Queen's Economics Department Working Paper No. 1447

Incentivized Mergers and Cost Effciency: Evidence from the Electricity Distribution Industry

Robert Clark
Queen's University

Mario Samano
HEC Montreal

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

7-2020

Incentivized Mergers and Cost Efficiency: Evidence from the Electricity Distribution Industry*

Robert Clark† Mario Samano‡

July 10, 2020

Abstract

In an effort to lower costs of provision, authorities have encouraged the consolidation of providers for a number of services such as electricity distributors, school boards, hospitals, and municipalities. In this paper we propose an endogenous merger process to evaluate the impact of government-provided incentives on consolidation patterns, and to evaluate the resulting outcomes. The process takes as input estimates from a stochastic frontier cost model, which yields an average cost curve for the industry. Policy parameters are used to simulate final configurations using offers that are the output of a Nash Bargaining problem. The efficiency of candidate merged entities is determined by a relative-influence function that measures the degree to which the combination of the involved firms' levels of efficiency results in cost-increasing amalgamations, and an interconnection cost that measures the impact of the size of the conglomerate that is formed. We calibrate parameters by applying the merger process to replicate the observed industry reconfiguration and then use these parameters to simulate the consolidation patterns that would have resulted from different policy incentives. We apply the method to the case of Ontario, where past mergers of local electricity distribution companies were incentivized by transfer tax reductions and a further round of mergers was recently proposed. Our findings suggest that the proposed tax incentive would have no impact on efficiency levels and consolidation patterns, and that even a substantial subsidy would still leave about five times as many LDCs as desired by policy makers.

*We thank Andrew Eckert, Gautam Gowrisankaran, Ignatius Horstmann, Andrew Leach, Angel Lopez, and seminar participants at the University of Alberta, the Dartmouth Winter IO conference, the ZEW Energy Conference, the IIOC 2016, the CEA 2016, the EEA 2016, and the CRESSE 2017 for helpful comments. We thank Claudia Berube and Marianne Laurin for excellent research assistance.

†Queen's University. Email: `clarkr@econ.queensu.ca`

‡HEC Montreal. Email: `mario.samano@hec.ca`

1 Introduction

In an effort to lower costs of provision, political authorities throughout the world have encouraged the consolidation of providers for a number of services. Examples include local electricity distribution companies, school boards, hospitals, and even entire municipalities. In some cases, governments have forced the amalgamation of specific entities, but in many others they have provided incentives to encourage participants to voluntarily consolidate.

The objective of this paper is to study the impact of policies designed to encourage consolidation. More specifically, our aim is to develop a framework for computing the counterfactual set of mergers that would have arisen following the enactment of a policy designed to incentivize voluntary consolidation, and for evaluating the consequences of these amalgamations. Determining whether mergers will take place and who merges with whom involves serious methodological challenges. In particular, since any firm can merge with any other, if there is initially a large number of firms, the set of possible partnerships can be significant. Furthermore, merger decisions are interdependent in the sense that firm A's acquisition of firm C, prevents firm B from acquiring firm C.

Studies trying to tackle these issues have mostly resorted to predictions based on retrospective analysis of merger waves. Gaynor et al. [2012] look at the consequences of government-led hospital mergers in the U.K., Harman and Harman [2003] study amalgamations in higher education, Brasington [1999] analyzes school district consolidation, while Saarimaa and Tukiainen [2014] consider municipal amalgamations in Finland.[1] However, although retrospective analysis is useful for studying the consequences of specific merger waves, it cannot necessarily inform as to the impact of unconsummated amalgamations or predict whether and which mergers will occur under different policy scenarios (Einav and Levin [2010]).[2]

[1]There is also a literature on the causes and consequences of non-incentivized merger waves. See Mitchell and Mulherin [1996] and Harford [2005] for financial markets; Harrison [2006], and Park and Town [2014] for hospitals; Pesendorfer [2003] for the paper industry; Okazaki et al. [2019] for electricity distributors in pre-WWII Japan; Eliason et al. [2020] for the dialysis industry. Some papers have also focused on the effect of mergers on cost efficiency (i.e. Focarelli and Panetta [2003], Fee and Thomas [2004], Maksimovic et al. [2011], Ashenfelter et al. [2015] and Bloom et al. [2015]).

[2]There is also a literature on prospective mergers that tries to simulate the impact of individual mergers

In contrast, our approach is to appeal to the theory literature on endogenous mergers, initiated by Perry and Porter [1985], Deneckere and Davidison [1985], and Kamien and Zang [1990], and advanced by Gowrisankaran [1999] and Qiu and Zhou [2007]. We develop a sequential merger algorithm that predicts which mergers should occur and we compare these to the actual consolidation patterns observed over the same period. This allows us to pin down parameters governing the merger process, which we can then use to analyze the effect of the consolidation-incentivizing policies.

Our particular focus in this paper is on local electricity distribution companies (LDCs) in Ontario, where the provincial government is considering incentivizing significant reorganization by subsidizing consolidation. At this time, there are just over 60 LDCs in the province and the government would like to see this number fall to about 10. Mergers in this industry involving private firms have been penalized with a transfer tax of 22% (while public utilities are exempt from corporate taxes but provide a payment in lieu of taxes designed to pay down the stranded debt) and the government is considering lowering this to 0%. Our aim is to determine what sort of consolidation will occur under the proposed subsidy scheme, and whether this reorganization results in hoped-for efficiency gains. Since our estimates suggest that the proposed tax reduction is insufficient to achieve the desired level of consolidation, we also ask what other incentives would be required to achieve significant consolidation.

A benefit of studying the electricity distribution industry is that some decisions are controlled by the government, and so are simple to model. Gowrisankaran [1999], Jeziorski [2014], Igami and Uetake [2019], Mermelstein et al. [2020], and Hollenbeck [2020] embed a merger algorithm within a dynamic framework that allows also for pricing/production, entry/exit, and investment. In the case of LDC mergers, there is no price or quantity competition, since LDCs are heavily regulated and they cannot charge a price above a given cap price. Furthermore, there is no entry/exit, since a single LDC serves each area.

This allows us to focus only on the static merger decision faced by firms. When firms merge they combine their respective customer bases, their distribution lines, and their effi-

(as opposed to waves) that have been proposed. See for instance Nevo [2000] and Collard-Wexler [2014]. For a discussion of the accuracy of simulations see Peters [2006] and Hosken and Weinberg [2013].

ciency levels (their proximity to the cost frontier). Revenue gains may accrue from increasing the number of customers, but costs could rise or fall depending on whether the merged entity reaches a point of diseconomies of scale or density, and on the relative efficiency levels of the merging parties. Combining LDCs also involves interconnection costs that may be rising in the number of firms being grouped together and the distances between them. According to our algorithm, LDCs move sequentially with the active LDC deciding which firm(s) to acquire by considering the benefits from consolidation with each potential merging partner. These benefits are net of the acquisition payment, which is determined as the outcome of a Nash bargaining problem.

