that a one-step increase in the technology gap reduces the relative marginal cost of the leader by approximately 10%.

6. \(\hat{n}\): I use two value for this parameter: \(-1\) and \(-2\). The value of \(-1\) implies more neck-and-neck environment and is meant to represent the US economy and \(-2\) implies slightly less neck-and-neck environment and is meant to represent the UK economy. This is the only parameter that I allow to differ between the two countries.
7. $w$: This is the wage rate. I set $w = 0.1$.

8. $N$: This is the maximum allowable technology gap between the leader and the laggard. I set $N = 100$. The maximum gap is large enough to ensure that in equilibrium both the leader and the laggard have almost no incentive to innovate if the gap between them is close to the maximum.

The values for $a$, $\eta$, $\gamma$ and $w$ are arbitrary. However, the qualitative results are not sensitive to changes in these parameters. Here my purpose is not to take a stand on any particular parameter value. Instead, it is to show that the model can reconcile most of the empirical results for a certain set of parameter values.

I solve the model numerically using the above parameter values. In Section 6.3.1 I examine the equilibrium solution of the model to see how much of the empirical evidence that we have seen so far it can explain. I find that the equilibrium of the model can reconcile the empirical evidence on Predictions 1 and 3. However, we need to forward simulate the model to see its implications for Prediction 2. In Section 6.3.2 I present the simulation results. I use them to reconcile evidence from the UK and the US datasets on all three predictions. The key result is that when I restrict the technology gap to very low levels in the simulations for the US industries, the model can explain almost all the empirical results in this paper.

### 6.3.1 Results from Equilibrium of the Model

I plot equilibrium innovation intensity against competition for six different values of technology gap in Figure 5. It is immediately clear that for very small technology gap the equilibrium relationship is positive (see panels (a) and (b)). As the technology gap increases, the relationship resembles an inverted-U (see panels (c) and (d)). And for very large technology gap the relationship is negative (see panel (e)). Another important feature of the equilibrium is that the innovation intensity is generally higher in low gap industries. These findings can reconcile some of the conflicting empirical evidence from the UK and the US datasets. In the empirical section, we saw that the relationship between competition and innovation was inverted-U in the UK data and positive in the US data. The equilibrium of the model suggests that if the manufacturing industries in the US are more neck and neck than those in the UK, then the model can explain why we see an inverted-U relationship in the UK and a positive relationship in the US. The equilibrium also helps us reconcile the empirical results for Prediction 3. We see in Figure 5 that more neck-and-neck industries tend
to innovate more. The mechanism behind both these results is the same as explained by ABBGH: as
technology gap narrows, the escape competition effect becomes stronger and eventually dominates
the Schumpeterian effect.

Figure 5 can also be used to understand the theoretical findings in ABBGH. They show that
in neck-and-neck industries the relationship is positive. We see that is panels (a) and (b). They
also show that in unlevelled industries the relationship is negative. We see that in Panel (e). Since
ABBGH restricted the maximum technology gap to one, they do not get the inverted-U relationship
shown in panels (c) and (d). This also explains why they use the assumption of the invariant
distribution. They do so because it enables them to show a nice inverted-U relationship at the
aggregate level. However, by relaxing their assumption of the maximum technology gap of one
step, I am able to get the inverted-U relationship at the industry level from the partial equilibrium
model. Hence I do not need to impose the restrictive assumption of the invariant distribution of
industries by technology gap.

The results in Figure 5 do not tell us any thing about the relationship between competition
and technology gap. This is because we fix the level of competition and then solve the model for
equilibrium innovation intensity at various levels of technology gap (see the definition of equilibrium
in Section 6.2.2). To study the relationship between competition and technology gap using the
partial equilibrium model, we need to simulate the model.

6.3.2 Results from Simulations of the Model

For simulation purposes, I assume that competition follows an exogenous process. In fact, in both
the UK and the US datasets, AR(1) processes approximate the evolution of competition very well.
I estimate the AR(1) processes from the two datasets and use them in simulations below.

