

Brewed in North America: Mergers, Marginal Costs, and Efficiency

Paul Grieco* Joris Pinkse* Margaret Slade[†]
Penn State UBC

December 2016

Abstract:

We show how production and pricing data can be used to estimate merger-related efficiencies using pre merger data. We assess the changes in static and dynamic firm efficiency and marginal costs that are expected to occur post merger. To do this, we jointly estimate firm level returns to scale, technical change, TFP growth, and price-cost markups. We implement our empirical model using data for the North American (US and Canadian) brewing industry, and we use the estimated model of firm technology to evaluate the changes that are associated with the merger between Molson and Coors. We forecast nontrivial increases in returns to scale and declines in marginal costs and we verify those results by analyzing the impact of the merger retrospectively using post merger data.

Keywords: Mergers, Efficiencies, Marginal costs, 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 the impact of a proposed merger on market power and welfare. Their analysis usually proceeds in two stages: first they predict changes in prices holding marginal costs constant, and second, if competitive concerns are raised, offsetting efficiencies are evaluated by the merging parties. This two stage process is incomplete, since in reality prices and costs are jointly determined. Moreover, if marginal costs are not constant, firm costs can change for two reasons: first, changes can be due to potential “synergies” realized by merging the two firms (e.g., returns to scale) and second, they can be due to changes in the optimal level of output. However, most merger simulations ignore potential changes in the marginal cost of production.¹

We propose a quantitative technique that can be used to forecast merger-related changes in returns to scale, marginal costs, and total factor productivity (TFP) growth and can be implemented with only pre merger data. Whereas the standard simulation model defines a merger as a consolidation of control rights — the ability to set prices — and uses estimates of demand to predict the impact of a merger on prices, we consider a merger to be a consolidation of assets and use estimates of the firms’ technology to predict the impact of a merger on costs. We view this approach as an important first step to incorporating cost changes directly into merger analysis.

We use our model of production to evaluate the cost impacts of a merger between Molson and Coors that occurred in 2005 and united the second largest brewer in Canada with the third largest in the US. 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 production, marketing, and distribution, since the brands of both firms were already sold in both countries. Our goal is to quantify these effects. To do this, we perform ex ante forecasts using pre merger data. The advantage of using a consummated merger for this exercise is that we are able to verify our forecasts using data from subsequent periods.

Our model of technology is a standard production function in the spirit of [Olley and Pakes \(1996\)](#) and [Akerberg, Caves, and Frazer \(2015\)](#). However, since we are interested in estimating returns to scale, we specify a more flexible functional form than the commonly used Cobb–Douglas. This is important for analysis of mergers since a merger implies a substantial change in the size of firms. Moreover, since our estimates of returns to scale vary across firms and time, technical change and TFP growth are not the same.² We show how they differ and how each can be estimated.

Given data on revenues and expenditures, which are available from most standard firm level data sources, we can also estimate markups. In doing this, we 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. As a prelude to a more detailed merger analysis, we

¹An important exception to this is [Fan \(2013\)](#), who models cost reductions due to consolidation of printing in mergers of local newspapers. However, this is a very specialized application and does not make use of data on production inputs.

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

compare efficiency and markups over time and across sectors and countries.

Unfortunately, just as data on demand alone cannot be used to forecast changes in costs, data on supply alone cannot be used to forecast changes in prices. As a first cut, however, we use our structural estimates of the merged firm's marginal costs and a statistical model of margins (price minus marginal costs) to predict post merger average price and markup. Specifically, we estimate margins as a nonparametric function of the firm's state variables. This forecast, which is clearly an imperfect solution, is only a first step. If one is interested in how prices and markups of different brands change in different geographic markets, one can combine our marginal cost estimates with a full demand model to conduct a merger price simulation.

Our approach unifies and extends three literatures: the use of discrete Divisia indices to measure TFP growth that began with [Jorgenson and Griliches \(1967\)](#),³ the estimation of productivity in the presence of endogenous variable factors and selection bias that dates to [Olley and Pakes \(1996\)](#), and the calculation of markups without specifying how firms compete in product markets that was initiated by [Hall \(1988\)](#). Each of those ideas has been extended in many directions by numerous researchers.⁴ However, by combining the three strands, we think that we provide a unifying supply-side approach to efficiency and market power measurement that can be complementary to the traditional demand-focused approach to merger analysis.

The heart of our approach is the estimation of a flexible firm-level production function. To estimate the production function, 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 returns to scale and markups across those groups. Moreover, even within sectors and countries, there could be substantial cross sectional variation in efficiency and market power. We therefore estimate a translog production function that allows returns to scale and markups to vary across firms and time, and we use the estimated function to forecast efficiency changes.

We chose to use firm rather than plant level data for two reasons. First, production and packaging costs account for only slightly over half of the costs of a large modern brewer. The remainder are attributable to advertising, marketing, sales, transportation, and administration, some of which are performed at the head office and others in different geographical markets. Second, if the parties to the merger were minimizing production costs at each plant prior to the merger, much of the merger related cost reduction will come from outside the plant.

The use of standard firm level 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 especially important. Indeed, 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

³For example, the Tornqvist index is a discrete Divisia index.

⁴[De Loecker and Warzynski \(2012\)](#) combine the second and third strands.

that distinguishes among those possibilities.

We assume that, at least from the production point of view, beer is relatively homogeneous. Although brands differ with respect to alcohol strength and flavor additives, those differences are small when compared to total production and marketing costs. In contrast, consumers have strong preferences over brands. This means that, at least to a first approximation, the product is homogeneous on the supply side but differentiated on the demand side. Our approach to efficiencies measurement is best suited for products that are relatively homogeneous on the supply side. However, there are many other industries, such as gasoline, soft drinks, and dairy products like yogurt, in which products are relatively homogeneous in supply but highly differentiated in demand.

Once the model is estimated we are able to compute the relevant performance measures for each firm. We then simulate the impacts of a merger by analyzing the implications for these variables of combining the merging firms' fixed and variable inputs. We allow for multiple models of "synergies" between firms based on the merging firms' recovered technical efficiency. Finally, we verify these forecasts by comparing them to the estimated performance measures recovered from post-merger data.

To anticipate, our structural simulation of the Molson Coors merger forecasts that scale economies will become stronger and marginal costs will fall. When we compare our forecasts to ex post estimates, we find that, by the second year post merger, forecast efficiencies had been largely realized. Furthermore, our statistical model of margins forecasts that prices will fall as a consequence, and prices did indeed fall in years following the merger. These findings should be contrasted with predictions from typical demand based simulations with constant marginal costs, either differentiated products Bertrand or homogeneous product Cournot, where prices can only rise post merger. The sole question is by how much.

The next section, which describes the North American brewing industry, the Molson Coors merger, and the data, 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, and finally, our empirical findings and conclusions.

2 The North American Brewing Industry

2.1 Background and History

Brewing consists of combining hops, 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.

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.

Historically the industry was relatively unconcentrated. In the US, antitrust vigilance meant that the large national brewers grew internally. However, early mergers between small brewers led to the creation of regional brewers. Those mergers, which occurred in a shakeout that allowed the remaining firms to achieve economies of scale, 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). Mergers for asset stripping, which were common during the wars of attrition in the 1970s, allowed failing firms to exit gradually.

In contrast to the US, the national Canadian brewers expanded through mergers with smaller brewers. The Canadian situation was 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.