To operationalize our approach we must first estimate the industry average cost curve in order to forecast the cost to assign to the merged entity. We use detailed data on annual costs, revenue, total assets, sales, infrastructure, and service area characteristics from each LDC in Ontario, and a stochastic frontier analysis approach to recover a cost function that then is transformed into an average cost function. The cost assigned to the merged entity is based on its size, but also the relative efficiency levels of the two parties. Specifically, the combination of the two merging entities' levels of efficiency is done through a relative-influence function that modifies the predicted average cost of the merged entity at that given level of output and density of customers per km of line, as a function of their proximity to the cost frontier.

We use our merger algorithm and observed data to calibrate the relative influence of the firms' efficiencies on the merged entity, and the interconnection cost. More specifically, we employ a minimum distance approach similar to that in Stahl [2016] to select the parameter values yielding the industry reconfiguration that is closest to what is observed in the data. We then compare this baseline scenario, achieved under the current tax regime, to counterfactual scenarios in which the tax regime is altered.

Our approach is related to those of Gordon and Knight [2009], who look at school district consolidation in Iowa, and Weese [2015] who considers municipal amalgamations in Japan. These papers develop matching estimators that allow them to determine which factors influenced the observed consolidation, and that can be used to predict how outcomes would change in the event that policy parameters are adjusted. Relative to Gordon and Knight [2009], we allow for mergers of more than two firms, since this is consistent with what is observed in the data. Relative to Weese [2015], our algorithm uses the model to decide which merger combinations to consider rather than selecting pairs at random. However, our approach comes at a cost in that we cannot account for all possible permutations (since our potential merging partners sets are sometimes very large) and our solution cannot be considered an equilibrium in the way these other methods can. This being said, our solution is stable in the sense that no LDC wishes to acquire additional firms (in line with the approaches taken by Jeziorski [2014] and Fan and Yang [2020]).

Our analysis suggests that buyers have a slightly smaller influence (41%) on the newly merged firm's cost efficiency than do sellers. This, combined with the fact that sellers are more efficient than buyers, implies that the latter are targeting more efficient LDCs for acquisitions. Our findings also suggest that interconnection costs represent 68% of the average revenue in our sample, for two-firm conglomerates. Our merger algorithm does a good job replicating the consolidation patterns observed in the data. The actual survival ratio was 70% and we predict 67%, the observed average number of customers was 54,200 and our model predicts 47,970, and the average size of the conglomerates formed was 2.23 while we predict 2.26 firms per merged entity. As robustness checks, we calibrate our model under six additional alternative specifications that yield survival ratios between 0.59 and 0.78, well within a reasonable distance of the targeted outcome value.

In our first counterfactual we consider the proposed tax-incentive policy that reduces the transfer tax for transactions involving private entities. This has almost no impact on consolidation patterns. In light of this, we also consider counterfactual scenarios in which subsidies of varying size are offered. Even a substantial subsidy reduces the number of LDCs by only 22%, nowhere near the stated objective (a reduction of 84%). Our simulations also reveal that consolidation does not lead to the desired average cost reductions, and, in fact, can even lead to cost increases. Overall, we conclude that further consolidation in this industry would be inefficient. As the size of the subsidy increases, the number of conglomerates formed is lower, but each of them contains on average a larger number of firms. The overall average efficiency level decreases for low levels of the subsidy, but stagnates quickly as most cost efficiencies are utilized early on. This disincentivizes further consolidation.

The rest of this paper proceeds as follows. In the next section we provide background on the electricity distribution sector and describe the data. In Section 3 we explain the semi-parametric estimation of the cost function. Section 4 describes the model, while Section 5 explains the calibration procedure. Section 6 describes the empirical results, goodness of fit and the findings from our policy evaluation. Finally, Section 7 concludes.

2 Electricity Distribution and Data

2.1 Electricity distribution

Many electricity markets around the world went through a deregulation process in the 1990s, mostly characterized by the breaking up of vertically-integrated firms into production, transmission, and distribution segments. It is typical that in such cases, the only segment that is partially handled as a market is the production sector through an Independent System Operator (ISO).[3] Transmission and distribution are typically left as regulated natural monopolies given the economies of scale assumed to exist in these segments.[4]

This is the case in Ontario, where a large number of Local Distribution Companies (LDCs) exist, each serving a specific location. LDCs buy electricity from the ISO and from transmission companies to distribute to end consumers according to reliability standards and government-imposed caps on price.[5]

In a number of jurisdictions, LDCs have experienced merger waves, sometimes resulting from government policies to incentivize consolidation. In New South Wales, laws are currently

[3]Conversations between the authors and some regulators in Ontario revealed that only about 10% of wholesale electricity is sold through ISO auctions and the rest through long-term bilateral contracts.

[4]The natural monopoly nature of electricity distribution has started to change with the wider adoption of solar and wind technologies, as well as with the introduction of low-scale storage capacity devices and distributed generation. As more consumers opt for installing solar panels and batteries at home, the distribution companies will no longer serve the entire market, opening the market for a small amount of competition. For a discussion see Gowrisankaran et al. [2016] and Brown and Sappington [2017].

[5]We refer to the standards imposed by the Canada's National Energy Board (NEB) and by the North American Electric Reliability Corporation (NERC).

being approved to merge the three electricity distribution companies of the region.[6] The US Energy Policy Act resulted in roughly 23 mergers per year through the 1990s.[7]

In the case of Ontario there were over 300 municipal electric utilities in the 1990s. These were operated like departments within municipalities. Two episodes of consolidation saw the number of LDCs drop by over 200 over the next ten years. First was a series of forced acquisitions made by Hydro One, the provincially owned distribution and transmission service provider, and by the amalgamation of cities during the late 1990s and early 2000s. A second wave of mergers was realized when the government put in place a holiday in the transfer tax for the purchase of companies. Without this exception, an LDC that bought another would pay a transfer tax equivalent to 33 percent of the value, designed to help pay down the stranded debt.[8] Since 2009 the transfer tax has been eliminated for all amalgamations combining two or more publicly owned LDCs, but is still in place for those involving at least one privately-owned LDC.

