Brewed in North America:
Mergers, Efficiency, and Market Power

Paul Grieco*

Joris Pinkse*

and

Margaret Slade†

April 28, 2016

Abstract:

We show how production and pricing data can be used to jointly estimate firm level returns to scale, technical change, TFP growth, and market power (in the form of price-cost markups). Our approach can be used to forecast merger related changes in efficiency and markups ex ante. We estimate our model using data for the North American (US and Canadian) brewing industry, and we use the estimated model to forecast changes that are associated with the merger between Molson and Coors. The model forecasts nontrivial increases in returns to scale and markups that tend to offset one another. We also verify these results by analyzing the impact of the merger retrospectively using post merger data.

Keywords: Mergers, Efficiencies, Markups, Returns to scale, Technical change, Productivity

JEL Classifications: D22, D24, L49, L66

*Department of Economics, The Pennsylvania State University, and the Center for Auctions, Procurements, and Competition Policy

†Corresponding author, m.slade@mail.ubc.ca, Vancouver School of Economics, The University of British Columbia

1 Introduction

Competition authorities and defendants in North American merger cases often use quantitative tools to forecast merger–related changes in market power and efficiency. Their analysis, however, usually proceeds in two stages: first changes in prices are predicted holding costs constant, and second, if competitive concerns are raised, offsetting efficiencies are evaluated. This is a curious process, since firms do not ignore costs when setting prices. Indeed, the two are jointly determined.[1]

We propose a quantitative technique that can be used to estimate merger–related changes in markups, returns to scale, technical change, and total factor productivity (TFP) growth simultaneously and can be implemented with only pre merger data. Our approach considers a merger as a consolidation of firms' physical assets and uses estimates of a flexible production function to predict the impact of a merger on firms costs and market power.

We use our model of production to evaluate the performance of the North American brewing industry. In particular, we compare productivity and markups across US and Canadian firms as well as across craft and mass production. In addition, we take a more detailed look at the US mass market where we assess the implications of market structure for firm and industry performance. Specifically, we look at how industry concentration and firm market share influence markups, scale economies, technical change, and TFP growth. Finally, we use our structural model to perform ex ante and ex post evaluations of the 2005 merger between Molson and Coors. We focus on brewing in North America because it is an industry that has witnessed positive growth in concentration, firm size, and import penetration and we wish to see how those changes in the structure of the market have affected industry performance as well as how a specific change affected the merging firms.

The merger between Molson and Coors united the second largest brewer in Canada with the third largest in the US and created the fifth largest brewer in the world. Over time, the Canadian and US markets had become more integrated due to trade liberalization in general and to the North American Free Trade Agreement in particular. For example, at the time of the merger, Coors Light was the largest selling light beer in Canada. The Molson Coors cross border merger rationalized marketing and distribution, since the brands of both firms were already sold in both countries. However, it is also possible that the merger led to increased market power. Our goal is to quantify these two countervailing effects.

Our model of technology is a standard production function in the spirit of Olley and Pakes (1996) and Ackerberg, Caves, and Frazer (2006). However, since we are interested in estimating markups and returns to scale, we specify a more flexible functional form than the commonly used Cobb–Douglas. We can then compare market power and efficiency over time and across individual or groups of firms. Furthermore, with our model returns to scale vary across firms, which implies that technical change and TFP growth are not the same.[2] We therefore show how each can be estimated and use our flexible model of production to evaluate several aspects of performance pre and post merger.

[1] For example, firms in a competitive industry set price equal to marginal cost and do not change one side of the equation while holding the other constant.

[2] When a Cobb Douglas is estimated it is standard to equate the two.

A typical merger simulation, for example, one that forecasts merger–related price changes, requires an assumption about the nature of competition in the market. In other words, it is necessary to specify the game that the firms are engaged in.[3] If the assumption about the nature of competition is correct, imposing the equilibrium restrictions increases (estimator) efficiency. However, if it is not, the restrictions are misspecified and the estimates of price changes are inconsistent. We, in contrast, make no assumptions about the way in which prices are determined. Instead, using the insight of De Loecker and Warzynski (2012), we estimate markups as wedges between output elasticities and expenditure shares in a production function context. Since we think that the simple models that are commonly used for merger simulations (i.,e., Bertrand or Cournot in a static context) are only rough approximations to real world competition, we prefer not to rely on equilibrium restrictions.

Whereas the standard simulation model defines a merger as a consolidation of control rights — the ability to set prices — we in contrast model a merger as a consolidation of physical assets, both variable and fixed. This difference allows us to forecast not only increases in markups but also changes in efficiency.

Our approach unifies and extends three literatures: the use of discrete Divisia indices to measure TFP growth that began with Jorgenson and Griliches (1967),[4] the calculation of markups without specifying how firms compete in product markets that was initiated by Hall (1988), and the estimation of productivity in the presence of endogenous variable factors and selection bias that dates to Olley and Pakes (1996). Each of those ideas has been extended in many directions by numerous researchers.[5] However, by combining the three strands, we think that we provide a unifying supply–side approach to market power and efficiency measurement that can be complementary to the traditional demand-focused approach to merger analysis.[6]

To estimate the production function and evaluate the Molson Coors merger we use an unbalanced panel of publicly traded US and Canadian brewing companies between 1950 and 2012. During the first half of that period, all brewers were in the mass production sector. However, after 1980, the craft sector gradually began to assume importance. The firms in our data are therefore heterogeneous because they produce in different countries and sectors and we hypothesize that there will be systematic differences in markups and returns to scale across those groups. Moreover, even within sectors and countries, there could be substantial cross sectional variation in technology and market power. We therefore estimate a translog production function that allows returns to scale and market power measures to vary across firms and time.

The use of standard data sets such as Compustat is fraught with problems. For example, typically revenue from output and expenditures on inputs are recorded but not prices and quantities. We therefore constructed firm specific input and output prices that allow us to measure physical inputs and output accurately. In addition, with data on publicly traded firms, selection is an important issue since a firm can disappear from the data not only because it fails but also because it merges with another firm that need not be in the data (e.g., a firm that is not North American) or because it goes private, We therefore specify a model of selection that distinguishes among those possibilities.

[3] See, e.g., Nevo (2000), Pinkse and Slade (2004), Ivaldi and Verboven (2005), and Jeziorski (2014a).

[4] For example, the Tornqvist is a discrete Divisia index.

[5] De Loecker and Warzynski (2012) combine the second and third strands.

[6] In particular, our simulation does not forecast price changes of firms that are not involved in the merger.

To summarize, we find that returns to scale increase over time and are higher in Canada compared to the US and lower in the craft sector compared to mass production. When we focus on US mass production, we find that a large market share is associated with higher markups and greater returns to scale. A more concentrated industry, in contrast, is associated with lower TFF growth. Finally, our simulation of the Molson Coors merger forecasts higher markups and increased scale economies. When we compare our forecasts to realizations, we find that, by the second year post merger, forecast efficiencies had been realized. Markups, in contrast, remained higher than forecast and we attempt to explain this discrepancy.

The next section, which describes the North American brewing industry and the Molson Coors merger, is followed by discussions of antitrust policy towards mergers in North America and the use of merger simulations. We then discuss related literature and present the theoretical and empirical models, the data, and finally, our empirical findings and conclusions.

2 The North American Brewing Industry

2.1 The Industry

The North American brewing industry consists of brewers in the US and Canada. Historically the US brewing industry, which was relatively unconcentrated, had many national and regional brewers. For example, in 1950, the US four firm concentration ratio — the percent of industry output that is produced by the four largest firms — was equal to 22. However, two factors changed the structure of the market: due to antitrust vigilance, the national brewers grew internally rather than through mergers, whereas some regional brewers, who were not subject to antitrust scrutiny, grew by merging.[7] As a result, by 1985, the US four firm concentration ratio had risen to 82.

Early mergers in US brewing, which occurred in a shakeout that allowed the remaining firms to achieve economies of scale in production, distribution, and marketing, were of two sorts: mergers to achieve synergies and growth (the fate of Pabst, for example) and mergers for asset stripping (the fate of, e.g., Stroh, Schlitz, and Heileman).[8] Mergers for asset stripping, which were common during the wars of attrition in the 1970s, allowed failing firms to exit gradually as, without investment, their physical and intellectual (brand) capital deteriorated.

In Canada, in contrast, the big three — Molson. Labbatt, and Canadian Breweries (later Carling O'Keefe) — became dominant in the 1950s. Moreover, in contrast to the US where national firms grew internally, the major Canadian firms expanded through mergers with smaller brewers and, by 1985, the three firm concentration ratio had risen to 96. The Canadian situation was very different because interprovincial trade was banned, which meant that the only way to become a national brewer was to acquire or establish a brewery in each province.

The situation has changed dramatically in both countries in the last three decades. First, there have been mergers and joint ventures among the largest firms (e.g., Molson/Carling O'Keefe in Canada and Miller/Coors

[7] See Elzinga and Swisher (2005) for an analysis of brewing mergers and US antitrust policy during this period.

[8] Tremblay and Tremblay (2009) call mergers for asset stripping the devolution strategy.

in the US). Second, cross border mergers such as that between Molson and Coors have also occurred. Third, mergers with non North American brewers (e.g., Labatt/Interbrew in Canada and Anheuser Busch/Inbev (ABI) in the US) have become common. Fourth, the craft beer movement was born and gained popularity in both countries. And finally, the share of imports has risen. As a result, the brewing industries in the two countries look rather similar today. Both are essentially duopolies with a fringe of small regional and craft brewers and both are dominated by foreign owned firms. In addition, the two markets are more integrated than they were in the past.

In the same period, the brewing industry witnessed technical changes that increased efficiency. For example, table 1 in Kerkvliet, Nebesky, Tremblay, and Tremblay (1998) contains estimates of the minimum efficient scale (MES) in brewing that were obtained by various researchers, and those estimates show how MES increased with time, particularly in the 1970s. The authors attribute that increase to the introduction of super breweries and to advances in packaging techniques, particularly in canning. In addition, improvements in shipping, such as the widespread availability of refrigerated trucks, allowed brewers to expand their geographic markets.

Brewing consists of combining malt, barley, and other ingredients with water and allowing the liquid to ferment. When this has been accomplished, the fermented beverage is packaged into bottles, cans, and kegs and the packaged goods are shipped to market. Since there can be economies of scale in brewing and packaging, a medium sized brewer faces a tradeoff between having one large brewery, which involves lower production but higher shipping costs, versus several smaller breweries with lower shipping but higher production costs. In contrast, a large national brewer with several large breweries can achieve economies of scale in both production and distribution. Moreover, in addition to the standard merger motives, mergers are more apt to result in synergies when there are large overlaps in markets but few overlaps in brewery locations.

Brewing is the first phase in a three tier system that consists of production, distribution, and sales. In the US, with the exception of microbreweries, federal law prohibits integration of the three phases. In Canada, in contrast, distribution and sales are regulated by the provinces. However, in most provinces, the downstream phases are separate from brewing and are often handled by a provincial liquor control board. Nevertheless, in both countries, the brewer incurs the freight costs of shipping product to distribution points that are closer to markets.

Between 1950 and 1980, North American brewing consisted almost entirely of mass production. However, at some time around 1980, the craft beer sector took off. The Brewers Association defines a craft brewer as small — production less than 6 million barrels per year,[9] independent — not owned or controlled by an alcohol industry member that is not itself a craft brewer, and traditional — using traditional or innovative brewing ingredients. The craft sector, which focuses on darker beers such as ales rather than the lagers that are the mainstay of mass production, has grown until, in 2015, it constituted 12% of the US market. Its popularity is due in part to a reaction against the light and rather flavorless beers that had become the mainstay of conventional brewing.

[9] This restriction changed recently from 2 to 6 million to accommodate the Boston Beer Co. (brewer of Samuel Adams), which had grown rapidly.

During the same time period, the share of imported beers began to rise until today it is nearly 15%. Early on, most imported beers were European. Today, however, of the five most popular brands imported into the US, four come from Mexico.

The North American industry has thus changed from one with two isolated geographic markets, each producing a rather homogeneous product that was protected from imports, to a more integrated geographic market with two sectors, conventional and craft, each facing substantial import competition. These changes mean that consumers have more choice. Indeed, even the largest brewers now produce craft like brands. The changes could also mean, however, that consumers pay higher prices.