For each simulation, in the initial period (period 0) I start from an arbitray combination of
competition ($\alpha_0$) and technology gap ($n_0$). Given the ($\alpha_0, n_0$) combination, I use the equilibrium
solution to find the equilibrium research intensity of both firms [$x(n_0), x(-n_0)$]. These research
intensities together with (6) give the equilibrium probabilities of successful innovation. Based
on these probabilities, I draw from a [0,1] uniform distribution to determine the outcomes of
research activity. Based on the outcomes, I update the technology gap to $n_1$. I then use the
estimated AR(1) process to get $\alpha_1$. I use this updated combination ($\alpha_1, n_1$) and repeat the above
calculations for period 1. I repeat this process for 10,000 periods. For each simulation period \( \tau \in \{0, 1, 2, \cdots, 10000\} \), I record the level of competition \( (\alpha_{\tau}) \), technology gap \( (n_{\tau}) \) and the sum of the equilibrium research intensities \( x(n_{\tau}) + x(-n_{\tau}) \). These simulations are repeated 100 times for each country. The only important difference between the simulations for the two countries is that when I simulate the UK economy, I use the solution of the model with \( \hat{n} = -2 \) and when I simulate the US economy I use the solution based on \( \hat{n} = -1 \). One hundred simulations of 10,000 periods each give one million observations on competition, technology gap and innovation for each country. I use these simulated data to run linear or quadratic regressions and report the results in Figure 6. Figure 6 is the theoretical counterpart of Figure 1. I shall compare the two figures panel-by-panel.

Figure 6(a) shows a nice inverted-U relationship. This is meant to theoretically explain the inverted-U relationship that we observe in the UK data (see Figure 1(a)).

Figure 6(b) is meant to replicate the positive relationship that we observe in the US data (see Figure 1(b)). Here although the relationship is mostly positive, at very high levels of competition, it turns negative. This is because in the simulated industries the average technology gap is 3.21 and hence the firms are not always in a neck-and-neck situation. Still the relationship in Figure 6(b) is more positive and less inverted-U compared to Figure 6(a). Hence the simulated model does explain that if technology gap is small, the relationship between competition and innovation tends to be more positive.

In Figure 6(c), we see a positive relationship between competition and technology gap. This replicates the observed relationship for the UK data very well (see Figure 1(c)).

In Figure 6(d), again we see a positive relationship between competition and technology gap. This contrasts with what we see in the US data (see Figure 1(d)). In the US data, there is no relationship between competition and technology gap. The theoretical model is unable to match this observation. However, the theoretical relationship is quantitatively weak: a one standard deviation (SD) increase in competition leads to a mere 0.16 SD increase in the technology gap. It could be that there are some other factors at work in the US economy that we do not capture in

---

13 The number of periods was chosen to be 10,000 because it is large enough to ensure that the arbitrarily chosen initial values of \( \alpha \) and \( n \) have no effect on simulation results.

14 For each country, I use a separate \( AR(1) \) process for \( \alpha \). The \( AR(1) \) processes are estimated from the two datasets. However, this difference is not important. Even if I use the same \( AR(1) \) process for both countries, the results are very similar.
our model. One possibility is that when competition increases, technologically weaker firms leave the industry and are replaced by relatively stronger firms. This will reduce the average technology gap in the industry and hence flatten the relationship. However, since our model does not feature entry or exit, we do not capture this characteristic of the US economy.

Figure (6e) is meant to replicate the empirical results from the UK data on Prediction 3 (see Figure (1e)). The red (grey, if viewed in greyscale) line represents the industries with technology gap less than the median and the black line represents the remaining industries. The simulation results replicate the empirical results from the UK data: the inverted-U is steeper for more neck-and-neck industries.

Figure (6f) is meant to replicate the empirical results from the US data on Prediction 3 (see Figure (1f)). Once again, the red (grey, if viewed in greyscale) line represents the industries with technology gap less than the median and the black line represents the remaining industries. The simulation results show that at low levels of competition the neck-and-neck industries innovate slightly less. This is inconsistent with the empirical results from the US data. However, at high levels of competition the simulation results replicate the empirical results from the US data.

To wrap up, the first simulation results shown in Figure 6 replicate the empirical results from the UK data very well. This is evident when one compares the simulation results in panels (a), (c) and (e) of Figure 6 with the empirical results in the corresponding panels in Figure 1. However, the simulation results only partially replicate the empirical results from the US data. When we compare the simulation results in panels (b), (d) and (f) of Figure 6 with the empirical results in the corresponding panels in Figure 1 we see that simulation results partially replicate the results in panels (b) and (f) but do not replicate the results in panel (d).