This trend towards increasing concentration has continued in recent years due to, for example, mergers with non North American brewers (e.g., Labatt/Interbrew in Canada and Anheuser Busch/Inbev (ABI) in the US). In addition, domestic mega mergers have further increased concentration in the industry post Molson Coors, most notably the Miller Coors joint venture that occurred in 2008.⁵ Finally, if the proposed ABI Miller merger is allowed to go forward, which is looking increasingly likely,⁶ the merged firm will supply nearly one third of the world's beer. Compared to these mega mergers, the Molson Coors merger was rather small.

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.

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 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.

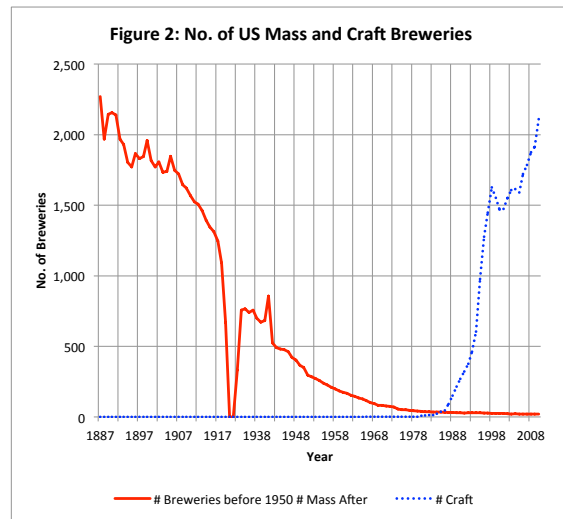
During the same time period, the share of imported beers began to rise until today it is nearly 15%. Early on, most imported beers were European. Today, however, of the five most popular brands imported into the US, four come from Mexico.

These trends in technology, sectoral growth, market structure, and imports are illustrated in figures 1 through 5. Figure 1 contains graphs of labor productivity — output per person — and materials intensity —

⁵This joint venture is studied by [Ashenfelter, Hosken, and Weinberg \(2015\)](#) and [Miller and Weinberg \(2015\)](#).

⁶In July of 2016, The US DOJ conditionally approved this \$106 billion proposal.

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.



Turning to the US market, figure 2 contains graphs of the number of US breweries in each sector from 1887 to recent times. 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 just 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.

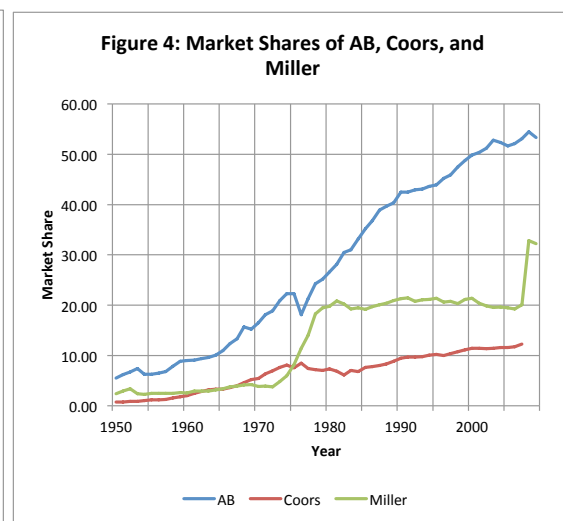
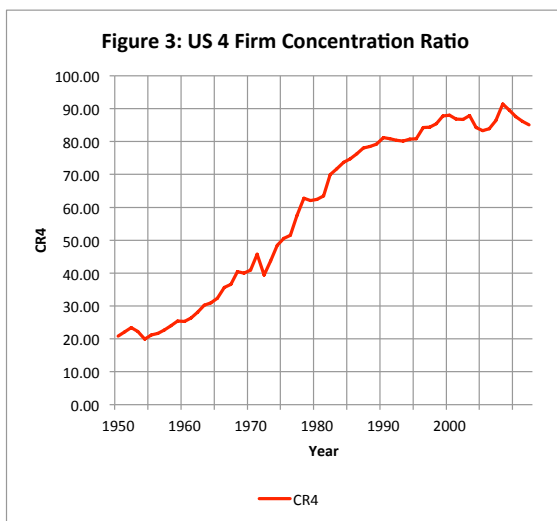
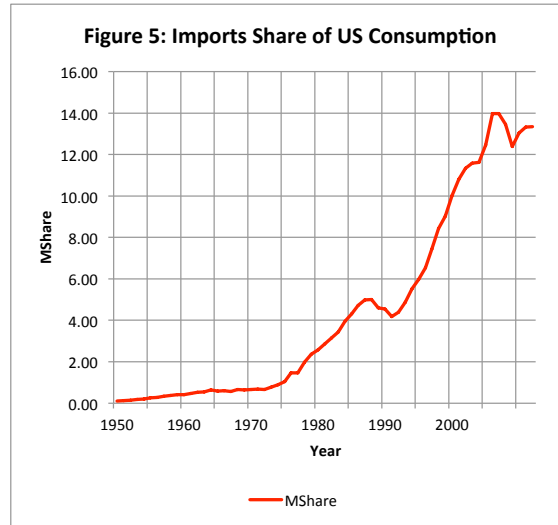


Figure 3 contains a graph of the US four firm concentration ratio between 1950 and 2012. It illustrates

the dramatic upward trend in concentration over the time period of the data. Moreover, figure 4, which graphs the 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.⁷ Finally, figure 5 illustrates the degree of import penetration, which grew dramatically starting in the late 1970s.



This brief snapshot of the industry suggests that industry concentration, returns to scale, and import penetration have increased over time. As a result, the market might have increasingly been characterized by both noncompetitive pricing and efficiencies associated with large size. Prior to our formal analysis of the merger, we assess those possibilities.

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 second largest brewer in Canada with the third largest in the US. Coors paid \$3.5 billion to acquire Molson. However, instead of consolidating the two head offices, Canadian operations, including Canadian sales of Coors brands, remain headquartered in Montreal, Que. whereas Coors operations, including US sales of Molson brands, continue to be headquartered in Golden, CO.

The Molson Coors merger was atypical among beer mergers in that rationalization of production was not a major motive. 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 for production, distribution, and marketing in the foreign country, and those entities were eliminated by the merger.

Despite the fact that headquarters would not be consolidated, Coors' 2004 Annual Report forecast 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

⁷The 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.

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.⁸ The other efficiency claims, which we assess, are more clearly tied to the merger.

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. On the other hand, there were large overlaps in markets but few overlaps in brewery locations, which are conditions that are conducive to efficiency gains.

2.3 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.⁹ 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.¹⁰ 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 craft brewer 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.¹¹ 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.¹²

We also constructed input prices for each 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

⁸Our retrospective analysis of the merger makes use of data from a period that is prior to the brewery openings and closures.

⁹The data are described in more detail and their sources are listed in appendix A.

¹⁰None 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.

¹¹The 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.

¹²Revenue is price times sales, not production. However, beer inventories are small since the period between brewing and the sell by date is quite short.

materials input prices and shares differ by sector and country.

Since we have data on total expenditures, which, for example, includes management and workers involved in advertising, R&D, and distribution, as well as total expenditure on other inputs, our efficiency measures that are based on returns to scale encompass economies of density, procurement efficiencies, savings from specialization, and the elimination of duplication as the firm expands.

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 (with 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 revenue 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 can be stored for a short period, 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.

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.

Table 2, which pertains to the US market alone, contains summary statistics for several US market structure variables. The statistics are averages over the firm-year observations on the US mass production in our data. They are moments that are used for the analysis of structure and performance in table 6. 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%.¹³

¹³Concentration ratios and market shares are fractions of US production, which includes exports, whereas import shares are fractions of US consumption, which includes imports.