We focus on the Molson Coors cross–border merger. However, subsequent mergers have further increased concentration in the industry, most notably the Miller Coors joint venture that occurred in 2008.[10] Moreover, if the proposed ABI Miller merger is allowed to go forward, the merged firm will supply about one third of the world's beer. It is still not clear what this mega merger would mean for Molson Coors. However, one possibility is that Molson Coors will become the sole owner of Miller Coors.

This brief snapshot of the industry suggests that industry concentration, returns to scale, and import penetration have increased over time. Moreover, the market might increasingly be characterized by both noncompetitive pricing and efficiencies associated with large size. In what follows we assess those possibilities more formally.

2.2 The Molson Coors Merger

The Molson Coors merger, which was announced in July of 2004 and consummated in early February of 2005, united the numbers two and three brewers in Canada and the US, respectively, and created the fifth largest brewer in the world. Coors paid $3.5 billion to acquire Molson. However, instead of consolidating the two head offices, Canadian operations, including Canadian sales of Coors brands, are headquartered in Montreal, Que. whereas Coors operations, including US sales of Molson brands, are headquartered in Golden, CO.

Coors 2004 Annual Report claims that, were the merger to occur, $175 million worth of synergies and merger–related cost savings would be realized. Furthermore, those efficiencies would principally be due to lower marketing and distribution costs, greater financial strength, facilitated geographic expansion, and increased tax benefits. In addition, the Coors brewery in Memphis, TN would close at the end of 2006 whereas two new breweries, one in the Shenandoah Valley and the other in Moncton, NB, would open in 2007. Although the brewery closure and openings have occurred as scheduled, it is not clear why they are merger–related. The other efficiency claims, which are more difficult to assess, are more clearly tied to the merger.

The Molson Coors merger was atypical among beer mergers in that rationalization of production did not play a major role. Indeed, Molson already produced Coors brands under license and Coors produced Molson brands. However, there were two separate entities, Coors Canada and Molson USA, that were responsible

[10] This joint venture is studied by Ashenfelter, Hosken, and Weinberg (2015) and Miller and Weinberg (2015) .

for production, distribution, and marketing in the foreign country, and those entities were eliminated by the merger.

Unlike the merging parties, who stress efficiencies, competition authorities are principally concerned with the possibility that a merger might lead to greater market power. For example, since Coors Light and Molson Canadian Light are very similar, with joint price determination, there might be an incentive to raise their prices.

In sum, although there were credible efficiency gains that could be expected, some of the usual gains from a merger, such as elimination of duplicate head offices, were not planned. Moreover, although all beer brands are substitutes for one another, no Molson brand is the closest substitute for a Coors brand and vice versa. For example, in the US, Bud Light is a closer substitute for Coors Light. We therefore expect to see some changes in both efficiency and market power but those changes might be modest.

2.3 Market Definition

With our merger simulation, there is no need to define a market. In particular, the model can be estimated using data from a subset of firms in the industry. However, when we assess US market structure and performance, we must specify one. Defining a market is always controversial, and with beer, submarkets in both the product and geographic dimensions have surfaced. Although in general antitrust agencies tend to take a narrow view of markets while merging parties take a broad view, in recent cases the US Department of Justice (DOJ) has advocated narrow geographic markets for beer — metropolitan statistical areas (MSAs) — but a broad product market. The firms, in contrast, both in documents and in testimony, are more interested in product submarkets (i.e., super premium, premium, value, etc.).

The agencies have in general favored narrow geographic markets due to their asymmetric treatment of costs and benefits. In particular, in the US, if efficiency is given any weight at all, it is given less weight than harm.[11] This means that the agencies focus on consumers, demand, and substitution possibilities, and clearly consumers in New York do not purchase beer in California. However, we focus on firms and their technologies. In particular, we wish to see if mergers change not just markups, but also returns to scale and productivity. The later two variables are inherently at the level of the firm. Moreover, firms produce all of their brands in a few breweries — Coors had just two — and coordinate production nationally. Finally, broad pricing strategies are centrally determined. Given our firm–centric focus, we think that a US beer market makes more sense. This does not mean that, for example, prices are the same in New York and San Francisco. However, there is probably as much price variation between inner cities and suburbs within MSAs as there is across MSAs, and the former are in the same DOJ market for beer.

There are precedents in the literature for adopting a broad definition of the geographic market for beer. For example, in their book *The US Brewing Industry*, Tremblay and Tremblay (2009, p. 44–45) claim that, although in earlier years beer markets were regional, they became national in the 70s, since by that time the large producers were selling in all parts of the country. Their stance is not surprising, given that their book is principally concerned with firms and markets rather than consumer tastes.

[11] For more on asymmetric merger standards, see Crane (2014).

3 Competition Policy Towards Mergers in North America

3.1 The Process

There are many similarities between the ways in which mergers are evaluated by competition authorities in the US and Canada and one important difference. We therefore describe the process in the US and then discuss the essential difference between the two countries.

3.1.1 US Merger Policy

The Hart Scott Rodino Act of 1976 requires parties making asset acquisitions that meet certain dollar thresholds to file a pre merger notification. The agency, DOJ or Federal Trade Commission (FTC), then has 30 days in which to decide whether to file a second request for more information. After that, if the agency feels that there are significant competitive concerns, it can seek a preliminary injunction, and during the evaluation process, the parties must behave as arm's length competitors. Most litigated mergers are decided on the basis of preliminary injunctions rather than trials.

The 2010 US Horizontal Merger Guidelines, like previous Guidelines, are grounded in Section 7 of the Clayton Act of 1914, which prohibits a merger if "the effect of such acquisition may be substantially to lessen competition." This language does not mention efficiencies and makes it difficult to mount an efficiencies defense (Blair and Haynes, 2011). Nevertheless, merger practice in the US has gradually become more sympathetic towards considering countervailing factors (Kolasky and Dick, 2003). In practice, if the agency decides that there will be competitive harm, the merging parties can seek to establish that there are offsetting efficiencies.

There are three peculiarities about this process. First, price and cost changes are determined in a sequential process (with harm determined first) by different groups. In particular, harm is assessed by predicting price changes while holding costs constant. Firms, in contrast, do not ignore costs when setting prices. Instead, the two are jointly determined.

Second, the legal standard of proof is much more stringent for the establishment of efficiencies than for harm. Indeed, since most cases are decided on the basis of preliminary injunctions, the government's burden is merely to prove a substantial likelihood that it will eventually be able to show probable cause to block the merger. On the other hand, efficiencies must be proven to a high degree of certainty (Crane, 2014). This dual standard is justified in part by the claim that "efficiencies are difficult to verify and quantify, in part because much of the information is uniquely in the possession of the merging firms. Moreover, efficiencies projected reasonably and in good faith by the merging firms may not be realized." (Section 19, 2.2.1 of the 2010 Guidelines). There is no mention, however, that anticompetitive effects are also difficult to project accurately and also might not be realized.

Finally, US law has adopted a consumer welfare standard in which efficiencies can cancel anticompetitive effects only if they are likely to reverse the merger's competitive harm to customers. In practice this means that they must offset the price increases, and cost savings per se are given zero weight.

3.1.2 Canadian Merger Policy

Canadian merger policy is grounded in the Competition Act of 1986 (the Act), which permits mergers to be challenged when they are likely to substantially lessen or prevent competition. However, in contrast to the US, the Act act adopts a total welfare standard that balances efficiency gains against anticompetitive effects. In other words, whereas the US gives zero weight to cost savings per se, Canadian authorities take them into account. Like the US, however, pre merger notification is required for large transactions. Moreover, there is a two step procedure in which the Bureau of Competition must establish the anticompetitive damage that is associated with a merger, whereas the merging parties must show all other aspects of the tradeoff, including the nature, magnitude and timeliness of efficiency gains and whether such gains are greater than and offset the anticompetitive effects. In addition, they must show that the gains are likely to occur and that they are specific to the merger. Cost savings can include economies of scale, scope, and density, and savings from specialization and the elimination of duplication.

4 Related Literature

4.1 Merger Simulations

4.1.1 Models that Require Only Per Merger Data

Competition authorities in both countries must assess anticompetitive effects and, in many instances, they have done this with the aid of merger simulations that involve only pre merger data, (e.g., Nevo, 2000; Pinkse and Slade, 2004; Ivaldi and Verboven, 2005). Since these are the sorts of tools that competition authorities commonly use, we describe them in detail before discussing simulations that require data that is not usually available.

A typical merger price simulation is a model of firms that produce a number of differentiated products and engage in price competition.[12] Firms or players choose the prices of the products that they own, taking into account the choices of other players. A market structure is a partition of the product space where the jth element of the partition is the set of products that the jth firm owns. A merger is modeled as a reduction in the number of players — a coarser partition. In particular, the products that were produced by two firms are now produced by one, and a single player can choose their prices. Since firms internalize the increase in demand for their own products that result from their own price increases, but do not internalize the comparable increase in demand for rival products, the merged firm has a unilateral incentive to raise prices. The degree to which this is true, however, depends on substitutability among products, and accurate price forecasts depend on accurate estimates of own and cross price elasticities.

In order to implement a simulation, one must estimate the demand and marginal cost for each of the N products[13]. One can then solve the N first–order conditions for the N prices twice, first using the old and

[12] In contrast to these price simulations, Jeziorski (2014a) assesses mergers where advertising is assumed to be the choice variable.

[13] A number of demand specifications have been used. For example, Nevo (2000) and Jeziorski (2014a) estimate the random coefficients model of Berry, Levinsohn, and Pakes (1995); Pinkse and Slade (2004) use the distance metric model of Pinkse, Slade,

second the new ownership structure to obtain predicted pre and post merger prices.

Static equilibrium simulations are sensitive to a number of factors including the equilibrium assumption and the specifications of demand and marginal costs. Some researchers have tested their equilibrium assumptions and have not rejected them (e.g., Nevo, 2001; Slade, 2004b). Unfortunately, estimated demand elasticities and marginal costs can be very sensitive to the specification of the demand and cost functions, as shown in Slade (2009). Finally and most importantly for our purposes, with simulations that are commonly used by competition authorities, costs do not change post merger. In other words, efficiencies are not estimated and price changes can therefore be overestimated.

4.1.2 Models that Require More Complex Data

Some researchers have estimated merger simulation models that require data of the sort that is not usually available to competition authorities, either because one must observe a large number of changes in assets (i.e., mergers) or because post merger data is essential. Nevertheless, that literature is related because it attempts to assess merger–related competitive harm and/or efficiencies.

A number of interesting things can be done with data that include both pre and post merger periods. For example, Miller and Weinberg (2015) assume that pre merger competition is Bertrand but introduce a conduct parameter post merger, which allows them to test if the merger facilitated tacit collusion (coordinated effects). Although models with conduct parameters were heavily criticized (e.g., Makowski, 1987; Corts, 1999), recent research shows how they can be identified (Berry and Haile, 2014). In addition, Miller and Weinberg allow the parameters of the marginal cost function to vary with the merger, which allows them to evaluate short–run efficiencies.

In contrast, Jeziorski (2014b) looks more deeply at merger–related efficiencies, both long and short run. Like Jeziorski (2014a), the stage game is Nash in advertising. However, the one–shot game is embedded in a model of dynamic discrete choice and inference is based on revealed preference as in Bajari, Benkard, and Levin (2007). Specifically, when the model predicts a merger that does not occur, it is assumed that estimated efficiencies are too large, and when a merger is not predicted but occurs, it is assumed that efficiencies are too small. In this way, both marginal and fixed costs can be uncovered. Unfortunately, estimation requires data on a very large number of mergers.

In the next section, we develop a merger simulation model that is based on the firms' technology, not it's demand. Furthermore, as with standard merger simulations, it requires only pre merger data. However, our model allows us to forecast merger–related changes in markups that do not rely on an equilibrium assumption, and those markups are estimated jointly with merger–related efficiencies. Moreover, unlike equilibrium price simulations, we do not require data for all of the principal firms. This can be a major advantage since competition agencies can subpoena data from the merging parties but not from their rivals. Like typical static equilibrium price simulations, however, we assume that the relationship that determines markups does not

and Brett (2002); and Ivaldi and Verboven (2005) use a nested logit. Marginal costs are usually specified as parameters or functions of observable cost shifters in the first order conditions that rationalize the assumptions on demand and equilibrium.

change as a result of the merger. Finally, unlike equilibrium simulations, we do not forecast price changes for nonmerging firms.