The key insight from the equilibrium of the model and the simulation results is that in more neck-and-neck industries the relationship between competition and innovation is positive. However, in the simulation results that are meant to replicate the US results, the endogenously evolving technology gap has a mean value of 3.21. Intuitively, if we could restrict the technology gap to a lower level, our simulated results would better replicate the US data. I do so in Figure 7. The only difference between Figures 6 and 7 is that in panels (b), (d) and (f) of the latter, I restrict the technology gap to a maximum of two steps. Now the simulation results replicate the empirical results from the US data very well. To see this compare panels (b), (d) and (f) of Figure 7 with the corresponding panels of Figure 1.
To sum up, the simulation results from the partial equilibrium industry model can replicate empirical results from both the UK and the US datasets if we assume that the manufacturing industries in the US are, on the average, more neck-and-neck than their counterparts in the UK.

The assumption that the US industries are more neck and than the UK industries finds some support in the data. For example, the median technology gap in the US is 0.48 compared to 0.51 in the UK (see Table 1). The difference is even higher between the 10th percentile industries: 0.23 in the US compared to 0.28 in the UK. So the industries with less than median technology gap in the US are more neck and neck than the similar industries in the UK. If we use labor productivity instead of TFP to compute the technology gap then the US industries appear to be even more neck and neck: the mean gap is 0.47, the median is 0.45 and the 10th percentile is 0.20.

7 Concluding Remarks

I have shown that the relationship between competition and innovation is positive in the US data. This is in contrast with the empirical results in ABBGH who find an inverted-U relationship in the UK data. Building on the duopoly model in ABBGH, I show that a partial equilibrium industry model can explain these conflicting results if one assumes that the US manufacturing industries are technologically more neck and neck than their UK counterparts. There is some support for this assumption in the data.

If one agrees with these findings then a natural question is: why the US industries are more neck-and-neck than the UK industries? Although I do not have a definitive answer to this question, I would offer some possible candidates. One possibility is the legal institutions. If bankruptcy laws are more lenient in the US then the weaker firms are more likely to quit and the remaining firms will be more neck-and-neck. Another possibility is the more efficient allocation of resources between firms. A recent body of literature (see Hsieh and Klenow [2009], for example) uses dispersion in TFP as a measure of allocative efficiency. If TFP dispersion is lower in the US, it could be because the allocative efficiency is higher.

The last point leads to another important policy implication of the results in this paper. If a higher degree of neck-and-neckness leads to a more positive relationship between competition and innovation then any policy to reduce the allocative inefficiency will also promote innovation and hence growth. So in the cases of India and China that Hsieh and Klenow [2009] examine, if governments take steps to promote allocative efficiency, a byproduct of these policies will be a
market driven increase in research activity. This may already be happening. The World Intellectual Property Organization reports in a recent press release that the international patent filings (under WIPO’s Patent Cooperation Treaty) by China had increased by a whopping 313% in just four years from 2006 to 2010. Growth in Indian filings was 33% over the same period. Although this is just one-tenth of the growth in the Chinese filings, it is still four times the global growth in filings, which was 8.9%, over the same period[15]

References


A Data

In this appendix, I briefly describe the construction of the key empirical variables used in estimation.

1. Competition

Following ABBGH (see equation (1) on page 705 of ABBGH), the competition variable is defined as one minus the average Lerner’s index of individual firms in the industry. Specifically

\[ c_{jt} = 1 - \frac{\sum_{i=1}^{n_{jt}} L_{it}}{n_{jt}}, \]

where \( j \) denotes the industry and \( i \) denotes a firm in the industry. The Lerner’s index is computed in the same way as in ABBGH (see footnote 7, p 704 in ABBGH):

\[ L_{it} = \frac{\text{operating profit} - \text{financial cost}}{\text{sales}}. \]

The firm-level data are from Standard and Poor’s COMPUSTAT database. To compute real quantities I use industry-level price indexes from NBER’s productivity database.

2. Citation-weighted patents

Citation-weighted patents are computed using patent and citation data from the New NBER Patent Data Project. These data are available at https://sites.google.com/site/patentdataproject (last accessed on November 16, 2010). To match the patent data to the firm-level accounting data from COMPUSTAT I use the unique gvkey assigned to each firm in both datasets.