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

3 Competition Policy Towards Mergers in North America

There are many similarities between the ways in which mergers are evaluated by competition authorities in the US and Canada. In both countries, pre merger notification is required for large transactions. Moreover, if a merger is challenged, there is a two step process in which the authority 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.

In the US, the Department of Justice (DOJ) or Federal Trade Commission (FTC) has 30 days after the parties have filed a pre merger notification 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. The authorities make every effort to ensure a speedy decision and most litigated mergers are decided on the basis of preliminary injunctions rather than trials.

The 2010 US Horizontal Merger Guidelines prohibit 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. Nevertheless, 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.

In contrast to the US, the Canadian Competition 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. Nevertheless, the standard of proof is higher for efficiencies.

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.¹⁴ 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 returns to scale, and scale economies are inherently at the level of the firm. Moreover, brewers produce all of their brands in just 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 national beer market is more appropriate.

¹⁴For more on asymmetric merger standards, see Crane (2014).

4 Related Literature

4.1 Merger Analysis

4.1.1 Merger Simulations

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 substitute products and engage in price competition.¹⁵ 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 j th element of the partition is the set of products that the j th firm owns. A merger is modelled 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.¹⁶ Marginal costs are not estimated using input data, but are instead derived from estimates of the demand function and the optimality conditions of some assumed model of competition (e.g., Nash-Bertrand pricing). With this assumption on firm conduct and demand estimates in hand, one can invert the N first-order conditions to recover marginal costs for each product.¹⁷ In other words, they are the costs that rationalize the assumptions on demand and equilibrium. The merger simulation then consists of altering the merged firms' objectives to reflect the fact that merged firms will maximize joint (not individual) profits and solving the resulting first order conditions for product prices while holding marginal costs for each product at their recovered level.

Static equilibrium simulations are sensitive to a number of factors including the equilibrium assumption and the specifications of demand and the assumed model of competition. Some researchers have tested their equilibrium assumptions and have not rejected them for the markets studied (e.g., [Nevo, 2001](#); [Slade, 2004b](#)). Unfortunately, however, 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

¹⁵In contrast to these price simulations, [Jeziorski \(2014a\)](#) assesses mergers where advertising is assumed to be the choice variable. However, the mechanics of the model are similar.

¹⁶A 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, and Brett \(2002\)](#); and [Ivaldi and Verboven \(2005\)](#) use a nested logit.

¹⁷Marginal costs are usually specified as constant parameters or functions of observable cost shifters in the first order conditions.

our purposes, with simulations that are commonly used by competition authorities, costs *by assumption* do not change post merger. In other words, efficiencies are not considered and price changes are therefore apt to be overestimated.

4.1.2 Retrospective Merger Analysis

Some researchers have estimated merger 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 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.

4.2 Productivity Measurement

There is a large macroeconomic literature that is devoted to estimating aggregate productivity growth and much of that 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 variable 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 \(2003\)](#), [Akerberg, Caves, and Frazer \(2015\)](#), and [Gandhi, Navarro, and Rivers \(2016\)](#) 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 that is discussed above. Since this literature is summarized in [Akerberg, Benkard, Berry, and Pakes \(2007\)](#), we do not discuss it in detail here.

Most researchers in the OP tradition follow the macroeconomic literature and assume the production function is Cobb-Douglas. However, even within narrowly defined industries, output elasticities can be heterogeneous across firms, which is ruled out in the Cobb Douglas case as long as firm heterogeneity are modeled as Hicks-neutral technology differences.¹⁸ As a result, when estimating a Cobb Douglas production function, one can equate productivity growth with technical change — a shift in the production function holding inputs constant.¹⁹ 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 therefore borrow from the index–number productivity–measurement literature to derive a measure of TFP growth that captures both technical change and changes in returns to scale. [Diewert and Nakamura \(2007\)](#) survey the index–number literature.²⁰

4.3 Markup Measurement using Production Data

In an early paper, [Hall \(1988\)](#) demonstrates 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 ,²¹ 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.

[De Loecker and Warzynski \(2012\)](#) extend Hall’s method by relaxing the assumption of constant returns to scale. Their insight is that the output elasticity of a variable input equals its share of revenue only when price equals marginal cost. Furthermore, under imperfect competition, the two are equal 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 determines the markup, P_y/MC . in addition,

¹⁸Recently, [Kasahara, Schrimpf, and Suzuki \(2015\)](#), [Zhang \(2015\)](#), [Balat, Brambilla, and Sasaki \(2016\)](#), and [Hoderlein, Fox, Hadad, and Petrin \(2016\)](#) have considered modeling non-Hicks neutral sources of unobserved heterogeneity, although this is outside the scope of our paper.

¹⁹A 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.

²⁰See also [Denny, Fuss, and Waverman \(1981\)](#) who develop an index–number model of TFP growth based on a cost 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.

²¹[Hall \(1988\)](#), like most others in the markup literature, assumes that there is a single fixed factor.

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 addressing the endogeneity problem discussed above. This extension allows them to recover firm and time-specific markups rather than industry 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.²² Indeed, from the production point of view that we adopt, the product — beer — is relatively homogeneous, at least for mass production. For example, [Tremblay and Tremblay \(2009, p.8\)](#) 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.”

Our estimated production functions, which differ by country — the US and Canada — and sector — mass production and craft — yield estimates of scale economies and technical change for each firm in each year, and those estimates are combined to yield a measure of TFP growth. In addition we obtain estimates of markups. We first perform a descriptive analysis to assess how our measures vary across Canadian and US firms and across craft and mass production. After that, we narrow our focus to US mass production, where we assess associations between various measures of market structure and our measures of performance. This preliminary analysis is undertaken to assess whether the time series and cross sectional patterns that we find are intuitively plausible. We then turn to the main exercise of the paper — our forecasts of the changes in efficiency that are expected to be associated with the merger between Molson and Coors — and we compare our forecasts to realized estimates. Finally, we use our statistical model to assess possible merger related price and markup changes.

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

5.2 The Production Function

5.2.1 Specification

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

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

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

²²With 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.

mean independent of current and past inputs. In what follows, we drop the s and c subscripts from the production and investment functions.

Equation (1) can then be written as²³

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

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

$$\omega_{jt} = g(\omega_{j,t-1}) + \xi_{jt}, \quad k_{jt} = (1 - \delta_{t-1})k_{j,t-1} + i_{j,t-1} + \varepsilon_{jt}, \quad i_{jt} = i(k_{jt}, \omega_{jt}, t). \quad (3)$$

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.

Ideally, to capture changes in market structure and other market level variables, the investment function would include year fixed effects that are interacted with the state. However, with a long time series that is not feasible. Instead, we include a trend, t and assume investment is smooth in t . Figure 3 shows that this is not a bad approximation. Indeed, market concentration increases over time in a rather smooth fashion. In addition, the investment function could include factor prices, and we experiment with that formulation.

Finally, we make a strict monotonicity assumption on the investment function — that i is monotonic in the unobservable ω — which implies that the investment function can be inverted and we can write²⁴

$$\omega_{jt} = h(k_{jt}, i_{jt}, t). \quad (4)$$

To anticipate, there are two equations that identify the parameters of the model:

$$y_{jt} = \beta_0 + \beta_t t + f(x_{jt}, k_{jt}) + h(k_{jt}, i_{jt}, t) + \eta_{jt} \quad (5)$$

and

$$y_{jt} = \beta_0 + \beta_t t + f(x_{jt}, k_{jt}) + g[h(k_{j,t-1}, i_{j,t-1}, t - 1)] + u_{jt}, \quad (6)$$

where $u_{jt} = \xi_{jt} + \eta_{jt}$. We adopt the [Akerberg, Caves, and Frazer \(2015\)](#) framework and do not use the first equation to identify any production function parameters. In practice, we estimate the two equations jointly via GMM as suggested in [Wooldridge \(2009\)](#). Details of the estimation are discussed in section 6.2. Since our dataset is unbalanced, we will also control for selection of firms due to exit. Our procedure is a straightforward extension of [Olley and Pakes \(1996\)](#) that accounts for the fact that some firms exit due to mergers and acquisition, instead of failure, we postpone the details of this procedure to Section 6.3.