4.2 Productivity Measurement

There is a large macroeconomic literature that is devoted to estimating aggregate productivity growth and much of this research was inspired by the work of Solow (1957). Typically, researchers start with a constant returns to scale Cobb Douglas production function and equate productivity growth with the percentage change in output minus a share weighted percentage change in inputs — the Solow residual. Moreover, it is common to estimate the production function by OLS, which makes the exercise very simple

The potential bias in OLS estimates of production functions, however, has long been recognized (see e.g., Marschak and Andrews, 1944). This bias results from the possible correlation between input levels and firm level productivity shocks. Specifically, when firms experience a large productivity shock, they might respond by using more inputs. Applied economists have devised alternatives to OLS that attempt to circumvent this problem. Most use either a variant of the method developed by Olley and Pakes (1996) (OP) and extended by Levinsohn and Petrin (2003), Ackerberg, Caves, and Frazer (2006), and Gandhi, Navarro, and Rivers (2013) or the GMM methods proposed by Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (2000). We have chosen to focus on the former since those methods are based on a full behavioral model that uncovers unobserved productivity as a function of observed decisions. Specifically, they employ an inverse function of input choices to control for unobserved productivity in the production function and to overcome the endogeneity problem associated with OLS estimation discussed above. Since this literature is summarized in Ackerberg, Benkard, Berry, and Pakes (2007), we do not discuss it in detail here.

Most researchers in the OP tradition assume that firms are identical up to a productivity shock. However, even within narrowly defined industries, firms can be heterogeneous. In particular, they can differ because they have different production functions.[14] In addition, even when firms have the same production function, if the technology is flexible, some can be producing in the region of increasing returns whereas others can be in the opposite region. Finally, firms can have different degrees of market power.

When a Cobb Douglas production function is estimated, it is common to equate productivity growth with technical change — a shift in the production function holding inputs constant.[15] However, when the technology is flexible, the growth of TFP also depends on economies of scale, which can vary by firm and over time. We borrow from the index–number productivity–measurement literature to derive a measure of TFP growth that captures both aspects of the problem.[16] In particular, our TFP measure — an index of output divided by an index of inputs — changes when a firm produces in a different region of its production function as well as when the firm's stock of knowledge is augmented.

Denny, Fuss, and Waverman (1981) develop an index–number model of TFP growth based on a cost

[14] Kasahara, Schrimpf, and Suzuki (2015), Balat, Brambilla, and Sasaki (2016), and Hoderlein, Fox, Hadad, and Petrin (2016) consider unobserved heterogeneity whereas we model observed heterogeneity.

[15] A Cobb Douglas can exhibit nonconstant returns to scale. However, with that function, scale economies are the same for all firms and therefore the correction would not change productivity comparisons.

[16] See Diewert and Nakamura (2007) for a survey of the index–number literature.

function that incorporates both technical change, a shift in the cost function, and changes due to economies of scale, a movement along that function. We, however, do not estimate a cost function, and their results are therefore not directly applicable. In particular, economies of scale that are estimated from a cost function are based on the assumption that inputs expand along the cost minimizing or expansion path, whereas those estimated from a production function are based on the assumption that they expand in equal proportions. However, Caves, Christensen, and Swanson (1981) show that, locally (i.e., at a point), the two measures are the same. In section 5.3.4, we discuss how one can use these results to obtain a theoretically consistent decomposition of TFP growth from our estimated production function.

4.3 Markup Measurement

In an early paper, Hall (1988) demonstrated that markups can be estimated from a production function. His research relaxes one of the frequently used TFP growth assumptions — competitive output markets — but maintains another — constant returns to scale. Under constant returns, revenue shares are equal to cost shares, which sum to one. It is therefore possible to divide both sides of the production function by the single fixed factor, K,[17] leaving the variable inputs on the right–hand–side, which means that it is not necessary to measure the user cost of capital. However, noncompetitive pricing causes the Solow residual to deviate from the rate of growth of TFP. In particular, when price exceeds marginal cost, input growth is associated with disproportional output growth. Hall uses this insight to devise a test for deviations from competitive pricing and to show how average markups can be estimated.

Several decades later, De Loecker and Warzynski (2012) extended Hall's research by relaxing the assumption of constant returns to scale. Moreover, as with the original paper, their generalization is achieved without the need to measure the user cost of capital. Their insight is that the output elasticity of a variable input equals its share of revenue only when price equals marginal cost. However, under imperfect competition, the two are equal only if revenue is evaluated using the shadow price of output — marginal cost (MC) — instead of the market price, P_y. Imperfect competition therefore, drives a wedge between the two that depends on the markup, P_y/MC. Furthermore, instead of implementing their calculations in a Solow framework like Hall, they formulate a model of production that is based on the model of Olley and Pakes (1996) and its extensions, thus overcoming the endogeneity problem that is discussed above. Moreover, their extension allows them to recover firm and time–specific markups rather than simple averages.

5 The Theoretical Model

5.1 The Research Framework

In contrast to most merger price simulations, which look at multiproduct firms that produce differentiated products, we consider a single homogeneous product.[18] Indeed, from the production point of view that we

[17] Hall (1988), like others in the markup literature, assumes that there is a single fixed factor.

[18] With multiple outputs, one can construct a share weighted index of outputs, as is common in the production function/productivity literature. One would then obtain the markup on the firm's bundle of products.

adopt, the product — beer — is relatively homogeneous, at least for mass production. For example, Tremblay and Tremblay (2009) state that "Although individual consumers have strong opinions about which brands are best, it is difficult to identify real quality differences among different brands of the mass–producing brewers." (page 8). Nevertheless, there are differences across sectors and countries.

Our estimated production functions, which differ by country – the US and Canada – and sector — mass production and craft — yield estimates of markups, scale economies, and technical change for each firm in each year, and those estimates are combined to yield a measure of TFP growth. The measures are then used to assess a number of aspects of the industry. For example, we compare TFP growth, markups, and returns to scale across Canadian and US firms and across craft and mass production. In addition, for US mass production, we assess associations between various measures of market structure and our measures of performance. Finally, we use our structural model to forecast changes in efficiency gains and market power that are associated with the merger between Molson and Coors, and we compare the forecasts to realized estimates.

In what follows, we discuss the proposed framework more formally.

5.2 The Production Function

We adopt a fairly standard Olley and Pakes (1996) framework and suppose that a vector of variable inputs X and fixed inputs K are used to produce a homogeneous output Y according to the production function

$$Y_{jt} = A_{s(j)c(j)t} F_{s(j)c(j)}(X_{jt}, K_{jt}) e^{\eta_{jt}}, \tag{1}$$

where j is a firm, t is a year, A is the state of technology, F is a function of the conventional inputs, the subscripts $s(j)$ and $c(j)$ are the sector and country to which j belongs, and η is a shock that is conditionally mean independent of current and past inputs. In what follows, we drop the s and c subscripts.

Equation (1) can then be written as[19]

$$y_{jt} = a_{jt} + f(x_{jt}, k_{jt}) + \eta_{jt}, \qquad a_{jt} = \beta_0 + \beta_t t + \omega_{jt}, \tag{2}$$

where all variables are in natural logarithms and the state of technology consists of a constant, a trend, and unobserved productivity, ω. In addition, we assume that the state variables ω and k evolve according to

$$\omega_{jt} = g(\omega_{j,t-1}) + \xi_{jt}, \qquad k_{jt} = (1-\delta)k_{j,t-1} + i_{jt}, \qquad i_{jt} = i(k_{jt}, \omega_{jt}).$$

In other words, ω is a first order Markov process, whereas k decays at the depreciation rate δ and is augmented by investment i, which is a function of the state variables.

Finally, we make the OP strict monotonicity assumption on the investment function — that i is monotonic

[19] It is not always possible to write (1) as (2). However, it is possible with the production functions that are commonly used.

in the unobservable ω — which implies that the investment function can be inverted and we can write

$$\omega_{jt} = h(k_{jt}, i_{jt}).$$

To anticipate, there are two equations that identify the parameters of the model: $y_{jt} = \beta_0 + \beta_t t + f(x_{jt}, k_{jt}) + h(k_{jt}, i_{jt}) + \eta_{jt}$ and $y_{jt} = \beta_0 + \beta_t t + f(x_{jt}, k_{jt}) + g[h(k_{j,t-1}, i_{j,t-1})] + u_{jt}$. We adopt the Ackerberg, Caves, and Frazer (2006) framework and do not use the first equation to identify any production function parameters, and we estimate the model by GMM as suggested by Wooldridge (2009). Details of the estimation are discussed in section 7.2.

5.3 The Performance Measures

Given estimates of the production function, we can construct four performance measures: the price/cost markup (PCM), returns to scale (RTS), technical change (TECH), and the rate of growth of total factor productivity (TFPG). The first three are fairly standard, and we devote more attention to the last.

5.3.1 The Price Cost Markup

We define the price cost markup (PMC) as output price P_y divided by marginal cost, MC, and we use the method that is outlined in De Loecker and Warzynski (2012) to measure PCM. Specifically, let γ_{yv} be the elasticity of output with respect to some variable input, V, $\gamma_{yvjt} = \partial f(x_{jt}, k_{jt}) / \partial v_{jt}$, where v is the log of one of the variable inputs, X. Then

$$\mathrm{PCM}_{jt} = \gamma_{yvjt} \frac{P_{yjt} Y_{jt}}{P_{vjt} V_{jt}} = \frac{\gamma_{yvjt}}{r_{vjt}}, \tag{3}$$

where P_v is the price of the variable input. In other words, the price cost markup equals the elasticity of output with respect to the variable input times total revenue divided by expenditure on the variable input, which equals the elasticity divided by the input's share of revenue, r_v.

We use PCM as our measure of market power as it is a proxy for the related noncompetitive distortion. Indeed, a fundamental condition for efficienct allocation of resources is that price equal marginal cost, and any deviation is a efficiency loss. Moreover, with linear demand and costs in the spirit of Williamson (1968), the deadweight loss depends only on the deviation of price from marginal cost and the slope of the demand function.

5.3.2 Returns to Scale

Returns to scale is defined as the proportionate change in output that is due to an equiproportionate change in all inputs. If we define $Z = (X^T, K^T)^T$, then our measure of *local* returns to scale is

$$\mathrm{RTS}_{jt} = \left. \frac{\lambda \partial \log F(\lambda Z_{jt})}{\partial \lambda} \right|_{\lambda=1} = \iota' \frac{\partial f}{\partial z}(z_{jt}) \tag{4}$$

where ι is a vector of ones, the proportionate change in inputs is $\lambda - 1$ and the partial derivative is evaluated at $\lambda = 1$. With a Cobb Douglas production function, RTS is constant. When the technology is flexible, however, RTS varies with Z.

We define two measures of static efficiency gains — the percentage change in output holding inputs constant, which is is used ex ante, and the actual change in input and output usage from pre to post merger, which is used ex post — and both are based on returns to scale. Moreover, both pick broader organizational changes and not just changes that occur at the plant level.

5.3.3 Technical Change

Technical change is a temporal shift in the production function, holding all inputs constant. With the production function (2), technical change is

$$\text{TECH}_{jt} = a_{jt} - a_{jt-1} = \Delta a_{jt} = \beta_t + \omega_{jt} - \omega_{jt-1} = \beta_t + \Delta\omega_{jt}, \tag{5}$$

which consists of a deterministic and a stochastic component.

5.3.4 TFP Growth

Finally, we need a measure of total factor productivity growth and, to obtain one, we borrow from the index–number literature. Define $\Delta y_t = y_t - y_{t-1}$ and $\Delta z_t = \sum_k r_{kt}(z_{kt} - z_{kt-1})$, where z is the vector $(x^T, k^T)^T$ and r_k is the kth input's share of revenue. Under the assumptions of constant returns and competitive pricing in the output market, Jorgenson and Griliches (1967) show that TFP growth is $\Delta\text{TFP}/\text{TFP} = \Delta y_t - \Delta z_t = \Delta a(t)$. That is, the rate of growth of TFP is equal to the rate of technical change. However, we assume neither constant returns nor competitive pricing, so the two will be different in general.

Under the Jorgenson Griliches assumptions, revenue shares are equal to cost shares, $s_k = P_{zk}Z_k/C$, where C is total cost. However, violation of either assumption, constant returns or competitive markups, causes the two to diverge. In particular, whereas cost shares are invariant to markups, revenue shares are not. Moreover, cost shares always sum to one whereas revenue shares do not. Cost shares are therefore preferred.