3. Technology gap

The industry-level technology gap is the average of firm-level technology gap \( (m) \), which is computed as:

\[ m_{it} = \frac{TFP_{Lt} - TFP_{it}}{TFP_{Lt}}. \]

The TFP (total factor productivity) is computed as \( (y/L)/(k/L)^{\alpha} \), where \( y \) represents real sales multiplied by value-added to sales ratio for the industry, \( k \) capital stock and \( L \) number of employees. I use \( \alpha \) equal to 1/3. This measure of TFP is highly correlated with NBER’s five-factor productivity. The correlation between the two (at the four-digit level) is 0.88. This is despite the fact that the sample used to compute NBER productivity measures is not the same as the COMPUSTAT sample used in this study. If one computes TFP by applying the above definition to NBER’s dataset, the resulting TFP series has a correlation coefficient of 0.94 with NBER’s five-factor productivity.

4. Instruments

I use two instruments for competition. The first instrument is import-weighted average tariff rate, where the tariff rate is defined as: duties/(customs value). The second instrument is import-weighted freight rate, where the freight rate is defined as: cif/fob - 1. The acronym cif stands for cost, insurance and freight and the acronym fob stands for free on board. The data for these instruments are from Peter Schott’s online database. These data are publicly available at http://www.som.yale.edu/faculty/pks4/sub_international.htm (last accessed on November 16, 2010).
<table>
<thead>
<tr>
<th></th>
<th>US Data (three-digit)</th>
<th>ABBGH Data (two-digit)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Competition</strong> = 1 – LI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.76</td>
<td>0.95</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.11</td>
<td>0.02</td>
</tr>
<tr>
<td>( p_{10} )</td>
<td>0.63</td>
<td>0.92</td>
</tr>
<tr>
<td>Median</td>
<td>0.76</td>
<td>0.95</td>
</tr>
<tr>
<td>( p_{90} )</td>
<td>0.89</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Innovation</strong> = Citation-weighted Patents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>5.53</td>
<td>6.66</td>
</tr>
<tr>
<td>S.D.</td>
<td>9.98</td>
<td>8.43</td>
</tr>
<tr>
<td>( p_{10} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Median</td>
<td>1.59</td>
<td>3.35</td>
</tr>
<tr>
<td>( p_{90} )</td>
<td>17.02</td>
<td>20.19</td>
</tr>
<tr>
<td><strong>Technology Gap</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.20</td>
<td>0.16</td>
</tr>
<tr>
<td>( p_{10} )</td>
<td>0.23</td>
<td>0.28</td>
</tr>
<tr>
<td>Median</td>
<td>0.48</td>
<td>0.51</td>
</tr>
<tr>
<td>( p_{90} )</td>
<td>0.77</td>
<td>0.69</td>
</tr>
</tbody>
</table>

**Miscellaneous**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Industries</td>
<td>116</td>
</tr>
<tr>
<td>Time Period</td>
<td>1976-2001</td>
</tr>
<tr>
<td>Observations</td>
<td>2756</td>
</tr>
<tr>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>1974-94</td>
</tr>
<tr>
<td></td>
<td>354</td>
</tr>
</tbody>
</table>
substantially smaller sample, the coefficients are not statistically significant. Finally, we fit the relationship for each of the top four innovating industries in our sample, and in each case there is part or all of an inverted-U shape (see the earlier working paper version of this paper [Aghion et al. 2004]).

### II.D. Using Policy Instruments

The major obstacle to empirical research in this area is that competition and innovation are mutually endogenous. Without addressing this, any results we find are likely to be biased toward

| TABLE I | EXPONENTIAL QUADRATIC: BASIC SPECIFICATION |
|---|---|---|---|---|
| Dependent variable: citation-weighted patents | (1) | (2) | (3) | (4) |
| Data frequency | Annual | Annual | 5-year averages | Annual |
| Competition$_{jt}$ | 152.80 | 387.46 | 819.44 | 385.13 |
| | (55.74) | (67.74) | (265.63) | (67.56) |
| Competition squared$_{jt}$ | $-80.99$ | $-204.55$ | $-434.43$ | $-204.83$ |
| | (29.61) | (36.17) | (141.43) | (36.06) |
| Significance of: Competition$_{jt}$, | 7.60 | 38.34 | 9.97 | 32.59 |
| Competition squared$_{jt}$ | (0.02) | (0.00) | (0.01) | (0.00) |
| Significance of policy instruments | 10.11 | 5.00 | 4.38 | 0.801 |
| in reduced form | (0.002) | (0.000) | (4.04) | |
| Significance of other instruments | 5.00 | 4.38 | 0.801 | |
| in reduced form | (0.000) | (4.04) | | |
| Control functions in regression | 5.00 | 4.38 | 0.801 | |
| Year effects | Yes | Yes | Yes | Yes |
| Industry effects | Yes | Yes | Yes | Yes |
| Observations | 354 | 354 | 67 | 354 |