²³It is not always possible to write (1) as (2). However, it is possible with the production functions that are commonly used.

²⁴Pakes (1994) derives monotonicity as an outcome of optimizing behavior.

5.2.2 Identification

Akerberg, Caves, and Frazer (2015) and Gandhi, Navarro, and Rivers (2016) have raised identification concerns about both the Olley and Pakes (1996) and Levinsohn and Petrin (2003) approaches.²⁵ Gandhi, Navarro, and Rivers (2016) propose one solution to the problem: they obtain identification by adding a variable input demand equation to the model. However, they implicitly assume that firms are price takers in output markets, an assumption that is violated in our framework.²⁶

Both of their concerns are related to the fact that standard assumptions made in this literature suggest that the model variables $v_{jt}, k_{jt}, i_{jt}, k_{j,t+1}$ live in a lower-dimensional subspace. We make slightly different assumptions. First, we assume that managers choose variable inputs v_{jt} prior to choosing investment i_{jt} , and that they are imperfectly informed of their productivity when making the variable input choice. Specifically, managers observe $\tilde{\omega}_{jt}$ when choosing variable inputs, while they observe $\omega_{jt} = E(\omega_{jt} | \tilde{\omega}_{jt}) + v_{jt}$ when choosing investment. We further assume that v_{jt} is independent of the firms' observable variables when choosing variable inputs. This means that v_{jt} is a function of $(k_{jt}, t_{jt}, \tilde{\omega}_{jt})$ and not $(k_{jt}, t_{jt}, \omega_{jt})$, which addresses the identification concern pertaining to the choice of v_{jt} .²⁷ This assumption is reasonable for the beer industry where variable input decisions are likely to be taken at much higher frequency than investment decisions, so the information set is unlikely to be identical for each choice.

We further assume that not all investments are equally successful in the sense that capital accumulates according to $k_{j,t+1} = (1 - \delta_t)k_{jt} + i_{jt} + \varepsilon_{jt}$, where the noise term is idiosyncratic. Note that with this admittedly strong assumption i_{jt} is a strictly monotonic function of ω_{jt} for any given k_{jt} , but $k_{j,t+1}$ is now a function of k_{jt}, ω_{jt} and noise. With these additional assumptions, none of the four afore-mentioned variables is a function of the others.

Lastly, both supply and demand simulations identify the effects of mergers without observing mergers. Our structural model of production relies on functional form and variation in input usage to predict output at counterfactual inputs. Demand simulations, analogously, rely on functional form and variation in product characteristics to be able to predict demand elasticities for products at counterfactual prices. In both cases, it is important to have observations on firms that are similar to the merged firm in order to avoid out of sample predictions.

5.3 The Performance Measures

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

²⁵Gandhi, Navarro, and Rivers (2016, appendix A), contains a detailed, clear, and elaborate description of the two main concerns.

²⁶Their method could be extended to the case where firms face a Dixit-Stiglitz demand curve, but this would imply uniform markups in the industry and so is not ideal for analysis of mergers.

²⁷Alternatively, we could assume that v_{jt} is not chosen optimally due to some frictions or optimization error. However, this error would contaminate our estimation of efficiencies, marginal costs and markups below. These alternatives are discussed in the working paper version of Akerberg, Caves, and Frazer (2015).

5.3.1 Returns to Scale

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

$$\text{RTS}_{jt} = \left. \frac{\lambda \partial f(\lambda Z_{jt})}{\partial \lambda} \right|_{\lambda=1} = \mathbf{t}' \frac{\partial f}{\partial z}(z_{jt}), \quad (7)$$

where \mathbf{t} is a vector of ones, the proportionate change in inputs is $\lambda - 1$ and the first partial derivative is evaluated at $\lambda = 1$. Alternatively, returns to scale is the sum of the output elasticities. With a Cobb Douglas production function, RTS is constant. When the technology is flexible, however, RTS varies with Z .

5.3.2 Technical Change

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

$$\text{TECH}_{jt} = a_{jt} - a_{jt-1} = \Delta a_{jt} = \beta_t + \omega_{jt} - \omega_{jt-1} = \beta_t + \Delta \omega_{jt}, \quad (8)$$

which consists of a deterministic and a stochastic component.

5.3.3 TFP Growth

We also 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 r_k is the k th 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, in general, be different.

[Denny, Fuss, and Waverman \(1981\)](#) show that, with a total cost function C , the growth in total factor productivity is

$$\frac{d \log \text{TFP}_C}{dt} = \frac{da}{dt} + \left(1 - \gamma_{cy}\right) \frac{dy}{dt} = \frac{da}{dt} + \left(1 - \frac{1}{\text{RTS}_C}\right) \frac{dy}{dt} \quad (9)$$

where

$$\gamma_{cy} = \frac{\partial \log C}{\partial \log Y} \quad \text{and} \quad \text{RTS}_C = \frac{1}{\gamma_{cy}}. \quad (10)$$

Furthermore, in contrast to Jorgenson and Griliches, TFP_C is based on cost, not revenue shares. With this formulation, TFP growth consists of not only technical change but also a term that depends on returns to scale. For example, under increasing returns, $\text{RTS}_C > 1$ and 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 (9) 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 (9).

Equation (9) involves infinitesimal changes and we have discrete changes. Our measure of *local* TFP growth, which is a discrete approximation to (9), is

$$TFPG_{jt} = \Delta a_{jt} + \left[1 - \frac{1}{(RTS_{jt} + RTS_{jt-1})/2} \right] \Delta y_{jt}. \quad (11)$$

In formulating (11), we do not specify either equiproportionate input changes or changes along the expansion path but instead look at actual changes.

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

5.3.4 The Price Cost Markup

Following [De Loecker and Warzynski \(2012\)](#), we *define* the markup (PMC) to be output price P_y divided by marginal cost, MC, and we use their method to measure the markup. 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 one of the variable inputs, x . Then

$$PCM_{jt} \equiv \frac{P_{yjt}}{MC_{jt}} = \gamma_{yvjt} \frac{P_{yjt} Y_{jt}}{P_{vjt} V_{jt}} = \frac{\gamma_{yvjt}}{r_{vjt}}, \quad (12)$$

where P_v is the price of v . In other words, the price cost markup equals the elasticity of output with respect to the variable input divided by the input's share of revenue, r_v . In performing this calculation, we assume that firms are price takers in input markets and are profit maximizers.

We use PCM as a measure of market power as it is a proxy for the related noncompetitive distortion. Given that we are using a firm-level output price and estimating a firm-level production function, this markup is best understood as a firm-level average. Clearly, markups can vary by location if there is either significant variation in transport costs or substantial variation in demand elasticities across geography. On the supply side, we note that if brewers are minimizing total cost, which includes not only production and packaging but also distribution and sales cost, they will equate marginal cost across breweries. Moreover, since we are interested in the firm, using the average price — a volume weighted average of the prices that the brewer sets — in our markup calculation seems appropriate.