A standard measure of *local* returns to scale obtained from a cost function is the reciprocal of the percentage change in total cost due to a percentage change in Y, holding input prices constant,

$$\text{RTS}_C = \frac{1}{\frac{\partial \log C}{\partial \log Y}} = \frac{1}{\gamma_{cy}}.$$

where γ_{cy} is the elasticity of total cost with respect to output. Moreover, Denny, Fuss, and Waverman (1981) show that, under nonconstant returns, the growth in TFP is

$$\frac{d\log \text{TFP}_C}{dt} = \frac{da}{dt} + \left(1 - \gamma_{cy}\right)\frac{dy}{dt}. \tag{6}$$

In other words, TFP growth consists of not only technical change but also a term that depends on returns to scale. For example, under increasing returns ($\text{RTS}_C > 1$), the rate of growth of productivity is greater than the rate of technical change as firms realize scale economies. Indeed, TFP growth is due to both a shift in the production function and a movement along that function.

The expression in (6) depends on economies of scale that are obtained from a cost function and we have measures from a production function. Unfortunately, the two are not the same. In particular, scale economies obtained from a production function are based on the assumption that all inputs increase at the same rate whereas those obtained from a cost function are based on the assumption that inputs increase along the expansion, or cost minimizing, path. Nevertheless, Caves, Christensen, and Swanson (1981) show that, at any point (i.e., when one evaluates derivatives, not discrete changes), the two measures are equal, which means that we can substitute RTS for RTS_C in (6).

Equation (6) involves infinitesimal changes and we have discrete changes. Our measure of *local* TFP growth, which is a discrete approximation to (6), is

$$\text{TFPG}_{jt} = \Delta a_{jt} + \left[1 - \frac{1}{(\text{RTS}_{jt} + \text{RTS}_{jt-1})/2}\right] \Delta y_{jt}. \tag{7}$$

In other words, we do not specify either equiproportionate input changes or changes along the expansion path but instead look at actual changes.

Finally, note that, by constructing a measure of TFP growth that was derived from a cost function, we have purged that measure of spurious dependence on markups. For example, if one constructs a TFP growth measure using revenue shares and assumes that capital's share is a residual, as is often done, estimates of capital's share rise with output price and TFP growth is underestimated.

We use changes in TFP growth in response to a change such as a merger as our measure of dynamic efficiency gain, since TFP growth encompasses technical change.

Once we have our performance measures, we can relate those measures to a firm's country and sector in a regression framework. With this analysis, we do not interpret the coefficients as causal but simply test whether, for example, there is a systematic tendency for different groups of firms to have different markups.[20] In addition, we assess how the structure of a single market — US mass production — correlates with firm performance. Finally, we can use our structural model to assess how our performance measures respond to a specific merger or some other change in the market.

5.4 The Merger Simulation

We model a merger as a consolidation of assets, both variable and fixed. Once the production function has been estimated, the simulation exercise is simple and can be performed with the aid of a calculator or spreadsheet.

Suppose that we have estimated the production function (2) using a panel of firms that includes, but generally is not limited to, the merging firms.[21] Suppose further that a merger between i and j is proposed in

[20] This is analogous to the analysis of markups and export status in De Loecker and Warzynski (2012).

[21] In particular, it is advisable to have data on some of the largest firms so that the simulation does not involve extrapolating out of

year t^*. Let the inputs of firm ℓ in period t^* be $Z_{\ell t^*} = (X^T_{\ell t^*}, K^T_{\ell t^*})^T$, $\ell = i, j$. Then, with the simplest calculation, the merged firm, m, will have the inputs $Z_{mt^*} = Z_{it^*} + Z_{jt^*}$. One can reevaluate equations (3), (4), (5), and (7) using the new inputs to obtain forecasts of the performance of the merged firm. Moreover, with post merger data, one can compare the forecasts to realized performance ex post.

As noted earlier, we define two measures of static efficiency gain. The first, which we use for the ex ante simulation, is the forecast percentage change in output holding inputs constant. The second, which we use for ex post evaluation, involves actual changes from pre to post merger. That measure is the realized percentage change in output minus the realized percentage change in inputs.

Formally, if i and j merge, we evaluate the ex ante ratio $F(Z_i + Z_j)/[F(Z_i) + F(Z_j)]$, where Z_ℓ is the pre merger vector of variable and fixed inputs for firm ℓ, $\ell = i, j$. Under increasing returns, that ratio will be greater than one. Our ex ante forecast measure of efficiency gain is therefore

$$\text{EFF}_{\text{forecast}} = \frac{F(Z_i + Z_j)}{F(Z_i) + F(Z_j)} - 1. \tag{8}$$

Since the firm can reoptimize after a merger, our simulated efficiency gain should be an underestimate.[22] On the other hand, the new firm might find it difficult to coordinate operations such as information systems, which would imply an overestimate.

Ex post, we observe actual changes in production and input usage. Our ex post measure of efficiency gain is therefore the percentage change in output minus the share weighted percentage change in inputs,

$$\text{EFF}_{\text{realized}} = \frac{Y_m}{Y_i + Y_j} - \sum_k s_k \frac{Z_{mk}}{Z_{ik} + Z_{jk}}, \tag{9}$$

where s_k is the average of input k's pre and post expenditure shares and the subscript m stands for post merger values for the merged firm.

Both measures of static efficiency gains are based on returns to scale. In particular, under increasing returns, with the first, holding inputs constant the output of the merged firm should be greater than the sum of the outputs of the two separate firms. With the second, the percentage change in output should be greater than the percentage change in inputs.

Since we look at total expenditure on labor, including management and workers involved in advertising, R&D, and distribution, and total expenditure on other inputs, our measures encompass economies of density, procurement efficiencies, savings from specialization, and the elimination of duplication as the firm expands.

The simulation model can also be used to evaluate the consequences of claims of the merging parties. For example, if merger specific plant closures are expected, one can adjust the pre merger inputs to reflect this fact. In other words, one can subtract the capacity, work force, and materials usage of the plants that will close from the merged firm's inputs and evaluate equations (3)–(7) again. Finally, it is possible to assess the efficiency effects of mandated divestitures in the same way.

sample.

[22] However, firms were already choosing inputs optimally in each plant, which implies that most gains come from elimination of duplication in administration, marketing, and transport costs that come with increased size.

Although not our focus, we also forecast changes in markups and those forecasts deserve further discussion. Recall that the markup — price divided by marginal cost — depends on the output elasticity and revenue share of a variable input. Given inputs and prices, we can calculate new input shares. Moreover, since we have assumed that the production function is the same for all firms in the mass production sector, we can calculate the output elasticity for the merged firm. In order to avoid out of sample extrapolation, however, it is advisable to have data on firms that are larger than the merging parties.[23] Indeed, we are assuming that the merged firm will be similar to an existing firm of similar size.

Like price simulations, our simulation model is based on the assumption that the method of price determination does not change after a merger. In particular, it is assumed that the market does not become more collusive (i.e., there are no coordinated effects).

Finally, the proposed simulation, which involves evaluating several equations sequentially, is much simpler than a price simulation, which is an equilibrium or fixed–point calculation. Furthermore, our assumptions are less restrictive and our data requirements are less stringent. A downside is that, although our model of efficiencies is structural, our model of markups is not.

6 Data and Preliminary Analysis

6.1 The Data

The data that we use is an unbalanced panel of North American — US and Canadian — brewers between 1950 and 2012, with much of the information coming from Compustat. A firm is included in the Compustat North American data base if its shares trade on a US stock exchange. Some firms in the Compustat data are headquartered outside of North America, and those firms were eliminated because they do not produce in North America and therefore face different factor prices and economic conditions.[24] Firms with fewer than three consecutive time series observations were also eliminated. This process yielded 30 firms and 602 observations.

We classify firms according to whether they are US or Canadian and whether they are mass or craft type of production. However, since all firms in the data are publicly traded, none are small microbreweries or brew pubs. For example, the Boston Beer Company, which brews Sam Adams, is quite a large firm.

Our data consist of firm level prices and quantities of inputs — labor, materials, capital, and investment — and output — barrels of beer — for the years when a firm appears in the data.[25] Much of the Compustat data is in the form of revenues and expenditures rather than prices and quantities. Fortunately, we were able to obtain firm level data on production from various sources that are documented in appendix A, and we use that data to calculate output prices as revenue divided by production. We also constructed input prices for each

[23]This would not be possible if the two largest firms were to propose a merger. However, it is unlikely that such a merger would be allowed without requiring substantial divestitures, which could be evaluated.

[24]None of the brewers with headquarters outside of North America that were dropped have a North American brewery. The products of those firms are either imported or produced under license by a North American brewer.

[25]The fact that we have prices of inputs and outputs is somewhat unusual. In fact, many researchers use revenues rather than output as the dependent variable.

firm from a large number of data sources, and we use the price and expenditure data to obtain quantities. For example, an overall materials price index is calculated as a expenditure share weighted geometric mean of raw materials and packaging price indices, where the raw materials price index is calculated as an expenditure share weighted geometric mean of the prices of malt, hops, corn, rice, wheat, and sweeteners (sugar and corn syrup), and the packaging price index is constructed as an expenditure share weighted geometric mean of the prices of bottles, cans, and cartons. Furthermore, the materials input prices and shares differ by sector and country.

When estimating the production function, we found that labor and materials were highly correlated and it was not possible to disentangle their separate effects. For this reason, we created a single variable input, V (wth price P_v) as an expenditure share weighted geometric mean of employment and materials. This procedure implies that the function of conventional inputs in (1) can be written as $F(V(L,M),K)$, where, since input shares vary over time and firms, V is a translog. When doing this we found that the expenditure shares of three small firms were substantially greater than one, and we dropped those firms. This left us with 573 observations.

Capital, K, is our fixed input and we have assumed that V, which is a combination of employment and materials, is variable. Although it seems clear that materials are variable, the situation with labor is less obvious. Verbal communications with strategy officers of brewers led us to the following conclusions: Brewers have long term labor forces that are adequate for the minimum expected demand. Seasonal and other fluctuations are handled by hiring temporary workers. For example, since the demand for beer is largest in the summer, in that season, brewers hire students who work full or part time. Furthermore, since beer is storable, very short term fluctuations can be handled through changes in inventories. We are therefore comfortable with the assumption that, at least with yearly observations, labor is a variable factor.

Finally, we determined the reason for exit — failure, merger for synergies, merger for asset stripping, or going private — for all exiting firms.

The data are described in more detail and their sources are listed in appendix A.

6.2 Preliminary Analysis

Table 1 contains summary statistics for the variables in the North American data. Although many of the variables do not have natural units, output and employment are well defined. Production averaged over firms and years is 11 million barrels per year and the average price is $74 per barrel. In addition, the average wage is $25,500 per year and average employment is nearly 5,000 people. The table also shows that there is substantial overall variation in all of the variables. In addition, although one cannot see this in the table, there is substantial cross sectional as well as time series variation in prices as well as quantities. The table also shows that approximately 20% of the observations are on Canadian firms and about 20% are on craft brewers.

Figure 1 contains graphs of labor productivity — output per person — and materials intensity — materials use per unit of output — for the the North American industry as a whole. The figure shows that labor productivity increased dramatically over the entire period. Materials intensity, in contrast, fell somewhat after 1975, probably due to the rising popularity of light beer, which contains more water.