Despite the important rationalization that occurred in Ontario, in the Fall of 2012 the Ontario Distribution Sector Review Panel appointed by the provincial government arrived at the conclusion that LDCs should consolidate further, into just eight to twelve major companies in order to reduce costs and incentivize investment in distribution lines. The implementation of this recommendation would impose a drastic change in the market structure of electricity distribution, since the current number of LDCs in Ontario is about seven times the Review Panel's target.[9]

As a response to the public debate regarding the appropriate number of LDCs, it has been suggested to make use of a tax holiday for amalgamations involving private companies. The current proposition is to lower the transfer tax from 22 percent to 0 for mergers involving municipal electric utilities (MEUs). There are, however, only a small number of privately

[6] http://www.dailytelegraph.com.au/ofarrell-government-to-merge-states-electricity-distributors/story-e6freuy9-1226302672417?nk=b0ce0df7e55765d1c048ace947d7caed

[7] See Joskow [2000], Kwoka [2005], Kwoka and Pollitt [2010] for discussions.

[8] As mentioned in the Introduction, the transfer tax was designed to protect the payments in lieu of taxes that Ontario requires electricity utilities, which are at least 90% municipally owned to pay, since they are exempt from corporate taxes. Sales of municipal LDCs would lower revenues for the provincial government.

[9] From the Ontario Distribution Sector Review Panel (Elston et al. [2012]).

owned LDCs and so it may be necessary to provide subsidies if further mergers between MEUs are to take place.

Since the initiative was made public, a number of LDCs and think tanks have expressed skepticism about the potential efficiency gains from these proposed amalgamations.[10] The effects of the policy recommendation in the early 2000s clearly did not produce a number of LDCs close to the current panel's target. If LDCs' decisions were far from the current target even with a tax incentive, further amalgamations may be putting the companies in situations where they are operating on the upward-sloping portion of their average costs curves. Studies on past experiences in the electricity distribution industry have found that related policies resulted in inefficient LDCs buying more efficient LDCs, and that the efficiency of purchased firms fell after the merger (see Kwoka and Pollitt [2010]).[11] A study by Becker-Blease et al. [2008] found similar results when looking at stock prices before and after amalgamations in electric utilities.

In addition to any tax incentives, the Ontario Energy Board (OEB) has, since 2015, attempted to encourage amalgamation by allowing consolidated distributors to defer rate rebasing for up to 10 years (see Ontario Energy Board [2015]). Furthermore, plans for rate harmonization of consolidated distributors are only expected at the time of rebasing (Ontario Energy Board [2016]).[12] We assume in the empirical analysis that no harmonization occurs

[10]See for instance (Fyfe et al. [2013]). The Review Panel wrote "While some stakeholders argued for mandatory consolidation, others told the Panel that they preferred voluntary consolidation. The Panel's preference is for voluntary consolidation, but action must be swift. The Panel recommends that licence applications of all new regional distributors be submitted to the OEB within two years of the government adopting the recommendations of this report." (Elston et al. [2012])

[11]They study the US electricity industry during the period 1994?2003 during which time 75 mergers took place. The authors use data envelopment analysis to evaluate each firm's production efficiency.

[12]The Ontario Energy Board uses price-cap incentive regulation to reward LDCs with higher efficiency levels. It consists of allowing a percentage increase in rates equal to the inputs inflation rate (same for all LDCs) minus the sum of the industry productivity factor expressed as a percentage (same for all LDCs) and a stretch factor expressed as a percentage. The latter includes extraordinary events that increase production costs and is firm specific. As of 2015, the input inflation factor was set at 1.6%, the productivity factor was set at 0%, and the stretch factor was allowed to range between 0 and 0.6%. Then the maximum allowed increase in rates is 2.2%. In our model, we will assume that this maximum allowed change in rates is small enough to require a mapping from our estimated inefficiency levels into changes in the price cap. For a discussion of regulation of the distribution sector see Lim and Yurukoglu [2016]. See also Dimitropoulo and Yatchew [2017] for an analysis of the effect of the price-cap incentive regulation on productivity growth in Ontario's electricity distribution sector.

and that customers continue to pay the price cap associated with their initial LDC even after a merger.

2.2 Data

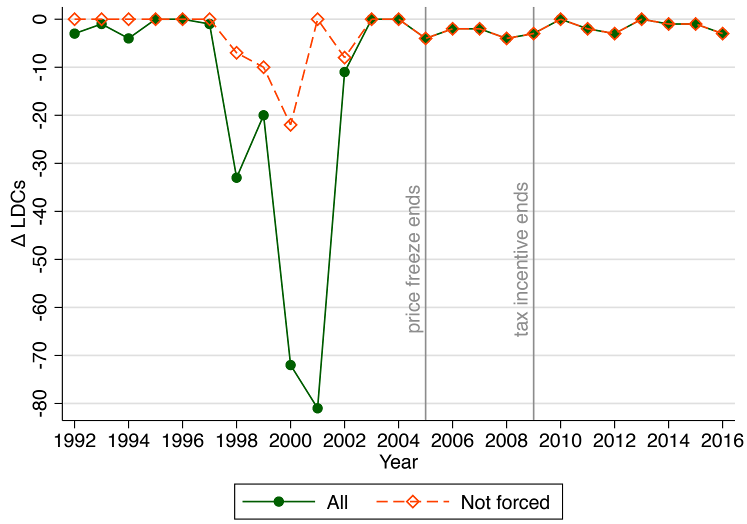
We obtained almost all accounting books for each LDC from the Ontario Energy Board. From the data we can calculate distribution costs per unit of electricity and per customer. We can also disaggregate these costs into administration, operation and maintenance, and others. The costs of power purchased are also available. We also have information on the length of distribution lines and the fraction of coverage that is urban or rural as well as the fraction of lines that are underground. From these data we can observe the set of mergers that occurred during our sample period. In addition to the accounting data, we use information on the location of the LDCs to construct distance measures between each pair of companies. We use these measures to determine feasible merger sets in our algorithm.

Our study concentrates on the 2003-2016 time period, after the two big consolidation episodes mentioned in the previous section. The merger rate is defined as the change in the number of LDCs from year to year in Ontario, Figure 1 shows its evolution. After 2002, only a few (at most four) LDCs were acquired each year. The vertical lines indicate changes in the policy environment. In 2005 the price freeze ended and 2009 was the last year of the tax incentive for private LDCs.[13]

Table 1 presents summary statistics for the first and final years of our sample in Panels A and B, respectively. In 2003 there were 93 LDCs in operation, including Hydro One (Figure A.1 presents a map of the remaining LDCs). The average cost per MWh was \$12.86, but this masks tremendous heterogeneity (minimum of of \$4.06/MWh and maximum of \$45.96/MWh). There is also important heterogeneity in size, with LDCs ranging in size from 569 customers to 668,625 (not including Hydro One Inc. which is owned by the provincial

[13]Ontario deregulated the generation segment of the industry on May 1, 2002, which caused electricity spot prices to spike (similarly to what happened during the California crisis). A price ceiling was imposed on the spot prices (generation). This affected the cost of purchased power LDCs faced and some of them passed-through these price increases onto consumers. This may have had an impact on the incentives to delay or expedite mergers among LDCs.