5.3.5 Marginal Costs

Marginal costs are a key input to a merger price simulation. Moreover, the change in marginal cost is used as the key measure of efficiency gains in the calculation of upward pricing pressure of a merger (Farrell and Shapiro, 2010). Therefore, it is useful to point out that our model can recover estimates of marginal costs that do *not* rely on assumptions about competition. We use the markup equation (12) to derive an expression for marginal cost.²⁸ Dividing both sides by P_y and inverting yields,

$$\text{MC} = \frac{P_y V}{\eta_{yv} Y}. \quad (13)$$

Equation (13) shows that η_{yv} is V 's expenditure share when it is evaluated at the shadow price of output, MC, not its market price, P_y . In a competitive industry, the two are equal. Under imperfect competition, in contrast, the two diverge. We use this relationship to assess marginal cost changes from pre to post merger.

As with markups, (13) yields a single marginal cost for the firm, not one for each plant. Nevertheless, as noted above, if a firm is minimizing total cost, marginal costs will be equated across plants.

5.4 Forecasting Efficiency Gains for Mergers

We model a merger as a consolidation of assets, both variable and fixed. Our goal is to determine the immediate impact of the merger on firm efficiency. There may also be long-run effects of the merger as firms adjust their investment and input demand strategies to react to the new competitive environment, however we do not attempt to capture these as they would require a complete model of demand and investment costs.

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. Suppose further that a merger between i and j is proposed in year t . Let the inputs of firm ℓ in period t be $Z_{\ell t} = (X_{\ell t}^T, K_{\ell t}^T)^T$, $\ell = i, j$. Then, with the simplest calculation, the merged firm, m , will have the inputs Z_m equal to the sum of the inputs of the parties to the merger. We must also deal with how the merger affects technical efficiency. We consider two measures of ω_m . The first, $\bar{\omega}$ is the output share weighted average of ω_i and ω_j , which is appropriate if technical know-how is difficult to transfer across firms. The second, ω_{max} , is the maximum of ω_i and ω_j , which is appropriate if know-how is fully transferrable. We expect the actual ω to lie between the two. We do not forecast the evolution of ω since the question that we ask is what would happen if the merger were to occur today — at the time of the announcement. Formally, define $\hat{Y} = F(Z)e^{\omega} = e^{(f(Z)+\omega)}$. If i and j merge, let the inputs of the merged firm be $Z_m = Z_i + Z_j$ and ω_m be the new firm's unobserved productivity.

An obvious concern with this method is that demand effects of the merger can lead the merged firm to restrict output and therefore Z_m will be less than $Z_i + Z_j$. Ideally, our approach could be coupled with a demand estimation so that both demand side and supply side effects of a merger could be simulated jointly. That said, our approach is analogous to holding costs constant in a price simulation. If one expects that prices will rise and output will be curtailed, it is straight forward to perform a sensitivity analysis that evaluates the

²⁸In appendix B, we derive this relationship directly for a general production function.

efficiency consequences of output reductions of different magnitudes, which is similar to what Hausman, Leonard, and Zona (1994) do with a merger price simulation when they evaluate the pricing consequences of efficiency gains (marginal cost reductions) of 0, 5, and 10 percent.

This approach also assumes that brewers are price takers in input markets, which means that the merger does not affect the demand functions for variable inputs. Inputs to brewing are mostly agricultural and energy products that are sold in world markets and, in addition, are not specialized to brewing. Although the market for hops is worldwide, it is used almost exclusively in brewing. However, the cost of hops is about 1% of total brewing cost.

Finally, in order to avoid out of sample extrapolation, one should have data on firms that are larger than the merging parties, which would not be possible if, for example, the two largest firms were to propose a merger. However, it is unlikely that such a merger would be allowed without requiring substantial divestitures, and divestitures could be evaluated.

Given these assumptions, one can reevaluate equations (7), (8), and (9) using the new inputs and technical efficiency to obtain forecasts of the performance of the merged firm. For concreteness, we also calculate the overall forecast of static efficiency gains of the merger as,

$$\text{EFF}_{\text{forecast}} = \frac{\hat{Y}_m}{\hat{Y}_i + \hat{Y}_j} - 1 \quad (14)$$

where Y_m is our forecast output of the merged firm and Y_i and Y_j are pre-merger (observed) outputs.

Since the firm can reoptimize after a merger, our simulated efficiency gain should be an underestimate. On the other hand, the new firm might find it difficult to coordinate operations such as information systems, or more generally to optimize within firm organization, which would imply an overestimate.²⁹

With post merger data, we can 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,

$$\text{EFF}_{\text{realized}} = \frac{Y_m}{Y_i + Y_j} - \sum_k s_k \frac{Z_{mk}}{Z_{ik} + Z_{jk}}, \quad (15)$$

where i and j are pre, m is post merger, and s_k is the average of input k 's pre and post expenditure shares. We use this measure to verify our forecast efficiency gains.

5.5 Prices and Markups

Although not our focus, we also forecast changes in prices and markups. To do this, we specify a statistical function for price minus estimated marginal cost. Specifically, we estimate a margin equation,

$$P_y - \widehat{\text{MC}} = m(k, \omega, t, \text{DC}), \quad (16)$$

²⁹For a discussion of within firm maximization after a merger, see Michel (2016).

where m is a nonparametric function and DC is an indicator for Canada. This is clearly not an economic model. However, we simply assume that the margin that is chosen by the merged firm will be similar to those that are chosen by firms of similar size in the data. Once we have a margin forecast, we can forecast price and markup changes.

We see this as only a first cut. An economic analysis of markups can be performed by combining our marginal cost estimates with a demand based simulation model. However, since our marginal costs are not constant, the process would have to be iterative.

6 The Empirical Model

6.1 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 flexible³⁰ and because it nests the Cobb Douglas that is used by most researchers in this literature. That function is

$$f(v, k) = \beta_c DC + \beta_s DS + \beta_v v + \beta_k k + \beta_{vv} v^2 + \beta_{kk} k^2 + \beta_{vk} vk, \quad (17)$$

where all variables are in natural logarithms and DC and DS are country and sectoral dummy variables.

We experimented with many other specifications. For example, we considered nonneutral technical change, i.e., with trending coefficients, $\beta_\ell = (\beta_{\ell 0} + \beta_{\ell t} t)$, $\ell = v, k$, 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 .³¹ We also considered coefficients that varied by sector and country and nonlinear trends. However, we ran into problems of multicollinearity and over fitting and finally settled on the simple specification in equation (17). Although the specification is simple, it captures the features that we are interested in. In particular, it provides flexible estimates of RTS, TFP growth, and MC, that vary by firm, over time, and with output. Most other studies, whether production function estimation in the spirit of Olley and Pakes or merger simulation, do not incorporate this flexibility.

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. We tried including factor prices in h . However, those variables were not significant.

6.2 Estimation

The following two equations are estimated (see section 5.2)

$$y_{jt} = \beta_0 + \beta_1 t + f(v_{jt}, k_{jt}) + h(k_{jt}, i_{jt}, t) + \eta_{jt} \quad (18)$$

³⁰A 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.

³¹We 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.

and

$$y_{jt} = \beta_0 + \beta_t t + f(v_{jt}, k_{jt}) + g[h(k_{jt-1}, i_{jt-1}, t - 1)] + u_{jt}, \quad (19)$$

where j is a firm, $u_{jt} = \omega_{jt} - E(\omega_{jt} | \omega_{jt-1}) + \eta_{jt} = \xi_{jt} + \eta_{jt}$.³²

Following [Akerberg, Caves, and Frazer \(2015\)](#), 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 (18) 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_{jt} in the second equation is correlated with v_{jt} and i_{jt} because it contains the current productivity shock, ξ_{jt} .