Table 1: Summary Statistics

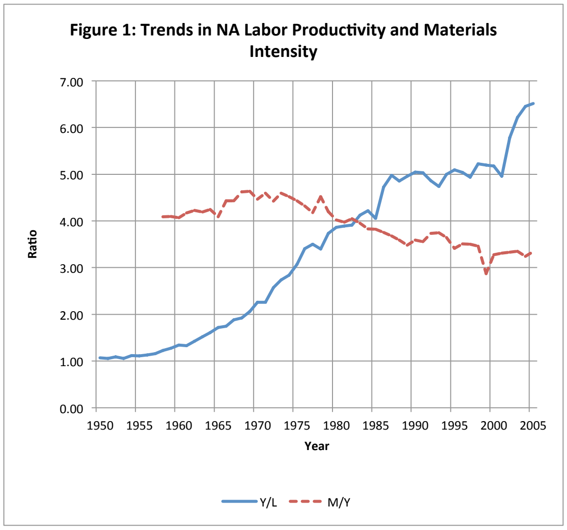
Variable	Description	Mean	St Dev	Minimum	Maximum
Y	Beer production	11.19	23.31	0.005	132.5
PY	Beer price	74.33	69.08	13.19	319.7
EMP	Employment	4.83	7.65	0.006	46.61
WAGE	Average wage	25.50	20.61	2.64	93.88
MAT	Materials use	656.3	1383.3	0.580	7933.7
PMAT	Materials price	0.717	0.474	0.275	2.14
V	Variable input	169.7	172.2	3.56	981.0
PV	Price of V	1.30	0.655	0.539	8.29
I	Investment	9.19	20.84	0.001	124.78
K	Capital stock	62.36	155.2	0.050	875.1
KPRICE	Price of capital	0.576	0.397	0.175	1.54
DC	Indicator for Canada	0.197	0.398	0	1
DS	Indicator for craft	0.190	0.393	0	1

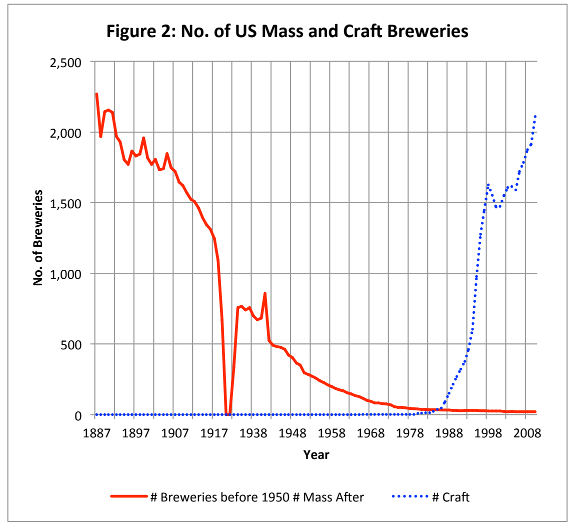
573 observations

Table 2: Summary Statistics for US Market

Variable	Description	Mean	St Dev	Minimum	Maximum
CR4	US 4 Firm Concentration Ratio	0.455	0.209	0.199	0.880
MShare	Firm's Share of US Production	0.072	0.112	0.0014	0.531
IShare	Import's Share of US Consumption	0.019	0.027	0.001	0.140

382 observations





Turning to the US market, figure 2, which contains graphs of the number of US breweries in each sector from 1887 to recent times, illustrates a bit of history. Since the distinction between mass and craft production was not meaningful in the early years, the solid line represents the total number of breweries prior to 1950 and the number of mass or conventional breweries thereafter. The figure shows that, whereas the number of breweries (in the conventional sector after 1950) fell from over 2,000 to 20, the number of craft breweries rose from zero to over 2,000. However, in spite of the large number of craft establishments, during the period of the data, the craft sector accounted for at most 10% of production. Finally, the sharp dip in breweries in the late 1920s was due to prohibition.

Table 2 contains summary statistics for the US market structure variables. The statistics are not time averages but instead are averages over the observations on the firms in US mass production that are in our data. In other words, they are moments that are used for the analysis of structure and performance in table 5. Table 2 shows that the four firm concentration ratio — the output of the four largest firms divided by industry production — varies from a low of 20 to a maximum of 88%. In addition, firm market shares, ratios of firm divided by industry production, range from extremely small to 53%. Finally import shares rose from almost nothing to 14%.[26] These trends are illustrated in figures 3–5.

Figure 3 contains a graph of the US four firm concentration ratio between 1950 and 2012. It illustrates the dramatic upward trend in concentration over the time period of the data. Moreover, figure 4, which graphs market shares of three of the largest US firms, Anheuser Busch (AB), Miller, and Coors, shows that the shares of all three have increased with time.[27] Finally, figure 5 illustrates the degree of import penetration, which grew dramatically starting in the late 1970s.

[26] Concentration ratios and market shares are fractions of US production, which includes exports, whereas import shares are fractions of US consumption, which includes imports.

[27] The dramatic increase in Miller's share in 2008 was due to the Miller Coors joint venture, when the US operations of Coors were deconsolidated from Molson Coors.

Figure 3: US 4 Firm Concentration Ratio

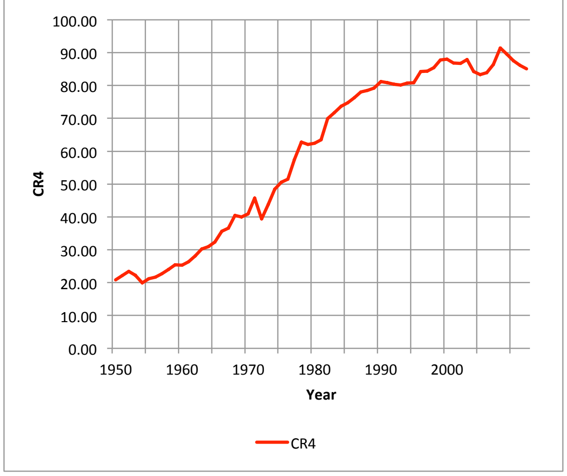


Figure 4: Market Shares of AB, Coors, and Miller

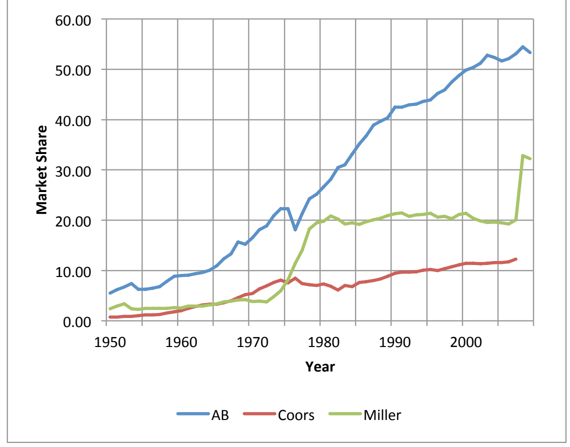
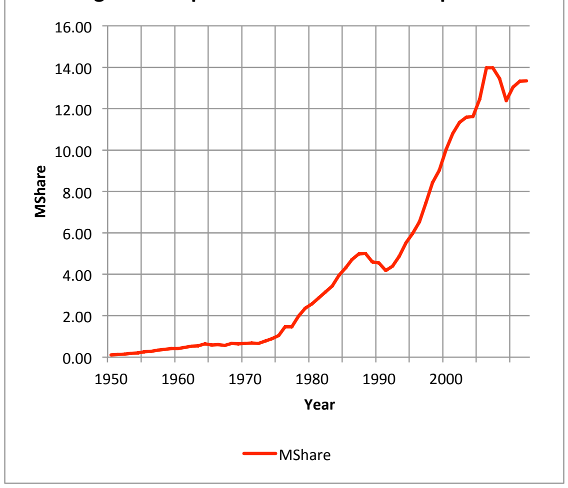


Figure 5: Imports Share of US Consumption



7 The Empirical Model

7.1 Specification of the Production Function

We must specify a functional form for f, the function of conventional inputs. We chose a translog because it is flexible[28] and because it nests the Cobb Douglas that is used by most researchers in this literature. That function is

$$f(v,k) = \beta_c \mathrm{DC} + \beta_s \mathrm{DS} + \beta_v v + \beta_k k + \beta_{vv} v^2 + \beta_{kk} k^2 + \beta_{vk} vk,$$

where all variables are in natural logarithms and DC and DS are country and sectoral dummy variables. We experimented with specifications with nonneutral technical change, i.e., with trending coefficients, $\beta_\ell =$

[28] A function is flexible if it provides a second–order approximation to an arbitrary technology. In particular, the matrix of elasticities of substitution between inputs is unconstrained.

$(\beta_{\ell 0}+\beta_{\ell t}t)$, $\ell = v,k$, but found that it was not possible to obtain robust estimates of the separate trends due to high correlation between v_t and k_t.[29]

Finally, the functions h, the inverse of the investment function that determines unobserved productivity ω, and g, the function of lagged ω that determines how ω evolves, are low order polynomials of their arguments.

7.2 Estimation

The following two equations are estimated (see section 5.2)

$$y_{jt} = \beta_0 + \beta_t t + f(v_{jt}k_{jt}) + h(k_{jt}, i_{jt}) + \eta_{jt}$$

and

$$y_{jt} = \beta_0 + \beta_t t + f(v_{jt}, k_{jt}) + g[h(k_{jt-1}, i_{jt-1})] + u_{jt},$$

where j is a firm and $u_{jt} = \omega_{jt} - E(\omega_{jt}|\omega_{jt-1}) + \eta_{jt} = \xi_{jt} + \eta_{jt}$.

Following Ackerberg, Caves, and Frazer (2006), we do not assume that the coefficients of v and v^2 are identified by the first equation alone. Instead, we estimate the two equations jointly by GMM as suggested by Wooldridge (2009). In particular, we specify different instruments for the two equations, the difference being that all variables in (**??**) and their interactions and lagged values are valid instruments for the first equation, whereas contemporaneous values of v and i and any variables formed from them must be excluded from the second instrument set. The error u_t in the second equation is correlated with v_t and i_t because it contains the current productivity shock, ξ_t. Since identification is discussed in detail in Ackerberg, Caves, and Frazer (2006) and Wooldridge (2009), we do not discuss it further here.

As noted by Wooldridge, the joint GMM estimation has several advantages over the two–step approaches that most researchers use. First, equation (**??**) contains identifying information even when it does not identify any production function parameters by itself. Second, the errors in the two equations are allowed to be contemporaneously correlated, and finally, no bootstrapping is required as fully robust standard errors are easy to obtain.

7.3 Selection

With our data on publicly traded firms, selection is particularly important. Indeed, firms exit the data when they cease to trade on a North American stock exchange. This can happen for three reasons: they merge with another firm, they go private, or they fail. However, in spite of the fact that there are three methods of exit, there are really only two outcomes: success or failure. We therefore classify firms as 'successful' if they remain in the data, if they go private, or they undergo a merger with a synergy motive. Firms are classified as

[29] We found that the performance measures obtained from the two specifications, one with trends and the other without, were very similar with one exception. The levels of technical change and thus TFP growth obtained from the model with neutral technical change were lower. However the rankings of those variables across firms were similar, and we are not interested levels per se but only in comparisons.

'failing,' in contrast, if they go bankrupt of if they undergo a merger for asset stripping purposes (delayed failure).

Like Olley and Pakes (1996), we assume that failing firms are not a random set but are instead those with low productivity. We further assume that the other three outcomes, remaining in the data, going private, and undergoing a merger for synergies are independent of productivity. For example, motives for going private include avoidance of onerous reporting to regulatory agencies and revelation of sensitive information, obtaining freedom to concentrate on long term goals, and providing incentives to management, none of which is directly related to pre–merger productivity. Furthermore, motives for a successful merger include reducing transport costs, achieving economies of scale, and reducing unit advertising costs. These goals, all of which involve post–merger changes, do not depend on the productivity of the pre–merger firms. We therefore assume that the selectivity problem only arises for failing firms and we include the estimated probability of failure in the function that determines the evolution of unobserved productivity (g).

We model exit in two stages, in the first success or failure is determined. In the second, if firms have been successful, they choose among the three successful outcomes.[30] Let the first outcome be represented by a dummy variable D_{j1t}, where 1 indicates success, and the second outcome by D_{j2t}, which equals 1 if the firm remains as an independent publicly traded firm and 0 if it either merges for synergy reasons or goes private. We adopt the OP assumption that a firm is liquidated ($D_{j1t} = 0$) when ω_{jt} drops below a threshold, $\omega_{jt} < \bar{\omega}_{jt}(k_{jt})$ for some monotonic function $\bar{\omega}_{jt}$. Hence, D_{j1t} is correlated with ω_{jt}.

Formally, Let Ω_{jt} be the information that is available in period t[31] and define

$$\chi_{jt} = D_{j1t} D_{j2t}.$$

In other words, $\chi_{jt} = 1$ if the firm remains in the data. We assume that, conditional on $\Omega_{j,t-1}$ and $D_{j1t} = 1$, the decision to remain ($D_{j2t} = 1$) is uncorrelated with ω_{jt}, i.e.

$$\text{Cov}(\omega_{jt}, D_{j2t} | \Omega_{j,t-1}, D_{j1t} = 1) = 0.$$

Then,

$$\begin{aligned} E[\omega_{jt} | \Omega_{j,t-1}, \chi_{jt} = 1] = E[\omega_{jt} | \Omega_{j,t-1}, D_{j1t} = 1, D_{j2t} = 1] &= \frac{E[\omega_t D_{j2t} | \Omega_{j,t-1}, D_{j1t} = 1]}{E[D_{j2t} | \Omega_{j,t-1}, D_{j1t} = 1]} \\ &= E[\omega_{jt} | \Omega_{j,t-1}, D_{j1t} = 1] = E[\omega_{jt} | \omega_{j,t-1}, \omega_{jt} \geq \bar{\omega}_{jt}(k_{jt})] = \tilde{g}(\omega_{j,t-1}, \bar{\omega}_{jt}(k_{jt})). \end{aligned}$$

Since the probability of failure is a sufficient statistic for the outcome of this process, the solution proceeds in the usual way.[32]

[30] We assume that the shareholders make these decisions. In other words, they not only choose whether to liquidate the firm but also whether to accept a merger offer by a publicly traded or private equity firm.