Competition$_{jt}$ is measured by (1-Lerner index) in the industry-year. All columns are estimated using an unbalanced panel of seventeen industries over the period 1973 to 1994. Estimates are from a Poisson regression. Numbers in brackets are standard errors. The standard errors in column (4) have not been corrected for the inclusion of the control function. Significance tests show likelihood ratio test-statistics and $P$-value from the $F$-test of joint significance. The fourth column includes a control function. The excluded variables are policy instruments specified in Table II, imports over value-added in the same industry-year, TFP in the same industry-year, output minus variable costs over output in the same industry-year and estimates of markups from industry-country regression [Martins et al. 1996] interacted with time trend, all for the United States and France.
Table 3: Replication of Table I of ABBGH using ABBGH Data and Negative-Binomial Model

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citation-weighted Patents</td>
<td>Annual</td>
<td>Annual</td>
<td>5-yr avg</td>
<td>Annual</td>
</tr>
<tr>
<td>Data frequency</td>
<td>174.41</td>
<td>405.45</td>
<td>623.17</td>
<td>409.21</td>
</tr>
<tr>
<td></td>
<td>(200.85)</td>
<td>(92.37)</td>
<td>(257.94)</td>
<td>(92.11)</td>
</tr>
<tr>
<td>$c_{jt}$ (competition)</td>
<td>$-91.86$</td>
<td>$-214.34$</td>
<td>$-331.48$</td>
<td>$-218.04$</td>
</tr>
<tr>
<td></td>
<td>(106.58)</td>
<td>(49.39)</td>
<td>(137.46)</td>
<td>(49.33)</td>
</tr>
<tr>
<td>Joint significance ($c_{jt}, c_{jt}^2$)</td>
<td>0.87</td>
<td>22.58</td>
<td>5.88</td>
<td>19.96</td>
</tr>
<tr>
<td></td>
<td>(0.65)</td>
<td>(0.00)</td>
<td>(0.05)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Significance of political instruments in reduced form</td>
<td>3.70</td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significance of other instruments in reduced form</td>
<td>4.53</td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. of over-dispersion parameter</td>
<td>2128.49</td>
<td>43.01</td>
<td>$1.3e-5$</td>
<td>42.72</td>
</tr>
<tr>
<td>Control function in regression</td>
<td>5.53</td>
<td>(4.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$ of the reduced form</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year effects</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Industry effects</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>354</td>
<td>354</td>
<td>67</td>
<td>354</td>
</tr>
</tbody>
</table>
Table 4: Replication of Table I of ABBGH using US Data and Negative-Binomial Model