As noted by [Wooldridge](#), the joint GMM estimation has several advantages over the two-step approaches that most researchers use. First, equation (18) 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, heteroskedastic, and serially correlated. Finally, no bootstrapping is required as analytic standard errors are easy to obtain.

6.3 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 ‘failing,’ in contrast, if they go bankrupt or if they undergo a merger for asset stripping purposes (delayed failure).

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

³²When accounting for selection, the expectation in u_{jt} must also control for the fact that we only observe productivities for those firms that remain in the data between $t - 1$ and t . This adjustment is made explicit in the [Section 6.3](#). We address selection due to exit in the following section.

successful, they choose among the three successful outcomes.³³ Let the first outcome be represented by a dummy variable D_{j1t} , where $D_{j1t} = 1$ 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 assume that a firm fails ($D_{j1t} = 0$) when ω_{jt} drops below a threshold, $\omega_{jt} < \tilde{\omega}_{jt}(k_{jt}, i_{jt}, t)$ for some function $\tilde{\omega}_{jt}$. Hence, D_{j1t} is correlated with ω_{jt} .

Formally, let Ω_{jt} be the information that is available in period t ³⁴ and define

$$\chi_{jt} = D_{j1t}D_{j2t}. \quad (20)$$

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.

$$\text{Cov}(\omega_{jt}, D_{j2t} | \Omega_{j,t-1}, D_{j1t} = 1) = 0. \quad (21)$$

Then,

$$\begin{aligned} E[\omega_{jt} | \Omega_{j,t-1}, \chi_{jt} = 1] &= E[\omega_{jt} | \Omega_{j,t-1}, D_{j1t} = 1, D_{j2t} = 1] = \frac{E[\omega_{jt} D_{j2t} | \Omega_{j,t-1}, D_{j1t} = 1]}{E[D_{j2t} | \Omega_{j,t-1}, D_{j1t} = 1]} \\ &= E[\omega_{jt} | \Omega_{j,t-1}, D_{j1t} = 1] = E[\omega_{jt} | \omega_{j,t-1}, \omega_{jt} \geq \tilde{\omega}_{jt}(k_{jt}, i_{jt}, t)] = \tilde{g}(\omega_{j,t-1}, \tilde{\omega}_{jt}(k_{jt}, i_{jt}, t)). \end{aligned} \quad (22)$$

Since the probability of failure is a sufficient statistic for the outcome of this process, the solution proceeds in the usual way.³⁵

7 Results

7.1 Production Functions

When estimating the GMM specifications, we found that the preferred model for the evolution of unobserved productivity, g , 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 inverse function h 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.³⁶ 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.

³³We 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.

³⁴More precisely: Ω_{jt} is the sigma algebra generated by all random variables in all periods up to and including period t .

³⁵See [Olley and Pakes \(1996\)](#) or [Yasar, Raciborski, and Poi \(2008\)](#) for examples of how selection is incorporated.

³⁶The 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.

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. 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. It is worth noting that the RTS distribution produced by GMM follows the firm size distribution more closely than the unimodal OLS distribution.

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 assumed to be 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, indicating that there is significant technological dispersion across firms in the brewing industry.

Table 4 contains the distributions of the estimated performance measures. It shows that, in addition to the large successful brewers, there are many small brewers that are pricing below marginal cost. Furthermore, there are many with negative technical change and TFP growth, which explains the high exit rate.

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

Once we have our performance measures, we 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.³⁷ In addition, we assess how the structure of a single market — US mass production — correlates with firm performance.

7.2.1 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 market.

Table 5 looks at differences in performance over time and by size, sector, and country based on the full model (column 3) in table 3. It shows that, all else equal, markups have fallen and 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 addition, large firms have higher markups and benefit more

³⁷This is analogous to the analysis of markups and export status in [De Loecker and Warzynski \(2012\)](#).

Table 3: OLS and GMM translog Production Functions

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

Table 4: Distributions of the Performance Measures

		PCM	RTS	TECH	TFPG
Mean		1.19	1.20	0.007	0.011
Quartiles (%)	25	0.96	1.04	-0.037	-0.035
	50	1.11	1.18	0.006	0.014
	75	1.29	1.36	0.057	0.064

Based on the GMM Specification with Selection

PCM is the price cost markup

RTS is returns to scale

TECH is technical change

TFPG is total factor productivity Growth

from increasing returns. In contrast to PCM and RTS, there are neither trends nor country, sectoral or size differences in technical change and TFP growth. These findings are not surprising however, since those variables are rates of change, not levels.

7.2.2 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 are more complete and because the sector is larger in terms of production, revenues, number of firms, and observations in the data. In 2012, conventional brewers in the US produced over 165 million barrels of beer and earned revenue net of taxes of almost \$25 billion.

We are interested in whether the temporal increases in the market structure variables that we document in Section 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.³⁸ 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 light of this, we 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.

³⁸For a discussion of this literature, see Slade (2004a).

Table 5: Performance by Country, Sector, and Size

	(1)	(2)	(3)	(4)
	PCM	RTS	TECH	TFPG
<i>t</i>	-0.00550*** (-6.25)	0.00301*** (5.59)	-0.000712 (-1.30)	-0.000931 (-1.60)
DC	0.00698 (0.20)	0.0703*** (7.16)	-0.00668 (-0.56)	-0.00499 (-0.42)
DS	0.539*** (6.36)	-0.310*** (-14.05)	0.0375 (1.44)	0.0343 (1.33)
<i>k</i>	0.000838*** (18.77)	0.000610*** (13.92)	0.0000261 (0.92)	0.0000480 (1.59)
Constant	1.189*** (57.31)	1.122*** (104.47)	0.0203* (1.73)	0.0300** (2.37)
Observations	573	573	545	545

t statistics in parentheses

DC = 1 for US, = 0 for Canada

DS = 1 for mass produced, = 0 for craft

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The top half of table 6 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 6 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

Table 6: US Market Structure and Performance

	(1)	(2)	(3)	(4)
	PCM	RTS	TECH	TFPG
Mean	1.13	1.23	0.006	0.012
SD	0.210	0.190	0.113	0.120
Significance				
<i>CR4</i>	0.275 (1.53)	-0.0623 (-0.51)	-0.228* (-1.95)	-0.285** (-2.25)
<i>MShare</i>	1.113*** (19.13)	1.077*** (13.59)	0.0845 (1.22)	0.131* (1.80)
Magnitude				
<i>CR4</i>	0.058*	-0.013	-0.048	-0.061*
<i>MShare</i>	0.123**	0.119**	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

• (••) denotes an effect that is greater than 5 (10)%

CR4 is the 4-firm concentration ratio

MShare is a firm's share of US production

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 6 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 share of production is associated with a 12% increase in its 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. Increases in industry concentration are associated with lower TFP growth when controlling for market shares. In contrast, increases in a firm's market share is associated with higher markups and lower long run marginal and average costs due to increasing returns; the two might offset one another.

7.3 The Molson Coors Merger Simulation

7.3.1 *Static Efficiency and Marginal Costs*

Table 7 contains the results of the Molson Coors merger efficiency and marginal cost 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 part of the table contains ex ante forecasts of efficiency based on equation (14) using data from 2002 and 2003 and the two measures of unobserved productivity, $\bar{\omega}$, a volume weighed average of the pre merger ω 's, and ω_{max} , the maximum of the two. The table shows that simulated efficiency gains are between 8 and 11%, depending on the specification.