Figure 1: Annual change of number of LDCs in Ontario



Notes: Each point represents the change in the number of LDCs relative to the previous year. The vertical lines indicate changes in the policy environment. In 2005 the price freeze ended and 2009 was the last year of the tax incentive for private LDCs.

government and had about 1 million customers).[14] Figure 2 presents a histogram of LDC sizes.

From Panel B we can see that the number of LDCs in operation by the end of our sample was just 64. We can also see that consolidation resulted in much larger LDCs: the average number of customers increased from 35,758 in 2003 to 47,969 in 2016. Average cost increased by almost 25% from $12.86/MWh to $15.95/MWh.

Table 1 also presents statistics on the number of potential merging partners. We calculate this by considering pairwise distances between all LDCs in our sample. Figure 3 presents a histogram of these pairwise distances. While there are some pairs that are located far apart, there are many LDC pairs that are located quite close together in the more populated parts of the province. From Table 2, which provides information on the consummated mergers, we

[14] In its recommendations for consolidation the Review Panel suggested that Hydro One's assets be sold off in pieces to the conglomerates that would be formed under the restructuring. Therefore, we do not allow Hydro One to acquire other LDCs, and since, the panel provides no guidance as to how Hydro One should be partitioned, in our algorithm we do not allow Hydro One to be split up.

Table 1: Summary statistics

	Panel A: 2003				
	Mean	Std. Dev.	Min	Median	Max
Avg. cost ($/MWh)	12.86	5.70	4.06	11.87	45.96
Density line (cust./km)	46.37	19.22	6.27	45.66	85.39
Price of capital ($/km)	88,333	36,248	15,319	90,781	188,997
Electricity sold (kWh/customer)	25.81	6.93	10.67	24.93	45.07
Total customers	35,760	81,027	189	12,810	668,625
Net income (mill. $)	0.019	29.8	−274	0.8	42.6
Fraction urban serv. area	0.69	0.38	0.0	1	1
Fraction overhead lines	0.73	0.20	0.04	0.77	1
Publicly owned	0.98	0.15	0	1	1
Avg. # potential merging partners	27.4				
N	92				

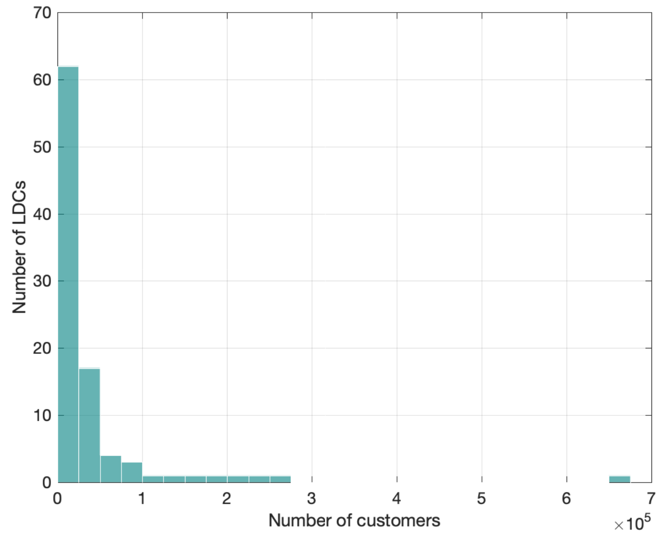
	Panel B: 2016				
	Mean	Std. Dev.	Min	Median	Max
Avg. cost ($/MWh)	15.95	7.83	6.19	14.59	69.62
Density line (cust./km)	47.30	17.32	6.33	47.17	81.97
Price of capital ($/km)	111,819	41,228	24,397	107,433	244,611
Electricity sold (kWh/customer)	22.57	4.91	10.02	21.90	36.57
Total customers	47,969	105,754	1,247	19,731	761,920
Net income (mill. $)	4.81	18.0	−0.26	1.15	142
Fraction urban serv. area	0.68	0.37	0.0	0.94	1
Fraction overhead lines	0.68	0.18	0.26	0.69	0.99
Publicly owned	0.97	0.17	0	1	1
N	64				

Notes: 2003 is the first year in our sample. 2016 is the last year in our sample. Statistics do not include Hydro One Networks Inc. The average number of potential merging partners is calculated by counting how many other LDCs are found within a radius of 300 km from a given LDC including geographical constraints (lakes) and this is the only statistic in this table that was calculated using the initial pool of LDCs in 2003 and not the data from 2016. The price of capital is defined as the total value of assets divided by the total length of lines.

can see that all mergers involved LDCs that are located fairly close together. Therefore, in our analysis we will impose a maximum distance for mergers. We select this to be 300 km, which allows for ample flexibility in the definition of the feasible sets as discussed in further detail below.[15] Figure 4 presents a histogram of the number of potential merging partners. The average is 27.4, but there is considerable dispersion with many firms having fewer than

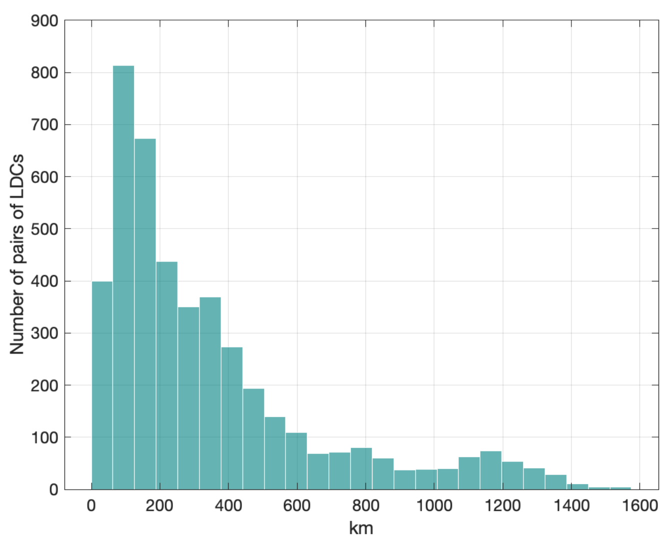
[15]From the map in Figure A.1 we can see that some LDC pairs may fall inside the 100 km limit by travelling across Lake Ontario. We forbid these links in our analysis.

Figure 2: Distribution of LDC sizes in 2003



Notes: See Table 1 for further details.