<table>
<thead>
<tr>
<th>Dependent variable: Citation-weighted patents</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data frequency</td>
<td>Annual</td>
<td>Annual</td>
<td>5-yr avg</td>
<td>Annual</td>
</tr>
<tr>
<td>$c_{jt}$ (competition)</td>
<td>$-13.46$</td>
<td>$-7.81$</td>
<td>$-7.85$</td>
<td>$3.95$</td>
</tr>
<tr>
<td></td>
<td>$(2.17)$</td>
<td>$(1.22)$</td>
<td>$(2.49)$</td>
<td>$(2.63)$</td>
</tr>
<tr>
<td>$c_{jt}^2$ (competition squared)</td>
<td>$8.70$</td>
<td>$4.64$</td>
<td>$4.64$</td>
<td>$4.02$</td>
</tr>
<tr>
<td></td>
<td>$(1.48)$</td>
<td>$(0.86)$</td>
<td>$(1.80)$</td>
<td>$(0.85)$</td>
</tr>
<tr>
<td>Joint significance ($c_{jt}$, $c_{jt}^2$)</td>
<td>$42.39$</td>
<td>$78.94$</td>
<td>$21.59$</td>
<td>$48.17$</td>
</tr>
<tr>
<td></td>
<td>$(0.00)$</td>
<td>$(0.00)$</td>
<td>$(0.00)$</td>
<td>$(0.00)$</td>
</tr>
<tr>
<td>Significance of instruments in reduced form</td>
<td>$30.06$</td>
<td>$30.06$</td>
<td>$30.06$</td>
<td>$30.06$</td>
</tr>
<tr>
<td>Sig. of over-dispersion parameter</td>
<td>$2.1e + 4$</td>
<td>$877.31$</td>
<td>$21.59$</td>
<td>$808.07$</td>
</tr>
<tr>
<td></td>
<td>$(0.00)$</td>
<td>$(0.00)$</td>
<td>$(0.00)$</td>
<td>$(0.00)$</td>
</tr>
<tr>
<td>Control function in regression</td>
<td>$-10.96$</td>
<td>$-10.96$</td>
<td>$-10.96$</td>
<td>$-10.96$</td>
</tr>
<tr>
<td></td>
<td>$(2.19)$</td>
<td>$(2.19)$</td>
<td>$(2.19)$</td>
<td>$(2.19)$</td>
</tr>
<tr>
<td>$R^2$ of the reduced form</td>
<td>$0.68$</td>
<td>$0.68$</td>
<td>$0.68$</td>
<td>$0.68$</td>
</tr>
<tr>
<td>Year effects</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Industry effects</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>$2756$</td>
<td>$2756$</td>
<td>$548$</td>
<td>$2756$</td>
</tr>
</tbody>
</table>
Two features stand out clearly. First, more neck-and-neck industries show a higher level of innovation activity for any level of product market competition. Second, the inverted-U curve is steeper for the more neck-and-neck industries, which accords well with our theoretical predictions. These differences are statistically significant, as shown in columns 3 and 4 of Table II, which reports the quadratic coefficients for the whole sample and the high neck-and-neck industry subsample including a full set of year dummies (third column) and a full set of year and industry dummies (fourth column). The interaction terms are jointly significant in both cases.

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology gap</td>
<td>2.858</td>
<td>0.942</td>
<td>183.81</td>
<td>424.46</td>
</tr>
<tr>
<td>Technology gap</td>
<td>(0.400)</td>
<td>(0.419)</td>
<td>(58.99)</td>
<td>(69.5)</td>
</tr>
<tr>
<td>Competition squared</td>
<td>-96.35</td>
<td>-222.9</td>
<td>(31.01)</td>
<td>(36.9)</td>
</tr>
<tr>
<td>Technology gap</td>
<td>1.43</td>
<td>3.82</td>
<td>(2.48)</td>
<td>(2.66)</td>
</tr>
<tr>
<td>Competition squared</td>
<td>-1.30</td>
<td>-3.84</td>
<td>(2.59)</td>
<td>(2.78)</td>
</tr>
</tbody>
</table>

Significance of:
- Competition
- Competition squared
- Competition squared

Year effects: Yes
Industry effects: Yes

Competition is measured by (1-Lerner index) in the industry-year. Technology gap is measured by the average distance to the TFP frontier firm across all firms in the industry-year, so it is an inverse measure of neck-and-neckness. All columns estimated using an unbalanced panel of 354 yearly observations on seventeen industries over the period 1973 to 1994. Significance tests show likelihood ratio test-statistics and P-value from the F-test of joint significance. Numbers in brackets are standard errors. The standard errors in columns 3 and 4 have not been corrected for the inclusion of the control function.
Table 6: Replication of Table III of ABBGH using ABBGH Data and Negative-Binomial Model