[31] More precisely: Ω_{jt} is the sigma algebra generated by all random variables in all periods up to and including period t.

[32] See Yasar, Raciborski, and Poi (2008) for an example of how selection is incorporated.

8 Results

8.1 Production Functions

When estimating the GMM specifications, we found that the preferred model for the evolution of unobserved productivity, ω, is a random walk with drift, which Wooldridge (2009) calls a 'leading case.' This finding greatly simplifies the estimation because it implies that the moment conditions are linear in the parameters. We assume that the function $h(k,i)$ is quadratic in its arguments.

Table 3 contains three specifications of the translog production function. With all three, the dependent variable is output, $y = \log(Y)$. The specifications differ according to the estimation method used. The first was estimated by OLS, the second by GMM, and the third is the full model, GMM with selection.[33] All three show that, all else equal, output has been increasing over time and that, conditional on inputs, production is lower in Canada and in the craft sector.

The bottom section of table 3, which contains means and standard deviations of estimated markups (PCM), returns to scale (RTS), technical change (TECH), and TFP growth (TFPG) for each specification, shows that average markups range from 0.96 with OLS to 1.19 with the full model, whereas average returns to scale range between 1.09 and 1.20. The difference in average markups between OLS and GMM is large. However, histograms of that variable show that the distributions are similar with both skewed to the right. Only the mean markup changes, and we are not interested in levels per se. The distributions of RTS obtained from OLS and GMM, in contrast, are symmetric. However, due to differences in estimates of the nonlinear coefficients, the distribution of GMM estimates is bimodal with a large number of firms with constant or slightly increasing returns and another group of firms with strongly increasing RTS. Since the RTS distribution produced by GMM follows the firm size distribution more closely than the unimodal OLS distribution, we find the GMM RTS distribution more credible.

With all three specifications, technical change is about 0.7% per year whereas TFP growth is 1.1%. Technical change is dominated by the trend, which is what one would expect given that ω is a random walk plus drift. TFP growth is higher than technical change because output per firm has been growing and most firms are characterized by increasing returns. Finally, the standard deviations of TECH and TFPG are very large relative to their means.

8.2 Analysis of markups, RTS, Technical Change, and TFP Growth

8.2.1 The North American Market

The averages that appear at the bottom of table 3 conceal considerable variation in all of the variables over time and firms. For this reason, we dig deeper into the determinants of PCM, RTS, TECH, and TFPG. In particular, we assess systematic variation by country and sector before looking more closely at the US mass

[33] The p values for the J statistics indicate that the over–identifying restrictions for the GMM specifications are not rejected. In particular, instrument validity is not rejected.

Table 3: OLS and GMM translog Production Functions

	(1) OLS	(2) GMM	(3) GMM Selection
t	0.00734***	0.00760***	0.00760***
	(5.15)	(6.17)	(6.25)
dc	-0.0683***	-0.101***	-0.101***
	(-2.67)	(-4.13)	(-4.16)
ds	-1.238***	-1.202***	-1.203***
	(-16.88)	(-19.72)	(-19.89)
v	0.735***	0.610***	0.612***
	(11.20)	(11.21)	(11.27)
k	0.502***	0.354***	0.321***
	(10.86)	(4.02)	(3.40)
v^2	-0.0153	0.00387	0.00335
	(-1.07)	(0.28)	(0.24)
k^2	0.0446***	0.0431***	0.0493***
	(12.38)	(3.78)	(3.43)
vk	-0.0642***	-0.00448	0.00222
	(-4.24)	(-0.14)	(0.07)
Constant	-1.781***	-1.550***	-1.565***
	(-18.26)	(-13.99)	(-13.95)
p value, J stat		0.95	0.96
Observations	573	545	545
PCM Mean	0.96	1.16	1.19
	(0.51)	(0.38)	(0.39)
RTS Mean	1.09	1.17	1.20
	(0.09)	(0.16)	(0.20)
TECH Mean	0.008	0.007	0.007
	(0.12)	(0.12)	(0.12)
TFPG Mean	0.011	0.011	0.011
	(0.13)	(0.12)	(0.12)

t statistics in parentheses, Robust standard errors, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard deviations in parentheses, Bottom section
v is the log of the variable input, k is the log of the fixed input, dc = 1 for Canada, ds = 1 for craft
The J statistic is Hansen's test of the over–identifying restrictions.

market. With these regressions, we do not interpret the coefficients as causal but simply ask if there are systematic differences in performance across different groups of firms.

Table 4 looks at trends and country and sectoral differences based on the full model (column 3) in table 3. Table 4 shows that, all else equal, only scale economies have increased with time. It also shows that markups are higher in the craft sector, whereas scale economies are lower in that sector and higher in Canada. In contrast to PCM and RTS, there are neither trends nor country or sectoral differences in technical change and TFP growth. These findings are not surprising however, since those variables are rates of change, not levels.

Finally, although not the focus of this paper, it is also true that the level of unobserved productivity, ω, is higher in the US and in the conventional or mass production sector.

Table 4: Performance by Country and Sector

	(1) PCM	(2) RTS	(3) TECH	(4) TFPG
t	-0.000503	0.00664***	-0.000549	-0.000633
	(-0.59)	(12.43)	(-1.33)	(-1.42)
dc	-0.0371	0.0381***	-0.00813	-0.00766
	(-1.03)	(3.36)	(-0.71)	(-0.68)
ds	0.341***	-0.454***	0.0310	0.0225
	(4.10)	(-21.69)	(1.44)	(1.06)
Constant	1.144***	1.089***	0.0187*	0.0271**
	(54.19)	(92.08)	(1.69)	(2.27)
Observations	573	573	545	545

t statistics in parentheses
Robust standard errors
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
dc = 1 for Canada, ds = 1 for craft

8.2.2 US Mass Production

Digging still deeper, we take a closer look at US mass production where we assess market structure effects. We chose to focus on this sector because the data is more complete and because the sector is larger in terms of production, revenues, number of firms, and observations in the data. Moreover, it is an important sector. For example, in 2012, conventional brewers in the US produced over 165 million barrels of beer and earned net revenue of almost $25 billion.

We are interested in whether the temporal increases in the market structure variables that we document in subsection 6.2 have been associated with changes in the performance measures that appear at the bottom of table 3. This exercise, however, is not straight forward. Indeed, much has been written about the endogeneity problem that is encountered when trying to uncover the causal effect of, for example, industry concentration and firm size on profitability.[34] To illustrate, high profits due to high concentration invite entry, which lowers concentration. Furthermore, when a firm's market share increases due to a fall in costs, its markups increase, which causes the firm to expand and further increase its market share. In other words, there is feedback between market structure and firm and industry performance.

We, however, do not attempt to uncover causal relationships. Our analysis is descriptive and is intended to determine if systematic differences in performance are associated with different market structures. In particular, as a check on our estimates, we wish to see if the correlations between market structure measures and our estimated performance indicators are intuitively plausible. We use industry concentration and firm market share as measures of market structure because those variables have been the focus of structure/conduct/performance studies. However, because import penetration has been increasing, we condition our estimates on that variable. Finally, we also include constants and trends in our regressions.

The top half of table 5 shows means and standard deviations of the performance measures for the US mass market. Compared to the entire sample (the bottom of table 3), markups are somewhat lower and returns to scale are slightly higher. The middle part of the table contains OLS regressions of the performance measures on the market structure variables. The table shows that market concentration is not associated with higher markups or returns to scale, which is interesting because, if one simply regresses markups and returns to scale on concentration (not shown), both relationships are positive and highly significant. However, when one conditions on the other explanatory variables, the relationships disappear. On the other hand, higher concentration does appear to be associated with lower technical change and TFP growth. Turning to the firm's market share, that variable is strongly correlated with both markups and scale economies and weakly associated with TFP growth.

The regressions in the middle of table 5 are informative about the statistical significance of coefficients. However, the coefficients, which depend on the units of the explanatory variables, are not measures of magnitudes, and we are interested in economic importance as well as statistical significance. In other words, we would like to know which effects are large. To investigate this issue, we define a notion of importance.

We say that an effect is very important (important) if a one standard deviation change in the explanatory variable is associated with at least a 10% (5%) change in the dependent variable. Specifically, we calculate the change in a performance measure that is associated with a one standard deviation change in an explanatory variable, and we divide that change by the mean of the performance measure to obtain a percentage if the measure is PCM or RTS. However, if the measure is TECH or TFPG, we do not divide by the mean, since those measures are already percentages. Although the choice of 10% and 5% is somewhat arbitrary, other critical values can be chosen.

The bottom part of table 5 contains the results of the importance calculations. The table shows that the positive effect of concentration on markups, though not statistically significant, is important. In particular, a one standard deviation increase in the four firm concentration ratio is associated with an overall increase in markups of 6%. An increase in concentration is also associated with a fall in the rate of technical change of 6%. Turning to the market share effects, they are fairly large; a one standard deviation increase in a firm's

[34] For a discussion of this literature, see Slade (2004a).

Table 5: US Market Structure and Performance

	(1) lnPCM	(2) lnRTS	(3) TECH	(4) TFPG
Mean	1.13	1.23	0.006	0.012
SD	0.210	0.190	0.113	0.120
Significance				
*CR*4	0.275	-0.0623	-0.228*	-0.285**
	(1.53)	(-0.51)	(-1.95)	(-2.25)
MShare	1.113***	1.077***	0.0845	0.131*
	(19.13)	(13.59)	(1.22)	(1.80)
Magnitude				
*CR*4	0.058$^{\bullet}$	-0.013	-0.048	-0.061$^{\bullet}$
MShare	0.123$^{\bullet\bullet}$	0.119$^{\bullet\bullet}$	0.009	0.014
Observations	382	382	366	366

t statistics in parentheses
Robust standard errors, $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Regressions contain a constant, trend, and import's share of US consumption
Magnitudes are the effect of a one SD change in the explanatory variable
$^{\bullet}$ ($^{\bullet\bullet}$)denotes an effect that is greater than 5 (10)%
*CR*4 is the 4–firm concentration ratio
MShare is a firm's share of US production

share of production is associated with a 12% increase in it's markup and a similar increase in economies of scale. On the other hand, the associations between market share and both technical change and TFP growth are not important.

It appears that, at least with respect to markups and returns to scale, the market share effect dominates the industry concentration effect. Since the correlations with market share are large, we look more closely at those relationships. A one standard deviation increase in market share is a substantial change; it moves the average firm from 7% to 18% of production, which is an increase of approximately 160%. In contrast, the standard deviation of concentration is less than half its mean.

It should not be surprising that markups grow in tandem with returns to scale. With increasing returns, a firm must price above marginal cost just to break even, and when scale economies become stronger, increases in markups are expected.

To summarize, we find many significant and important relationships between US market structure and our performance measures. Indeed, increases in industry concentration have no beneficial consequences and are associated with lower TFP growth. In contrast, increases in a firm's market share involve a tradeoff between higher markups and lower long run marginal and average costs due to increasing returns, and the two might offset one another.

8.3 The Molson Coors Merger Simulation

Table 6 contains the results of the Molson Coors merger simulations. The merger was announced in the middle of 2004, which means that only data for 2003 and earlier would have been available for the evaluation. The first half of the table contains ex ante evaluations of efficiencies based on equation (8) using data from 2002 and 2003. The table shows that simulated efficiency gains are between 8 and 9%.

Simulated efficiency gains can be compared to realized gains. The middle part of table 6, which contains estimates based on equation (9), shows that gains were only 4% in the year of the merger, considerably less than forecast. However, by the second post merger year, realized gains are very close to simulated gains. We do not look at later years because there was another period of adjustment in 2007 due to brewery openings and closures.