Figure 3: Distribution of pairwise distances for set of LDCs in 2003

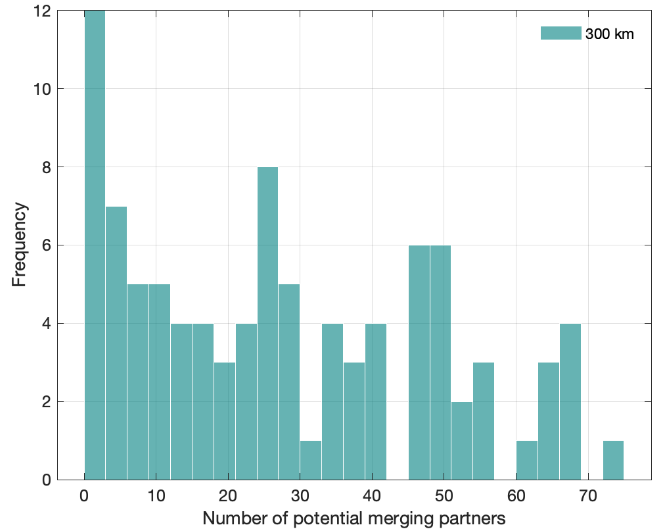


Notes: The distance for each pair among the LDCs in 2003 is calculated assuming that the headquarters are located in the most populous city or town within the geographic region of each LDC.

this, and some having as many as 75.

Table 2 provides details on the mergers that took place during the sample period. The

Figure 4: Potential merging partners in 2003



Notes: For each LDC, we count how many other LDCs are found within a radius of 300 km. This histogram shows the frequencies of potential number of merging partners. The average is 27.4 potential merging partners.

table provides information on the size of the LDCs, their average costs and the Euclidean distance between each (as measured by the city centers). From the table we can see that the mergers that took place were almost always between LDCs of different sizes and average costs, although towards the end of the period there are a few mergers between rather large LDCs. Mergers were also between LDCs that were located fairly close together (at most 295.6 kms). Note that this list does not include mergers involving Hydro One, since we will not attempt to model its acquisitions as it is owned by the provincial government.[16]

Table 2: List of observed mergers during the sample period

Year	Name	Nbr custs	Avg. cost	Distance
2004	1: Asphodel Norwood Distribution Inc.	682	13.12	(1, 2): 24.45
	2: Peterborough Distribution Inc	33,438	9.90	(1, 3): 26.01
	3: Lakefield Distribution Inc.	1,378	11.03	(2, 3): 13.04

Continued on next page

[16]Hydro One acquired Terrace Bay (2006), Hydro One Remote Communities (2007), Norfolk (2014), Haldimand County Hydro Inc. (2015), and Woodstock Hydro Services Inc. (2015). Reliable data for Hydro One Remote Communities are not available and so we do not even include it in the summary statistics presented in Table 1.

				Continued from the previous page
Year	Name	Nbr custs	Avg. cost	Distance
	1: Hamilton Hydro Inc.	177,495	7.03	52.16
	2: St. Catherines Hydro Utility Services	51,979	8.76	
	1: Scugog Hydro Energy Corporation	2,340	10.64	26.98
	2: Veridian Connections Inc.	101,867	11.08	
2005	1: Aurora Hydro Connections Limited	16,039	8.08	18.72
	2: PowerStream Inc.	203,749	13.15	
	1: Wellington Electric Distribution	3,416	27.57	0
	2: Guelph Hydro Electric Systems inc.	54,520	8.93	
	1: Gravenhurst Hydro Electric Inc.	5,928	18.39	120.57
	2: Veridian Connections Inc.	100,802	12.46	
2006	1: Newmarket Hydro Ltd.	26,647	15.11	84.58
	2: Tay Hydro Electric Distribution Company Inc.	4,037	27.78	
2007	1: Dutton Hydro Limited	600	17.52	33.98
	2: Middlesex Power Distribution Corporation	6,957	12.28	
	1: Niagara Falls Hydro Inc.	34,704	16.59	32.24
	2: Peninsula West Utilities Limited	15,491	22.58	
	1: West Nipissing Energy Services Ltd.	3,284	9.34	84.91
	2: Greater Sudbury Hydro Inc.	43,167	22.67	
	1: Grand Valley Energy Inc.	681	24.57	20.36
	2: Orangeville Hydro Limited	10,200	13.33	
2008	1: Barrie Hydro Distribution Inc.	69,628	14.85	61.44
	2: Powerstream Inc.	244,573	13.07	
	1: Newbury Power Inc.	199	16.25	33.86
	2: Middlesex Power Distribution Corporation	7,026	11.16	
2010	1: Erie Thames Powerlines Corporation	14,373	14.42	(1,2): 54.14
	2: West Perth Power Inc.	2,049	16.14	(2,3): 32.26
	3: Clinton Power Corporation	1,639	20.86	(1,3): 83.41
2011	1: Canadian Niagara Power Inc.	15,708	28.50	(1,2): 19.36
	2: Port Colborne Hydro Inc.	9,138	22.58	(2,3): 295.56
	3: Eastern Ontario Power Inc.	3,551	22.87	(1,3): 277.82
	1: Chatham-Kent Hydro Inc.	32,132	16.51	76.92
	2: Middlesex Power Distribution Corporation	7,988	12.23	
2013	1: Lakeland Power Distribution Ltd.	9,765	16.91	64.05
	2: Parry Sound Power Corporation	3,463	24.89	
2014	1: Brant County Power Inc.	10,058	14.27	28.09
	2: Cambridge and North Dumfries Hydro Inc.	53,106	12.18	
2015	1: Enersource Hydro Mississauga Inc.	204,728	11.56	(1, 2): 42.08
	2: Horizon Utilities Corporation	244,114	13.92	(1, 3): 13.19
	3: Hydro One Brampton Networks Inc.	158,630	11.47	(1, 4): 28.61
				Continued on next page

Continued from the previous page				
Year	Name	Nbr custs	Avg. cost	Distance
	4: PowerStream Inc.	364,505	13.53	(2, 3): 48.16
				(2, 4): 70.69
				(3, 4): 27.15

Notes: No observed mergers in 2009 or 2012. Average total cost is in \$ per MWh. Distance is measured in km. The notation $(m, n) : x$ means that the distance between LDCs m and n from the corresponding set of merged entities is x.