Simulated efficiency gains can be compared to realized gains. The second part of table 7, which contains estimates based on equation (15), shows that gains were only 4% in the year of the merger, considerably less than forecast. However, by the second post merger year, realized are very close to simulated gains. This is quite promising given the fact that reoptimization of inputs post-merger is accounted for in the retrospective efficiency analysis whereas our forecast simply combines assets to arrive at a prediction. We do not look at later years because there was another period of adjustment in 2007 due to brewery openings and closures.

Table 7: Molson Coors Merger Efficiency Simulation and Ex Post Verification

Ex Ante Forecasts of Efficiency Gains (%)			
	Year	Using $\bar{\omega}$	Using ω_{max}
	2002	8	10
	2003	9	11
Ex Post Estimates of Efficiency Gains (%)			
	Year		
	2005	4	
	2006	9	
Ex Ante Forecasts of MC Reductions (%)			
	Year	Using $\bar{\omega}$	Using ω_{max}
	2002	10	12
	2003	8	10
Ex Post Estimates of MC Reductions (%)			
	Year		
	2005	8	
	2006	10	

Ex ante efficiency gains are based on equation (8)

Ex post efficiency gains are based on equation (9)

$\bar{\omega}$ is an output share weighted average of ω_{Coors} and ω_{Molson}

ω_{max} is the maximum of ω_{Coors} and ω_{Molson}

MC stands for marginal cost

Marginal cost reductions are relative to an output share weighted average of MC_{Coors} and MC_{Molson} pre merger

Ex ante forecasts use pre merger data

Ex post estimates use post merger data

Turning to magnitudes, it is common to measure the value of efficiencies as the merger-related reduction in costs, and indeed Coors forecast a cost reduction of \$175 million at the time of the merger. With our production function based analysis, however, we measure efficiencies as the value of the extra output that can be produced from the same level of inputs. To do this, we compare the value of pre and hypothetical post merger output holding prices and input usage constant at pre-merger levels. This is clearly an approximate gain since output price and quantity can adjust post merger. However, since those adjustments will be negatively related, the change in revenue should be less sensitive to our assumptions. This calculation indicates that the merger increases the value of output by \$80 and \$120 million, depending on whether $\bar{\omega}$, or ω_{max} is used.³⁹ The large width of this interval underscores how the potential gain depends on whether the merged firm is able to transfer know-how across its operations.

The third part of table 7 contains the marginal cost forecasts. Those forecasts, which are based on equation (13) and use pre merger data, range between an 8 and a 12 percent decline. This calculation represents the decline in marginal cost holding inputs fixed, which would be appropriate for use in a calculation of the upward pricing pressure of a merger. In a full merger analysis, one would want to account for how changes in the competitive environment induce the firm to alter output, as this would in general affect marginal costs. Nonetheless, the ex post estimates are quite close to the forecasts. This appears to support our approach as doing a reasonable job forecasting first-order changes in costs.

A fall in marginal costs of around 10% may seem large. However, inframarginal costs, as well as average variable costs, are expected to fall by smaller amounts. We highlight marginal cost as it is the relevant variable for a merger price simulation. Overall, we believe that these results are a strong indication that merger simulations that assume that marginal cost are held fixed are likely to overestimate post-merger prices.

7.3.2 *Dynamic Efficiency*

We next turn to dynamic efficiency gains, that is, changes in the growth rate of TFP. Our measure of dynamic efficiency gain after a merger is based on equation (9). However, we do not include the trend because it does not change with the merger. Our measure therefore has two parts: growth that is due to changes in returns to scale and growth that is due to changes in ω . If we use $\bar{\omega}$, which is merely a weighted average of pre merger unobserved productivities, only the first part matters. However, if we use ω_{max} , the maximum of the productivities of the merging parties, the second part contributes as well.

Using $\bar{\omega}$, our estimate of the change in TFP growth that is due solely to changes in returns to scale 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. On the other hand, if we use ω_{max} , the change is 0.3%, which is substantial. These results indicate that the dynamic benefits from the merger are largely dependent on the merged firm being able to adopt the “best practices” of its most technically proficient ancestor.

³⁹Our estimates are less than the cost reductions forecast by Coors. However, they were including cost savings due to brewery openings and closures in their estimates.

7.3.3 Markups

Due to the lack of demand data, we must replace a demand-side merger price simulation with a statistical forecast to predict markups. As described in section 5.5, we do this by specifying a functional form for the equation that relates the state variables of the firm to its margins (16). We chose a cubic of t , k , ω , and DC, the indicator for Canada, and we estimated the equation using data for the largest brewers.

Table 8, which contains ex ante forecasts and ex post estimates of changes in prices and markups, also includes marginal cost changes for comparison purposes. The table shows that, with both forecasts and estimates, prices and marginal costs fall and markups rise. If we compare our forecasts to the ex post estimates for 2006, which allow more time for the parties to adjust to the merger, we see that the ex post rise in markups (4%) is twice the forecast change. However, this fact appears to be due to lower realized marginal costs. Nevertheless, it is possible that higher markups are not solely due to efficiency — a firm effect — but are at least partially due to the change in market structure — an industry effect.

Table 8: Molson Coors Merger Markup Simulation and Ex Post Verification

Ex Ante Forecasts of Price Changes (%)		Ex Post Estimates of Price Changes (%)	
Year		Year	
2002	-8	2005	-5
2003	-6	2006	-6

Ex Ante Forecasts of MC Changes (%)		Ex Post Estimates of MC Changes (%)	
Year		Year	
2002	-10	2005	-7
2003	-8	2006	-10

Ex Ante Forecasts of PCM Changes (%)		Ex Post Estimates of PCM Changes (%)	
Year		Year	
2002	2	2005	2
2003	2	2006	4

Ex ante prices are based on a statistical model of margins

Margins are a cubic function of k , ω , t , and DC

Estimates use $\bar{\omega}$, a volume weighted average of ω_{Coors} and ω_{Molson}

MC is marginal cost

Marginal cost reductions are relative to an output share weighted average of MC_{Coors} and MC_{Molson} pre merger

PCM is price divided by marginal cost

Ex ante forecasts use pre merger data

Ex post estimates use post merger data

One can investigate this hypothesis by looking at the performance of Anheuser Busch (AB), one of the merged firm's two major competitors.⁴⁰ During the merger period, AB's output remained on trend, its markups fell by about 3%, and returns to scale remained constant. This evidence is inconsistent with an increase in collusion among the major players, or even a strong Cournot effect, both of which would imply a reduction in output and an increase in markups.

Finally, we want to reemphasize that, unlike our efficiency and marginal cost forecasts, our price and markup forecasts are not based on an economic model of price formation and should be viewed as merely a first approximation. A more detailed investigation of markups on individual brands in different geographic markets would be required for a full competitive impact analysis, and we provide supply side estimates that are complementary but typically neglected in merger analysis.

8 Final Remarks

Most analyses of mergers hold marginal cost fixed and are therefore unable to capture the impact of projected cost efficiencies. This is despite the fact that firms almost always posit that cost reductions are a key motivation for the proposed merger. We leverage production function estimation techniques to propose a very simple method for forecasting merger-related efficiencies. Our approach provides a structural model to forecast a merger's affect on returns to scale, marginal costs, and TFP growth. These forecasts can complement existing merger price simulations that evaluate competitive harm due to changes in market structure.

Our model of technology is based on the production function/productivity model that was pioneered by [Olley and Pakes \(1996\)](#) and is summarized in [Akerberg, Benkard, Berry, and Pakes \(2007\)](#). We employ a translog production function that is able to capture both changes in returns to scale and changes in technical proficiency that can result from a merger. We draw upon the index-number/productivity literature to derive a theoretically consistent measure of TFP growth. In particular, we modify the standard measure to incorporate differences in economies of scale (a movement along the production function holding technology constant).