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>(1) Technology gap</th>
<th>(2) Technology gap</th>
<th>(3) Citation-weighted patents</th>
<th>(4) Citation-weighted patents</th>
<th>(5) Citation-weighted patents</th>
<th>(6) Citation-weighted patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation Procedure:</td>
<td>Linear regression</td>
<td>Linear regression</td>
<td>Negative Binomial</td>
<td>Negative Binomial</td>
<td>Negative Binomial</td>
<td>Negative Binomial</td>
</tr>
<tr>
<td>$c_{jt}$ (competition)</td>
<td>2.86 (0.40)</td>
<td>0.94 (0.41)</td>
<td>204.43 (215.13)</td>
<td>441.03 (94.76)</td>
<td>184.38 (217.53)</td>
<td>454.00 (94.73)</td>
</tr>
<tr>
<td>$c_{jt}^2$ (competition squared)</td>
<td>−106.67 (113.04)</td>
<td>−231.52 (50.26)</td>
<td>−96.13 (114.31)</td>
<td>−240.17 (50.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c_{jt} \times d_m$ (competition × low gap dummy)</td>
<td>1.60 (8.50)</td>
<td>4.28 (3.78)</td>
<td>3.17 (8.84)</td>
<td>5.45 (3.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c_{jt}^2 \times d_m$ (competition squared × low gap dummy)</td>
<td>−1.46 (8.90)</td>
<td>−4.34 (3.95)</td>
<td>−3.13 (9.27)</td>
<td>−5.55 (4.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joint significance ($c_{jt}, c_{jt}^2$)</td>
<td>0.94 (0.63)</td>
<td>22.73 (0.00)</td>
<td>0.76 (0.68)</td>
<td>23.12 (0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joint significance ($c_{jt} \times d_m, c_{jt}^2 \times d_m$)</td>
<td>1.43 (0.49)</td>
<td>3.16 (0.21)</td>
<td>1.26 (0.53)</td>
<td>4.13 (0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. of over-dispersion parameter</td>
<td>2113.14 (0.00)</td>
<td>38.06 (0.00)</td>
<td>2111.96 (0.00)</td>
<td>37.51 (0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control function in regression</td>
<td>6.38 (10.13)</td>
<td>7.35 (4.94)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year effects</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Industry effects</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>354</td>
<td>354</td>
<td>354</td>
<td>354</td>
<td>354</td>
<td>354</td>
</tr>
</tbody>
</table>
Table 7: Replication of Table III of ABBGH using US Data and Negative-Binomial Model

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>(1) Technology gap</th>
<th>(2) Technology gap</th>
<th>(3) Citation-weighted patents</th>
<th>(4) Citation-weighted patents</th>
<th>(5) Citation-weighted patents</th>
<th>(6) Citation-weighted patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation Procedure:</td>
<td>Linear regression</td>
<td>Linear regression</td>
<td>Negative Binomial</td>
<td>Negative Binomial</td>
<td>Negative Binomial</td>
<td>Negative Binomial</td>
</tr>
<tr>
<td>( c_{jt} ) (competition)</td>
<td>(-0.30)</td>
<td>(-0.01)</td>
<td>(-8.65)</td>
<td>(-4.98)</td>
<td>(-6.80)</td>
<td>(3.66)</td>
</tr>
<tr>
<td></td>
<td>((0.04))</td>
<td>((0.04))</td>
<td>((2.09))</td>
<td>((1.25))</td>
<td>((2.18))</td>
<td>((2.50))</td>
</tr>
<tr>
<td>( c_{jt}^2 ) (competition squared)</td>
<td>(5.18)</td>
<td>(2.17)</td>
<td>(4.47)</td>
<td>(1.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((1.45))</td>
<td>((0.91))</td>
<td>((1.49))</td>
<td>((0.90))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c_{jt} \times d_m ) (competition × low gap dummy)</td>
<td>(-4.66)</td>
<td>(-2.18)</td>
<td>(-4.81)</td>
<td>(-1.99)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((0.65))</td>
<td>((0.39))</td>
<td>((0.65))</td>
<td>((0.39))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c_{jt}^2 \times d_m ) (competition squared × low gap dummy)</td>
<td>(4.93)</td>
<td>(3.02)</td>
<td>(5.05)</td>
<td>(2.76)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((0.82))</td>
<td>((0.48))</td>
<td>((0.81))</td>
<td>((0.48))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joint significance (( c_{jt}, c_{jt}^2 ))</td>
<td>(28.76)</td>
<td>(112.72)</td>
<td>(9.77)</td>
<td>(13.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((0.00))</td>
<td>((0.00))</td>
<td>((0.01))</td>
<td>((0.00))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joint significance (( c_{jt} \times d_m, c_{jt}^2 \times d_m ))</td>
<td>(129.49)</td>
<td>(60.68)</td>
<td>(141.33)</td>
<td>(49.40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((0.00))</td>
<td>((0.00))</td>
<td>((0.00))</td>
<td>((0.00))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. of over-dispersion parameter</td>
<td>(2.0e + 4)</td>
<td>(708.92)</td>
<td>(2.0e + 4)</td>
<td>(673.98)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((0.00))</td>
<td>((0.00))</td>
<td>((0.00))</td>
<td>((0.00))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control function in regression</td>
<td>(-1.99)</td>
<td>(-8.28)</td>
<td>(-1.99)</td>
<td>(-8.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((0.53))</td>
<td>((2.18))</td>
<td>((0.53))</td>
<td>((2.18))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year effects</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Industry effects</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2756</td>
<td>2756</td>
<td>2756</td>
<td>2756</td>
<td>2756</td>
<td>2756</td>
</tr>
</tbody>
</table>
Table 8: Robustness Analysis: Prediction 1