3 Average Cost Curve and Stochastic Frontier for LDCs

In Section 4 we develop a sequential merger algorithm that we use to compare predicted mergers to actual consolidation patterns observed over the same period. In order to operationalize this approach we must first estimate the industry average cost curve so that we can forecast the cost to assign to the merged entity. Following the work in Kwoka [2005] we estimate an average cost curve in two dimensions. The first argument of this function is electricity sold, while the second is a measure of density, defined as the number of customers per kilometer of line. To do so, we estimate a stochastic frontier for costs (see Kumbhakar and Lovell [2003] for an extensive treatment of the subject and Knittel [2002] for an application to the U.S. electricity industry). We start with a cost function for firm i in year t of the form

$$C(q_{it}, density_{it}) = f(q_{it}, density_{it}, \mathbf{W}_{it}; \theta)\xi_{it} \exp(\epsilon_{it}),$$

where q_{it} is electricity output, $density_{it}$ is defined as the ratio of the number of customers and the kilometers of line, $\mathbf{W}_{it}$ is a vector of observable factors that influence costs, θ is a vector of parameters to be estimated, and ϵ_{it} is the unobservable error term. $\xi_{it} \geq 1$ is the firm's level of inefficiency. If $\xi_{it} = 1$, we say that the firm is at the *cost frontier*. Deviations from this cost frontier are associated with values of $\xi_{it} > 1$. By taking logs on both sides and assuming a functional form for f we obtain

$$\log C(q_{it}, density_{it}) = \theta_0 + \theta_1 \log q_{it} + \theta_2 \log density_{it} + \sum_k \theta_k \log W_{itk} + \log \xi_{it} + \epsilon_{it},$$

where $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$ and the term $\log \xi_{it}$ follows a truncated normal distribution with parameter σ_ξ^2. Using the estimates from this regression we can recover the set of estimates $\hat{\xi}_{it}$, which gives a ranking of the different firms in terms of their inefficiency levels or inefficiency scores for each year t: higher values correspond to more inefficient firms.

Using the rest of the estimates and the predicted values for $C(q_{it}, density_{it})$ we compute average costs as $\widehat{AC}(q_{it}, density_{it}) \equiv \hat{C}(q_{it}, density_{it})/q_{it} = \widetilde{AC}(q_{it}, density_{it})\hat{\xi}_{it}$ where $\widetilde{AC}(\cdot) = \hat{f}(\cdot)/q_{it}$ is the predicted average cost associated with the cost frontier (at $\xi = 1$). Note that the predicted average cost, $\widehat{AC}$, is the product of the average cost at the frontier, $\widetilde{AC}$, times the inefficiency term. We then consider the set of scattered data points of predicted average costs over a two-dimensional grid formed by q_{it} on one axis and by the density of line on the other and find a surface that best interpolates these scattered data points. This is our estimate for the average cost curve.[17] Extrapolation to find the predicted average cost curve after a merger is done by using linear approximations out of sample based on the interpolated function.

It is important to point out that there are other methods that have been proposed in the literature and that we could have used to estimate the average cost curve. Such methods include partial linear regressions (see Yatchew [2000]), translog cost functions, and Data Envelopment Analysis.[18] While all these methods have their advantages, we opted for the stochastic frontier approach because it allows us to have a fully specified functional form to predict average costs at given points of output and density that do not significantly increase the computational time in our merger simulations.

4 Endogenous Mergers

Our objective is to provide a framework for determining which LDCs merge and with whom. Below we specify a sequential algorithm, which allows buyers to make acquisition offers to potential sellers. As mentioned in the introduction, because there is no price or quantity competition between LDCs, nor are there any entry and exit decisions, we are able to avoid dealing with the complications stemming from a dynamic framework that takes into account pricing/production, entry/exit, and investment. We set up a static model that captures the

[17] We use interchangeably the terms curve and surface.

[18] See Greene [2008] for a survey of these alternative methods.

industry reconfiguration that took place from 2003 to 2016, assuming that LDCs had the entire period in which to reorganize.

4.1 Model setup

Profits for an LDC with q_i customers and facing a price cap of $\bar{p}_i$ are given by

$$\pi_i = q_i \times (\bar{p}_i - AC(q_i, density_i)).$$

In principle, it is possible to determine the optimal size q^* that would maximize profits if we disregard for a moment the dependence on the density.[19] However, this optimal size might not be attainable given that each company adds a block of customers and not a divisible set of them and it is not possible to shuffle customers in order to attain a certain firm size.

Instead, our algorithm will involve LDCs deciding whether to make offers by comparing profits from merging to profits from staying alone. If the model only requires computation of an existing firm's profits, we use the observed price cap and the actual average cost observed in the data. The complication arises when firm i acquires firm j. In this case we specify the new firm's profits to be:

$$\pi_{ij} = \bar{p}_i q_i + \bar{p}_j q_j - \left(\widetilde{AC}(q_i + q_j, density_{ij}) \times H(\xi_i, \xi_j)\right) \times (q_i + q_j) - Z_{ij},$$

where the product $\widetilde{AC}(q_i + q_j, density_{ij}) \times H(\xi_i, \xi_j)$ mimics our definition of the predicted average cost. $H(\cdot)$ is a function of inefficiency levels as defined below and $density_{ij}$ is the sum of the number of customers divided by the sum of the km of line.

As discussed in Section 2.1, price harmonization is not required following the merger, and instead consumers pay the price cap associated with each of the LDCs that constitute the merged entity. Therefore, revenues are just given by: $q_i\bar{p}_i + q_j\bar{p}_j$.[20] We must also determine the appropriate AC value to assign to the merged entity. One possibility is to take the point on the estimated AC curve associated with the merged entity's size, $q_i + q_j$, and its combined

[19] To do so, take the first order condition of the profits function and solve for the number of customers as a function of the regulated price and the average cost curve to obtain $AC(q^*) + q^* \times AC'(q^*) = \bar{p}$, which simplifies to the usual relationship $C'(q) = \bar{p}$.

[20] With price harmonization, we would have to introduce a common price, $\bar{p}_{ij}$, and revenues would be $\bar{p}_{ij}(q_i + q_j)$.

new density. The problem is that doing so implies that the merged entity is exactly at the average for a firm of its size and density, regardless of the initial level of the inefficiencies of the merging parties. We denote buyer and seller inefficiencies by ξ_i and ξ_j and estimate them using the stochastic frontier analysis presented in the previous section. When merging, these inefficiencies are combined. We capture this using H, a relative-influence function that maps the efficiency levels of the merging firms into a positive factor that creates a deviation from the forecasted average cost $\widehat{AC}$ for the merged entity's size. We choose a linear parameterization for H of the form $H(\xi_i, \xi_j) = \alpha\xi_i + (1-\alpha)\xi_j$. The parameter α represents the relative influence of the buyer's performance deviation. One goal of our calibration exercise is to determine α.