The bottom portion of table 6 contains simulated markups for the pre merger years 2002 and 2003 as well as estimated markups for the post merger years 2005 and 2006. One can see that, in contrast to pre and post merger efficiencies, ex post estimated changes in markups are much higher than forecasts that use pre merger data — roughly double. The large differences between the simulated markups and those that were obtained using post merger data could be explained in at least two ways. First, the merger could have facilitated tacit collusion among the major players (coordinated effects), and second, Molson Coors's variable costs could have fallen farther than predicted.

One can investigate the first hypothesis by looking at the performance of Anheuser Busch (AB), one of the merged firm's two major competitors.[35] During the 2002–2007 period, AB's output remained on trend, it's

[35] Miller, the other major competitor, had been purchased by SAB, a South African brewer, and no longer appeared in the data. Fortunately, the merger of AB with Inbev, a Belgian brewer, occurred later — in 2009.

Table 6: Molson Coors Merger Simulation

	(1) Year	(2) $F(Z_C)$	(3) $F(Z_M)$	(4) $F(Z_C + Z_M)$	(5) Efficiency (%)
Efficiency Simulation					
	2002	50.71	16.00	72.11	8
	2003	53.62	17.32	77.54	9
	Year	Inputs	Output	Efficiency (%)	
Ex Post Reductions					
	2005	13.4	9.5	4	
	2006	14.8	6.0	9	
	Year	Change (%)	Year	Change (%)	
% Change in Markups	Simulated		Estimated		
	2002	4	2005	7	
	2003	4	2006	9	

F(Z) is the function of conventional inputs, Z
C stands for Coors
M stands for Molson
The ex ante gain is column (4) divided by the sum of columns (2) and (3) minus one
Input reduction is the the negative of the percentage change in input usage
Output reduction is the negative of the percentage change in production
The ex post efficiency gain is the difference (%)
PCM is the price cost markup
Simulated markup changes use pre merger data
Estimated markup changes use post merger data

markups actually fell, and returns to scale remained constant. This evidence is inconsistent with an increase in collusion among the major players, which would imply a reduction in output and an increase in markups.

The second hypothesis can be investigated by examining the performance of Molson Coors in greater detail. During the merger period, the merged firm reoptimized, cutting back on variable input usage, which caused v's share of revenues, r_v, to fall. Under perfect competition, this would also cause the output elasticity, γ_{yv}, to fall. However, with imperfect competition, this need not happen and, in fact, the output elasticity remained relatively constant. Since markups are given by γ_{yv}/r_v, they rose to cover the increased share of fixed costs.

It is also possible to investigate the magnitude of simulated efficiencies under the hypothesis that output remained constant and inputs fell by the estimated 8%. With these calculations, long run costs are predicted to fall by between \$300 and \$330 million, depending on whether 2002 or 2003 data are used. These numbers are substantially larger than the \$175 million in merger–related efficiencies that the merging parties forecast. However, subsequent annual reports, as well as interviews with people at Molson Coors, claim that realized efficiencies did in fact far exceed those that were forecast.

Finally, one can ask what happened to TFP growth. Our estimates of that measure are so noisy that it is not possible to obtain meaningful forecasts. Indeed, any forecasts would be almost totally determined by the years that were used in calculation. However, one can look at the systematic portion of TFP growth that is due to changes in returns to scale. Our simulated estimate of the change in this portion is 0.1%, which may seem small. However, this change would move an average firm from a 1.2 to a 1.3% annual growth rate, which is not negligible.

9 Final Remarks

We propose a very simple method for forecasting merger–related efficiencies and markups that uses only pre merger data. Our simulation model has several advantages. First, unlike standard merger price simulations that are often used in evaluating competitive harm, our simulation provides a structural model of efficiencies. Moreover, with our simulations, markups and efficiencies are jointly determined. Second, we do not rely on assumptions about the way in which markups are determined. In other words, we do not need to specify the game that the firms are playing. If the game is known, imposing equilibrium restrictions provides added efficiency. However, if the game is misspecified, those restrictions lead to inconsistent estimates of markups. Third, unlike equilibrium simulation models, we do not require data on all of the major players. Since competition agencies cannot subpoena data from firms that are not involved in a merger, this can be important. Moreover, it is not necessary to define the market in order to simulate the merger. Finally, our calculations are simpler than equilibrium calculations.

There are also drawbacks associated with our proposal. The main drawback is that, absent a model of demand, our simulation cannot be used for welfare calculations. However, if one had data on all of the major players, the sort of data that is required for a merger price simulation, one could estimate demand by firm and combine it with our model of technology.[36] Second, like equilibrium price simulations, we cannot forecast merger–induced changes in collusive behavior or product positioning. Third, unlike equilibrium simulations, we cannot forecast changes in the markups of nonmerging firms. Nevertheless, we see our simulation as primarily a method of forecasting efficiencies, and markup forecasts are secondary. Clearly there are complementarities between the two sorts of simulations.

Whereas merger price simulations are based on models of demand and substitution, our simulation is based on supply and technology. We use panel data on firms in the North American brewing industry to estimate a very flexible model of a firm's technology that admits U–shaped short and long–run average cost curves, variable returns to scale, and systematic country and sectoral differences. Moreover, not only do our firms differ with respect to country, sector, and scale economies, but also with respect to market power.

Our model of technology is based on the production function/productivity model that was pioneered by Olley and Pakes (1996) and is summarized in Ackerberg, Benkard, Berry, and Pakes (2007). However, it is common in that literature to assume that firms within a narrowly defined industry differ only with respect to a persistent productivity shock. Moreover, under that assumption the growth in total factor productivity is equated with technical change (i.e., a shift in the production function holding inputs constant). Under our assumptions, however, the two are not the same. We therefore draw upon the index–number/productivity literature to derive a theoretically consistent measure of TFP growth. In particular, we purge our measure of changes that are due solely to differences in markups and we modify our measure to incorporate differences in economies of scale.

The standard simulation model defines a merger as a consolidation of control rights — the ability to set prices. We in contrast model a merger as a consolidation of physical assets, both variable and fixed. We can then equate static efficiencies with changes in production and input usage. In particular, our scale based measure picks up more efficient use of variable and fixed inputs that is due to economies of density, procurement efficiencies, savings from specialization, and the elimination of duplication as the firm expands.

Turning to market power, we equate that concept with the price cost markup, which is related to the monopolistic distortion. For example, in a linear demand and cost model, the distortion depends only on the deviation of price from marginal cost and the slope of the demand function. Finally, we equate dynamic efficiencies with changes in TFP growth rates.

When we apply our proposed simulation method to evaluate the merger between Molson and Coors that occurred in early 2005, we find that our simulations yield fairly accurate forecasts of efficiencies. However, our forecasts of post merger markups are too low. Although in theory, higher markups could be due to facilitation of tacit collusion, we argue that this is not the case here. Instead, it is due to post merger reoptimization of variable inputs.

It is not too surprising that our forecasts of changes in markups are less accurate than our efficiencies forecasts. Indeed, our predictions of scale economies come from a structural model, whereas our markup predictions simply use estimated elasticities and variable expenditure shares to back out markups. Nevertheless, equilibrium merger price simulations do not perform any better (see, e.g., Peters, 2006; Weinberg and Hosken,

[36] SABMiller and Labatt are not in our data because, by the time of the merger, they had been acquired by foreign firms.

2013). Given this fact, one might question the wisdom of using quantitative methods to infer merger–related changes in prices and markups. However, if one side in a contested merger presents econometric evidence to demonstrate impact, or lack thereof, it is wise for the other to respond in kind. Otherwise, it is difficult to argue that more realistic assumptions are not just technical niceties but instead can alter conclusions in important ways.

When we attempt to assess dynamic efficiencies, which we measure as a change in TFP growth, we find that that performance measure is much noisier and thus more difficult to predict. Nevertheless, we evaluate the systematic portion of the merger–related change in TFP growth and find that it is positive and small but not negligible.

We also use our estimated production function to evaluate trends and country and sectoral differences in performance. When we do this, we find that returns to scale have been increasing over time, as expected. Moreover, they are greater in Canada compared to the US, and are smaller in the craft sector compared to mass production. In contrast, we find that markups are higher in the craft sector. Finally, we find no systematic differences in technical change and TFP growth across countries and sectors.

We then narrow our focus to US mass production, where we evaluate firm and industry performance. Specifically, we assess how performance varies with measures of market structure and we find that, compared to industry concentration, a firm's share of the market is a much more important determinant of markups and returns to scale. With hindsight, this finding is not surprising. Indeed, small firms in concentrated industries have little market power and do not benefit from economies of scale. Nevertheless, as summarized in Schmalensee (1986), the relative importance of industry concentration and firm market share was hotly debated. However, much of the empirical work in that literature at that time assessed cross sectional differences in concentration across industries whereas we have a panel of firms within a well defined industry. Schmalensee found large cross sectional differences in profitability (as measured by industry fixed effects) and small but significant market share effects.[37] We, in contrast, find that time series variation in concentration within this industry is less important than a firm's share of production.

More recently, Ashenfelter, Hosken, and Weinberg (2015) used a reduced–form analysis to assess the Miller Coors joint venture that occurred in 2008 and found that small but significant increases in both prices and efficiencies (2–3%) post joint venture roughly offset one other. The simulated markup and efficiency changes that we predict also work in opposite directions. However, they are larger (4–9%).

Miller and Weinberg (2015) also assessed the Miller Coors joint venture and found that markups, averaged across Miller, Coors, and ABI, changed from 3.5 to 5%, which is nearly double. However, they estimated a structural model with a conduct parameter and found evidence of a merger induced rise in tacit collusion. In other words, they found that the merger changed the market game and, post merger, they reject the assumption of Bertrand competition that underlies most merger simulation models. Using a descriptive analysis, we find no evidence of increased tacit collusion associated with the Molson Coors merger. However, that merger was very different from the Miller Coors joint venture. In particular, whereas Miller Coors involved firms that interacted at arm's length, Molson Coors was a cross border merger between firms that already produced

[37] Note that Schmalensee, like us, did not look for causal relationships but instead assessed the existence and importance of effects.

each others' brands under license.

As with our merger–related findings, based on our descriptive analysis we conclude that it is easier to explain cross sectional and time series variation in markups and scale economies compared to technical change and TFP growth. Nevertheless, our US mass market regressions indicate that, on average, increased industry concentration is associated with slower productivity growth. There is also an old literature on market structure, technical change, and productivity (see, e.g., the summary in Geroski, 1994), and the arguments run in both directions. For example, some researchers argue that rapid technical change will result in increased minimum efficient scales and higher rates of concentration, whereas others note that increased rivalry (lack of concentration) will stimulate technical change. Our evidence supports the latter, at least in this industry.

References

ACKERBERG, D., L. BENKARD, S. BERRY, AND A. PAKES (2007): "Econometric Tools for Analyzing Market Outcomes," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6. North Holland, Amsterdam.

ACKERBERG, D., K. CAVES, AND G. FRAZER (2006): "Structural Identification of Production Functions," working paper.

ARELLANO, M., AND S. BOND (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58(2), 277–297.

ARELLANO, M., AND O. BOVER (1995): "Another Look at the Instrumental Variable Estimation of Error–Components Models," *Journal of Econometrics*, 68(1), 29–51.

ASHENFELTER, O. C., D. S. HOSKEN, AND M. C. WEINBERG (2015): "Efficiencies Brewed: Pricing and Consolidation in the US Beer Industry," *The RAND Journal of Economics*, 46(2), 328–361.

BAJARI, P., C. BENKARD, AND J. LEVIN (2007): "Estimating Dynamic Models of of Imperfect Competition," *Econometrica*, 75(5), 1331–1370.

BALAT, J., I. BRAMBILLA, AND Y. SASAKI (2016): "Heterogeneous Firms: Skilled Labor Productivity and Export Destimations," Johns Hopkins University.

BERRY, S., AND P. A. HAILE (2014): "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, 82(5), 1749–1797.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 847–890.

BLAIR, R. D., AND J. S. HAYNES (2011): "The Efficiencies Defense in the 2010 Horizontal Merger Guidelines," *Review of Industrial Organization*, 39, 29357–68.

BLUNDELL, R., AND S. BOND (2000): "GMM Estimation with Persistent Panel Data: an Application to Production Functions," *Econometric Reviews*, 19(3), 321–340.