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>CWP Benchmark</th>
<th>CWP Benchmark</th>
<th>CWP Benchmark</th>
<th>CWP Alternative</th>
<th>R&amp;D Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition of Competition:</td>
<td>Benchmark</td>
<td>Benchmark</td>
<td>Poisson</td>
<td>NB</td>
<td>2SLS</td>
</tr>
<tr>
<td>Regression Model:</td>
<td>NB</td>
<td>NB</td>
<td>Poisson</td>
<td>NB</td>
<td>2SLS</td>
</tr>
<tr>
<td>Data Type:</td>
<td>3-digit</td>
<td>2-digit</td>
<td>3-digit</td>
<td>3-digit</td>
<td>3-digit</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c_{jt}) (competition)</td>
<td>3.95</td>
<td>4.99</td>
<td>10.31</td>
<td>5.38</td>
</tr>
<tr>
<td>(c_{jt}) (competition squared)</td>
<td>4.02</td>
<td>15.01</td>
<td>2.24</td>
<td>2.52</td>
</tr>
</tbody>
</table>

| Joint significance \((c_{jt}, c_{jt}^2)\) | 48.17 | 98.36 | 118.38 | 46.16 | 17.46 |
| Significance of instruments in reduced form | 30.06 | 4.42 | 30.06 | 23.11 | 30.06 |
| Sig. of over-dispersion parameter | 808.07 | 0.00 | - | 799.93 | - |
| Control function in regression | -10.96 | -31.68 | -15.04 | -10.40 | - |
| \(R^2\) of the reduced form | 0.68 | 0.77 | 0.68 | 0.62 | 0.68 |
| Observations | 2756 | 519 | 2756 | 2756 | 2756 |
Table 9: Robustness Analysis: Prediction 2

<table>
<thead>
<tr>
<th>Definition of Competition</th>
<th>Data Type:</th>
<th>Dependent variable:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3-digit</td>
<td>Tech Gap (TFP)</td>
<td>Tech Gap (LP)</td>
<td>Tech Gap (TFP)</td>
</tr>
<tr>
<td>Benchmark</td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Benchmark</td>
<td>3-digit</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.002</td>
</tr>
<tr>
<td>Alternative</td>
<td>3-digit</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Benchmark</td>
<td>2-digit</td>
<td>0.44</td>
<td>0.44</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Observations</td>
<td>2756</td>
<td>2756</td>
<td>2756</td>
<td>520</td>
</tr>
</tbody>
</table>

(cj_t (competition))
Figure 1: Empirical Results for the Three Predictions: ABBGH data vs. US data
Figure 2: Robustness Analysis: Prediction 1

(a) R&D as the measure of innovation

(b) Spline estimate of CWP
Figure 3: Robustness Analysis: Prediction 2
Figure 4: Robustness Analysis: Prediction 3

(a) The Poisson Model
(b) Alternative Definition of Competition
(c) R&D as the measure of innovation
(d) Two-digit Data
Figure 5: Equilibrium Relationship between Competition and Innovation in the Partial Equilibrium Industry Model at Different Levels of Technology Gap
Figure 6: Simulation Results from the Partial Equilibrium Industry Model: The three panels on the left (a, c and e) are meant to replicate the empirical results on the three predictions for the UK industries. The panels on the right (b, d and f) do the same for the US industries.
Figure 7: Simulation Results from the Partial Equilibrium Industry Model when the Technology Gap for the Simulated US Industries is Restricted to Two Steps