Note that H does not take into account the effect of the number of firms on costs, nor the distance between them. That is because, the cost function depends only on the number of customers, not on which LDCs were combined to achieve that size. In practice, the cost of serving a specific number of customers could depend on the number of LDCs that are combined to do so. Moreover, adding further LDCs to the combination increases the distance from the initial buyer, possibly generating additional costs. The inclusion in the profit function of the interconnection cost term, Z_{ij}, allows us to capture these costs. We parametrize the interconnection costs as a quadratic function of the number of firms in the conglomerate, I_{ij}, at the moment when i acquires j (and including j). Specifically $Z_{ij} = \lambda I_{ij}^2$ and we calibrate λ from the data.

4.2 A model for offers

We model the consolidation process as a sequence of decisions in which LDCs choose whether to make a proposal to buy another firm, that in turn must decide whether or not to sell. We use the index i for the word *buyer*, and the index j for the word *seller*. We discuss in Section 5.1 below the sequences that we consider.

For any given sequence, if a buyer acquires a seller, the net gains from the transaction are,

$$NG_{buyer} = \pi_{ij} - b_{ij} - \tau a_j - \pi_i + s_{ij}$$

for the buyer, and

$$NG_{seller} = b_{ij} - \pi_j$$

for the seller.

b_{ij} represents the price paid by the buyer and a_j is the annualized value of firm j's assets. The tax is calculated as a fraction τ of the annualized value of the seller's assets, following regulations on acquisitions in our case study. We take the tax environment as exogenous, so that τ is a fixed parameter in the calibration process. However, we adjust this parameter in the policy simulations once the structural parameters have been calibrated. Finally, s_{ij} represents a synergy shock that is generated by the merger and that represents a random benefit or cost that arises once the two LDCs become one and that is not directly captured by our cost function. s_{ij} is observed by both firms at the time of the acquisition. However, since s_{ij} is not an observable in our data, we simulate random draws from a known distribution and average the outcomes from each of those draws. We assume a uniform distribution for the cost/synergy random shock $s_{ij} \sim U[-s^{\max}, s^{\max}]$ as in Gowrisankaran [1999] and Jeziorski [2014].

The values $\{b_{ij}\}$ can be determined as the solution to the following Nash Bargaining problem with bargaining weight η:

$$\begin{aligned} \max_{b_{ij}} NB &= \max_{b_{ij}} \{(NG_{buyer})^{\eta}(NG_{seller})^{1-\eta}\} \\ &= \max_{b_{ij}} \{(\pi_{ij} - b_{ij} - \tau a_j + s_{ij} - \pi_i)^{\eta}(b_{ij} - \pi_j)^{1-\eta}\}. \end{aligned}$$

Its solution is

$$b^*_{ij} = (1-\eta)(\pi_{ij} - \tau a_j + s_{ij} - \pi_i) + \eta \pi_j. \tag{1}$$

The optimal offer b^*_{ij} is increasing in the profits of the seller and in the difference between the joint profits of the merged entity and the buyer's profits. It is also increasing in the synergy shock. On the other hand, it is decreasing in the size of the tax and the annualized value of firm j's assets. In the empirical analysis below we assume equal bargaining weight and set $\eta = 0.5$, but we present results for alternative values and show that our findings are robust to these changes.

The model specifies that, when its turn to make offers arises, a buyer behaves rationally, tests all feasible combinations, and chooses its best option. In contrast, the seller is assumed to be myopic at this stage. It simply compares the benefit from continuing alone to the benefit from being acquired. Obviously, a more appropriate representation would allow the seller to consider the possibility that it could be acquired by another firm, and/or that it might have the opportunity later to make an offer itself. Unfortunately it is not possible to incorporate this into the analysis, since our potential merging partners sets are sometimes very large (an average of almost thirty LDCs), meaning that there are too many permutations to consider. The sequential approach described above can, therefore, not be considered an equilibrium. However, we believe that our approach is reasonable, since it is stable in the sense that no LDC wishes to acquire additional firms. Jeziorski [2014], Fan and Yang [2020], and Seim and Waldfogel [2013] all use similar approaches.[21]

4.3 Acquisitions

Our merger algorithm can be summarized as follows:

1. For a given sorting of firms, label firms with indices such that if firm l appears in this sorting before firm l', then $l < l'$. The firm at the top of the list moves first.

2. An acquisition attempt occurs only if firm l (the buyer) has profits above the minimum profitability threshold, given by $\underline{\pi}$. If this condition is satisfied, l can make offers to firms $k > l$ in the feasible set, which is defined by a maximum distance between k and l given by a fixed threshold, $\bar{D}$. Otherwise we restart with firm $l + 1$.

3. The firm l under consideration computes b^*_{lk} for each firm in the feasible set and computes NG_{lk} for each of the potential acquirees.

[21] The first two use algorithms that share the following properties with ours: they are sequential over a pre-specified sorting of the firms, they are heuristic algorithms not based on an equilibrium condition, and their main purpose is to reduce the state space. The third also uses a heuristic algorithm to find profitable additions to the set of stores. In these three cases, the full examination of all the possible configurations is prohibitive due to the size of the set.

4. If every element of the vector formed with the different values of $NG_{l\cdot}$ is negative, no merger is possible and we start over with firm $l+1$ making offers. If the maximum of the vector of $NG_{l\cdot}$ values is positive, we execute the merger between l and the LDC that yields this maximal value, $\bar{k}$. We add the number of customers, compute the new density of customers per km of line, and compute the new average cost.

5. If a merger occurred in the previous step to create a new firm $\{l\bar{k}\}$, then if $\pi_{l\bar{k}}$ exceeds the profit threshold $\underline{\pi}$, $\{l\bar{k}\}$ computes $b^*_{\{l\bar{k}\}m}$ for every firm $m > l, m \neq \bar{k}$ in the feasible set of $\{l\bar{k}\}$. If the maximum of the new collection of NG values is positive, we execute the merger. We repeat this step until the maximum of the vector of NG values is negative or profits of the acquirer fall below the minimum profitability threshold. We call this particular sequence of acquisitions a path, or a conglomerate.

6. Once we have reached the end of a path, we move to the next firm $l+1$ (if $l+1$ has already been acquired, we continue with $l+2$, etc.) and return to step 1 keeping the sorting of unacquired firms the same as before any offers took place.

5 Model Calibration

In this section we discuss the calibration of our model. We first discuss the sequence in which LDCs are assumed to make offers. We then describe the fixed parameters used in the procedure and the tax structure that characterizes the baseline scenario. Finally, we describe the calibration approach, which follows a method similar to a minimum distance estimation procedure, searching over the values of α and λ to find the values that best replicate the change in industry configuration between 2003 and 2016 using the merger simulator.