CAVES, D., L. CHRISTENSEN, AND J. SWANSON (1981): "Productivity Growth, Scale Economies, and Capacity Utilization in U.S. Railroads, 1995–74," *American Economic Review*, 71, 994–1002.

CORTS, K. S. (1999): "Conduct Parameters and the Measurement of Market Power," *Journal of Econometrics*, 88, 227–250.

CRANE, D. A. (2014): "Rethinking Merger Efficiencies," *Michigan Law Review*, 110(3), 347–391.

DE LOECKER, J., AND F. WARZYNSKI (2012): "Markups and Firm Level Export Status," *American Economic Review*, 102(6), 2437–2471.

DENNY, M., M. FUSS, AND L. WAVERMAN (1981): "The Measurement and Interpretation of Total Factor Productivity in Regulated Industries, with Application to Canadian Telecommunications," in *Productivity Measurement in Regulated Industries*, ed. by T. Cowing, and R. Stevenson, pp. 179–218. Academic Press, New York.

DIEWERT, W. E., AND A. O. NAKAMURA (2007): "The Measurement of Productivity for Nations," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6, pp. 4501–4586. North Holland, Amsterdam.

ELZINGA, K. G., AND A. W. SWISHER (2005): "The Supreme Court and Beer Mergers: From Pabst/Blatz to the DOJ Merger Guidelines," *Review of Industrial Organization*, 25(3), 245–267.

GANDHI, A., S. NAVARRO, AND D. RIVERS (2013): "On the Identification of Production Functions: How Heterogenous is Productivity?," University of Wisconsin mimeo.

GEROSKI, P. (1994): *Market Structure, Corporate Performance, and Innovative Activity*. Oxford University Press, Oxford.

HALL, R. E. (1988): "The Relationship Between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96(5), 921–947.

HODERLEIN, S., J. FOX, V. HADAD, AND A. PETRIN (2016): "Heterogenous Production Functions, Panel Data, and Productivity Growth," Boston College.

IVALDI, M., AND F. VERBOVEN (2005): "Quantifying the Effects from Horizontal Mergers in European Competition Policy," *International Journal of Industrial Organization*, 23(9–10), 669–691.

JEZIORSKI, P. (2014a): "Effects of Mergers in Two–Sided Markets: The US Radio Industry," *American Economic Journal: Microeconomics*, 6(4), 35–73.

——— (2014b): "Estimation of Cost Efficiencies from Mergers: Application to US Radio," *RAND Journal of Economics*, 45(4), 816–846.

JORGENSON, D., AND Z. GRILICHES (1967): “An Explanation of Productivity Change,” *The Review of Economic Studies*, 34(3), 249–283.

KASAHARA, H., P. SCHRIMPF, AND M. SUZUKI (2015): “Identification and Estimation of Production Functions with Unobserved Heterogeneity,” UBC mimeo.

KERKVLIET, J. R., W. NEBESKY, C. H. TREMBLAY, AND V. J. TREMBLAY (1998): “Efficiency and Technical Change in the U.S. Brewing Industry,” *Journal of Productivity Analysis*, 10(3), 271–288.

KOLASKY, W., AND A. DICK (2003): “The Merger of Guidelines and the Integration of Efficiencies into Antitrust Review of Horizontal Mergers,” Wilmer Cutler Pickering Hale and Dorr Antitrust Series, Working paper 31.

LEVINSOHN, J., AND A. PETRIN (2003): “Estimating Production Functions Using Inputs to Control for Unobservables,” *Review of Economic Studies*, 70(2), 317–341.

MAKOWSKI, L. (1987): “Are ‘Rational Conjectures’ Rational?,” *Journal of Industrial Economics*, 36(1), 35–47.

MARSCHAK, J., AND W. H. ANDREWS (1944): “Random Simultaneous Equations and the Theory of Production,” *Econometrica*, 13(1), 143–205.

MILLER, N. H., AND M. C. WEINBERG (2015): “Mergers Facilitate Tacit Collusion: Empirical Evidence from the Brewing Industry,” mimeo, Georgetown University.

NEVO, A. (2000): “Mergers with Differentiated Products: The Case of the Ready–to–Eat Cereal Industry,” *RAND Journal of Economics*, 31(3), 395–421.

——— (2001): “Measuring Market Power in the Ready–to–Eat Cereal Industry,” *Econometrica*, 69(2), 307–342.

OLLEY, G. S., AND A. PAKES (1996): “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 64(6), 1263–1297.

PETERS, C. (2006): “Evaluating the Performance of Merger Simulations: Evidence from the US Airline Industry,” *Journal of Law and Economics*, 49(2), 627–649.

PINKSE, J., AND M. E. SLADE (2004): “Mergers, Brand Competition, and the Price of a Pint,” *European Economic Review*, 48(3), 617–643.

PINKSE, J., M. E. SLADE, AND C. BRETT (2002): “Spatial Price Competition: A Semiparametric Approach,” *Econometrica*, 70(3), 1111–1155.

SCHMALENSEE, R. (1986): “Do Markets Differ Much?,” *American Economic Review*, 75(3), 341–351.

Slade, M. E. (2004a): “Competing Models of Firm Profitability,” *International Journal of Industrial Organization*, 22(3), 289–308.

——— (2004b): “Market Power and Joint Dominance in UK Brewing,” *Journal of Industrial Economics*, 48(1), 133–163.

——— (2009): “Merger Simulations of Unilateral effects: What Can We Learn from the UK Brewing Industry?,” in *Cases in European Competition Policy: The Economic Analysis*, ed. by B. Lyons. Cambridge University Press, Cambridge, UK.

Solow, R. M. (1957): “Technical Change and the Aggregate Production Function,” *Review of Economics and Statistics*, 39(3), 312–320.

Tremblay, V. J., and C. H. Tremblay (2009): *The US Brewing Industry: Data and Economic Analysis*. MIT Press, Cambridge, MA.

Weinberg, M. C., and D. Hosken (2013): “Evidence on the Accuracy of Merger Simulations,” *Review of Economics and Statistics*, 95(5), 1584–1600.

Williamson, O. E. (1968): “Economies as an Antitrust Defense: The Welfare Tradeoffs,” *American Economic Review*, 58(1), 18–36.

Wooldridge, J. M. (2009): “On Estimating Firm–Level Production Functions Using Proxy Variables to Control for Unobservables,” *Economics Letters*, 104, 112–114.

Yasar, M., R. Raciborski, and B. Poi (2008): “Production function estimation in Stata using the Olley and Pakes method,” *Stata Journal*, 8(2), 221231.

A Data Appendix

Table 7: Data Sources

United States

Abreviation	Source
AR	Company Annual Reports
ASM	Annual Survey of Manufacturers, Census Bureau
BA	Brewers Almanac, The Beer Institute
BEA	Bureau of Economic Analysis, Department of Commerce
BLS	Bureau of Labor Statistics
BMI	Beer Marketer's Insights
Compustat	S&P Capital, Compustat
HSUS	Historical Statistics of the US, Cambridge University Press
NASS	National Agricultural Statistics Service, USDA
NBER	NBER-CES Manufacturing Industry Data Base
Weinberg	The Office of Robert S. Weinberg

Canada

Abreviation	Source
AR	Company Annual Reports
ASB	Annual Statistical Bulletin, Beer Canada
ASML	Annual Survey of Manufacturing & Logging, StatCan
CANSIM	Computerized data base, Statistics Canada
HSBC	Historical Statistics, Beer Canada
StatCan	Statistics Canada

International

Abreviation	Source
IMF-IFS	International Monetary Fund, International Financial Statistics

North American Data

Data for all firms consist of firm specific prices and quantities of output, labor, materials, capital, and investment. In addition, the mode of exit is included for all firms that left the data. Data sources can be found in table 7.

Output

All of the firms are brewers and most produce only beer. Beer prices and quantities are therefore the primary output data. Firm net revenues are from Compustat (Compustat pneumonic SALES). Beer production up to 2009 for each US firm is from Weinberg,[38] and after 2009 from BMI. US firm–level beer prices are then calculated as net revenue divided by production. Canadian production data for the later years (1990s and beyond) are from company annual reports, and firm–level prices for those years are calculated as in the US. Due to a lack of production data, a Canadian industry average price was used for the earlier years. This is calculated as beer industry value of shipments from ASML divided by industry production from CANSIM. The production of Canadian firms in the early years is then calculated as firm net revenue divided by industry price.

A few firms produce in multiple markets and those markets are quite diverse (e.g., from wine and soft drinks to sports and entertainment). However, in all cases, revenue from the other segments is a small fraction of the total. Moreover, in recent years, most brewers have sold their non–beverage assets in order to concentrate on brewing. Revenue by segment was obtained from the Compustat Product Segment data base. Segment prices for US non–beer manufacturing markets are from NBER, augmented with BEA prices after 2009. Canadian non–beer manufacturing prices are from CANSIM. Finally, the relevant CPI or PPI (US or Canadian from BLS or CANSIM) was used for the non–manufacturing segments, sports, entertainment, primary energy and retail. An output price index was then constructed as a revenue share weighted geometric mean of segment prices and output was calculated as net revenue divided by that price.

Labor

The number of employees in each firm is from Compustat (EMP) and the wage is calculated as expenditures on labor, also from Compustat (StaffExp), divided by the number of employees. When firm–level labor expenditures are missing, a beer industry average wage is used. The average wage for the US is calculated as industry expenditures on labor divided by the number of employees in the industry, both from ASM. For Canada, the industry data on employment and expenditures are from ASML.

Materials

A materials price index was calculated for four segments: US and Canadian mass production and craft. Materials prices differ by segment because both expenditure shares and factor prices differ. For each segment, a raw materials price index was computed as an expenditure share weighted geometric mean of the prices of malt, hops, corn, rice, wheat, and sweeteners (sugar and corn syrup). A packaging price index was created as a expenditure share weighted geometric mean of the prices of bottles, cans, and cartons. An overall

[38]This data was given to me by Carol and Victor Tremblay.

materials price index was then obtained as a expenditure share weighted geometric mean of raw materials and packaging price indices. This two–stage process was adopted to reflect the fact that substitution is easier within compared to across groups of inputs.

Expenditure shares for the US are from BA while those for Canada are from StatCan. When share data were missing, they were extrapolated. US raw materials prices are from HSUS and NASS and Canadian raw materials prices are from CANSIM. In some cases, Canadian prices are withheld due to confidentially. When this occurred, the US price in CAD was substituted, using exchange rates from IMF–IFS. Finally, packaging prices are producer price indices from BLS.

Compustat does not report materials expenditures directly. Following Bresnahan, Brynjolfsson, and Hitt (2002), materials expenditures are calculated by subtracting labor expenditures from total expenditures, where total expenditures are calculated as net revenue minus operating income before depreciation (OIBD). The quantity of raw materials is then materials expenditure divided by materials price.

Capital and Investment

Data on investment flows and capital stocks, both in dollars, are from Compustat (CAPEX and PPET). In a few cases where investment data were missing, investment was calculated from changes in capital stocks and depreciation (DAA) data or was obtained from annual reports. It is assumed that the capital equipment used by breweries trades in international markets and that all firms face the same investment prices. NBER brewery investment prices are used and, after 2009, are augmented using data from their sources. Canadian investment prices are USD prices in CAD, using exchange rates from IMF–IFS. Finally, for those firms that produce in multiple markets, an investment price index was created as an investment share weighted geometric mean of investment prices, using investment shares calculated from Compustat product segments data and NBER segment investment prices. Real capital and investment are constructed as dollar values divided by investment prices.

Currency

The currency for each observation is that reported in Compustat. This means that, with the exception of Carling O'Keefe who report values in USD in the early years, data for US firms are in USD and for Canadian firms are in CAD. In practice, as long as all values for a given observation are in the same currency, the currency is irrelevant. This is true because, for example, when a quantity variable is created by dividing expenditures by price, expenditures and price are in the same currency.

Exit

Exit from the data base occurs for three reasons: failure (a liquidation), purchase by another firm (a merger), or private purchase (a management or private–equity buyout). In addition, mergers can have two motives, synergies or asset stripping. An internet and literature search was used to determine the type and reason for each exit in the data.

US Data

Additional data for firms in the US mass production sector consists of industry concentration ratios, firms' shares of production, and import shares of consumption. Concentration ratios, the output of the four largest firms divided by industry production, and firms' shares of production are constructed from the Weinberg data on output by firm and data from BA on industry production. Import shares are constructed from data on imports and consumption from BA.