To illustrate our approach, we apply our proposed simulation method to evaluate the merger between Molson and Coors that occurred in early 2005 using panel data from the North American brewing industry. Because we have data for years following the merger, we are able to compare our forecast to retrospective measures of efficiency changes following the merger. We find that our simulations yield fairly accurate forecasts of efficiencies. Moreover, prices are expected to decline as a consequence. Given such strong evidence, it is quite possible that a full scale investigation would not be undertaken. One can contrast these findings with predictions from merger price simulations where prices can only rise.

While we are able to provide a first-order forecast of merger efficiencies, supply side simulation alone also has several important limitations. The main drawback is that, absent a model of demand, our simulation cannot be used for welfare calculations. Furthermore, although we provide a statistical model of price and markup changes, it is not an economic model. In particular, we simply assume that the merged firm will use a price setting rule that is similar to the rules used by firms of similar size in the data. It should therefore be

⁴⁰Miller, 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.

clear that the two types of simulation models are complements and that a complete analysis of competitive harm and offsetting efficiencies should involve using both price and efficiency simulations in an iterative fashion.

Of course, our procedure will require production data—which is not typically part of a traditional merger simulation. That said, 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. Rather, we need only a panel of firms sufficient to estimate the firm’s production function.

There are also limitations that both sorts of simulations share. For example, like equilibrium price simulations, we cannot forecast merger–induced changes in collusive behavior, product positioning, or the number of brands offered. A merger is a complex transaction and one cannot hope to capture all aspects of that transaction with a simple model of supply or demand. Given that fact, one might question the wisdom of using quantitative methods to infer merger–related changes in efficiencies 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.

We conclude by highlighting how our work complements two recent papers on the impact of mergers in the brewing industry. [Ashenfelter, Hosken, and Weinberg \(2015\)](#) used a reduced–form analysis to assess the Miller Coors joint venture that occurred in 2008 and found small but significant increases in both prices and efficiencies (2–3%) post joint venture that roughly offset one other. Our efficiency estimates for Molson Coors are larger, which is perhaps due to the fact that Ashenfelter, Hosken, and Weinstein focus principally on reductions in transport costs.

[Miller and Weinberg \(2015\)](#) also assessed the Miller Coors joint venture. In a manner similar to the traditional merger simulation literature, they estimate marginal cost effects of the merger by combining a structural model of demand with a model of differentiated product price competition between firms. Using data pre- and post- merger, they also find that the Miller Coors merger entailed significant cost reductions. While the cost reduction is more relevant to our work, Miller and Weinberg’s primary focus is to document the possibility that mergers can induce tacit collusion. They find that the merger changed the market game and they reject the assumption of post merger Bertrand competition that underlies most merger price simulations. Using a descriptive analysis of markups before and after the Molson Coors merger, we find no evidence of increased tacit collusion. However, our merger was very different from the Miller Coors joint venture. In particular, whereas Miller Coors involved firms that interacted at arm’s length pre merger, Molson Coors was a cross border merger between firms that already produced each others’ brands under license.

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 (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83(6), 2411–2451.
- 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 Imperfect Competition," *Econometrica*, 75(5), 1331–1370.
- BALAT, J., I. BRAMBILLA, AND Y. SASAKI (2016): "Heterogeneous Firms: Skilled Labor Productivity and Export Destinations," 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.
- BRESNAHAN, T., E. BRYNJOLFSSON, AND L. HITT (2002): "Information Technology, Workplace Organization, and the Demand for Skilled Labor," *Quarterly Journal of Economics*, 117, 339–376.
- 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.
- DI EWERT, 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.
- FAN, Y. (2013): “Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market,” *American Economic Review*, 103(5), 1598–1628.
- FARRELL, J., AND C. SHAPIRO (2010): “Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition,” *B.E. Journal of Theoretical Economics*, 10(1), Article 9.
- GANDHI, A., S. NAVARRO, AND D. RIVERS (2016): “On the Identification of Production Functions: How Heterogenous is Productivity?,” Western University mimeo.
- HALL, R. E. (1988): “The Relationship Between Price and Marginal Cost in U.S. Industry,” *Journal of Political Economy*, 96(5), 921–947.
- HAUSMAN, J., G. LEONARD, AND J. D. ZONA (1994): “Competitive Analysis with Differentiated Products,” *Annales d’Economie et de Statistique*, 34, 159–180.
- HODERLEIN, S., J. FOX, V. HADAD, AND A. PETRIN (2016): “Heterogenous Production Functions, Panel Data, and Productivity Growth,” Boston College.
- I VALDI, 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.
- KERKVL IET, 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.
- 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.
- 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.

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), 221—231.

ZHANG, H. (2015): “Non-Neutral Technology, Firm Heterogeneity, and Labor Demand,” University of Hong Kong.

A Data Appendix

Table 9: 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 9.

Output

All of the firms are brewers and most produce only beer. Beer prices and quantities are therefore the primary output data. Firm revenues net of taxes are from Compustat (Compustat pneumatic SALES). Beer production up to 2009 for each US firm is from Weinberg,⁴¹ and after 2009 from BMI. US firm-level beer prices are then calculated as revenue net of taxes 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 (4) produced in multiple markets for a decade or so 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 (full time equivalents) 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,

⁴¹This data was given to us by Carol and Victor Tremblay.

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 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 (33 points, all for small breweries). 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 confidentiality. 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 and Deflation

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. For the same reasons, it doesn't matter if monetary variables are in constant or current dollars. However, units of measurement matter when estimating margins. For those estimations, Canadian dollars were converted to US and all were deflated by the US PPI, all items. Real USD were also used in performing the ex post merger assessment.

Measurement Error

As with most empirical studies, our data are not perfect and measurement error is a possibility. In particular, for some observations, we use an industry, rather than a firm, output price or wage. However, this problem is less severe for us than for much research in this tradition, where industry wide input and output prices are used to deflate all monetary variables (e.g., [Olley and Pakes \(1996\)](#)) or where a single overall price index is used to deflate all variables (e.g., [Levinsohn and Petrin \(2003\)](#) and [Gandhi, Navarro, and Rivers \(2016\)](#)). Nevertheless, we discuss the issue. Measurement error in y does not create a problem as long as it is not correlated with the explanatory variables. On the other hand, measurement error in an explanatory variable tends to bias the coefficients and t statistics towards zero. As can be seen by examining table 3, we do not have those symptoms and therefore do not try to correct the problem.

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.

B Derivation of the Marginal Cost Equation

Suppose that the production function is $Y = F(X, K, t, \omega)$, where X (K) is a vector of variable (fixed) inputs and let the price of X_i be P_i . The firm's cost minimizing problem is to choose variable inputs to minimize

$$\sum_i P_i X_i + \lambda [Y - F(X, K, t, \omega)], \quad (23)$$

where λ is the multiplier on the feasibility constraint. The i 'th first order condition for this problem is

$$P_i - \lambda F_i = 0 \quad \text{where} \quad F_i = \frac{\partial F}{\partial X_i}. \quad (24)$$

Multiplying both sides by the input output ratio yields

$$\frac{P_i X_i}{Y} = \lambda F_i \frac{X_i}{Y} = \lambda \eta_{yi}, \quad (25)$$

where η_{yi} is the output elasticity of the i 'th variable input. Finally, λ is the shadow price of output, which is marginal cost. Substituting yields

$$\text{MC} = \frac{P_i X_i}{\eta_{yi} Y}, \quad (26)$$

which is equation (13) in the text.