5.1 Offer sequence

The merger algorithm described above requires us to input a particular sequence that specifies the order in which LDCs are chosen to get to make offers to others in their feasible sets. There is no obvious way to select this sequence. The literature has often elected to allow *the largest* firms to move first (see for instance Gowrisankaran [1999] and Jeziorski [2014]), and indeed most of our observed mergers involve at least one of the larger LDCs. However, as Igami and Uetake [2019] and Hollenbeck [2020] point out, there are no empirical or theoretical foundations that support one specific particular sorting. Both Igami and Uetake [2019] and Hollenbeck [2020] recognize that there are multiple configurations of mergers that can be optimal. If all firms were allowed to make simultaneous offers, we would have $N(N-1)$ bids, each of them being a function of the remaining offers. In our case $N = 92$, yielding more than eight thousand such functions. Moreover, we are not only interested in conglomerates consisting of two firms, but of as many as optimal for exploiting the benefits of economies of scale. The number of such sets is $2^N - 1$, which is computationally prohibitive for any type of calculation. To overcome these challenges, the authors consider random orderings.

Our approach is a compromise between these two perspectives, while maintaining computational tractability. We construct a large number of different sequences (100) that each allow bigger firms to move earlier than smaller ones, but that also allow for randomness within the two groups (bigger and smaller firms). Specifically, we require that the first ten movers in each sequence serve 60% of all customers in Ontario (which represents 1.99 million customers). That is, we only keep a sequence among all the possible permutations of our list of LDCs if this condition is satisfied. Flexibility arises because there are many different possible ways that larger firms can be sorted and that lead to this coverage, and because once the first ten movers are assigned, the remaining ordering is completely random. Once we have our 100 sequences, we run our algorithm for each and we report the average results based on these. In the empirical analysis below we present results for alternative sequences that also prioritize bigger LDCs and show that our findings are robust to these changes.

5.2 Fixed parameters for grid search

In this subsection we explain how we use the data to motivate our choices of values for the fixed parameters used for the grid search. There are five parameters to consider: (i) the bound of the random synergy shocks, $s^{\max}$, (ii) the price cap, $\bar{p}$, (iii) the maximum distance allowed between merging firms, $\bar{D}$, (iv) the profit threshold to be able to make acquisitions, $\underline{\pi}$, and (v) the transfer tax, τ.

$\boldsymbol{s^{\max}}$: From multiple sources we determined the amounts of money transferred for the mergers in our data. Since all of our computations are on an annual basis, we transform these purchase amounts into annual amounts by using a depreciation rate of 0.041 and a discount factor for the electricity distribution industry of 0.06 as in Lim and Yurukoglu [2016]. The first number implies a lifetime of 24.4 years and altogether we obtain a conversion factor of 0.37 for the observed transaction amounts. From our main dataset we extract the net income amounts for the firms involved in the acquisitions and, together with the value of the tax on the transaction, we can recover an implied value for $s^{\max}$ by using Equation 1, assuming $\eta = 0.5$. We then select the median of the implied values for s_{ij} to obtain \$4.88 million, as the value of $s^{\max}$ we take to our simulations. This is a conservative choice since the maximum of our implied values is well beyond the range of profits values in our sample.

$\boldsymbol{\bar{p}}$: For each year we take the average revenue in MWh as a proxy for the LDC-specific price cap, $\bar{p}_i$. It is not possible to extract one single price cap from the regulatory forms to the OEB because each LDC has multiple rates: residential, commercial, public lighting, and others. If LDC i merges with j, customers associated with buyer i continue to pay $\bar{p}_i$, while those associated with seller j continue to pay $\bar{p}_j$. Revenues in this case are given by $q_i\bar{p}_i + q_j\bar{p}_j$.[22]

$\boldsymbol{\bar{D}}$: We only allow mergers to occur in the simulations if the distance from centre to centre (of the respective municipalities) is less than a fixed threshold of 300 km, which is just slightly above the maximum distance among the observed mergers in the data (295 km). If a conglomerate has already formed, our algorithm will check within $\bar{D}$ of the centre of *each* of the municipalities making up the conglomerate.

$\boldsymbol{\underline{\pi}}$: We only allow LDCs to make offers if they had positive profits. This is consistent with what we observe in the data: eight out of the firms involved in actual mergers in the data had negative profits in the year of the acquisition, but in all but one of these cases the negative profits belonged to the smaller LDC, which we might imagine is not the acquirer.

$\boldsymbol{\tau}$: The tax, τ, imposed on the merging parties depended on whether the buyer and seller were public or private. As of 2003, a transfer tax of 22 per cent was imposed on the fair

[22] p_i is censored at 0 if the LDC's revenue was negative.

market value of the electricity assets sold to the private sector. Public to public acquisitions were not taxed. In the counterfactuals we adjust the value of τ to reflect the policy that was proposed to encourage consolidation. As we will discuss in more detail below, since the proposed policy has little impact on industry configuration, we also allow τ to become negative, reflecting a subsidy for consolidation.

The values of the fixed parameters are summarized in Table 3.

Table 3: Fixed parameters for grid search

Parameter	Description	Value
$s^{\max}$	Upper bound of the random synergy shocks	\$4.88 million
$\bar{p}_i$	Price cap for LDC i (ij)	i's average revenue
$\bar{D}$	Upper bound on distance between merging firms	300 km
$\underline{\pi}$	Lower bound on profits of buyer	0
τ	Policy parameter	0 if public to public, 22% if one of the firms is privately owned

Notes: See main text for further details.

Note that there are two other parameters that play important roles in our algorithm for which data do not exist to guide the choice of values: the Nash bargaining parameter, η, and the size/importance of the first movers in our sequence. Since these values are arbitrarily chosen, we perform some checks to confirm the robustness of our results to alternative levels.

5.3 Calibration of the merger algorithm

With the estimated average cost function and parameter values in hand we can use the algorithm described in section 4. Our calibration procedure is similar to a minimum distance estimator, searching over combinations of α and λ to find the values that best replicate the change in industry reconfiguration between 2003 and 2016. We then use the estimated values

$\hat{\alpha}$ and $\hat{\lambda}$ to simulate mergers under different tax incentive scenarios to evaluate the Review Panel's policy recommendations.

For a given pair (α, λ), the rest of model parameters, and the semi-parametric estimate of the average cost function $\widehat{AC}$, each merger simulation yields a single market configuration. As in the minimum distance estimator in Stahl [2016] and Lim and Yurukoglu [2016], we penalize for mergers predicted by our model that are not observed in the data, and for observed mergers that are not predicted by the model.

Consider for instance a scenario with five LDCs in which the specified sequence was ABCDE and where a conglomerate comprised of LDCs B,C, and D (i.e. BCD) is the only one observed in the data. Suppose further that the model predicts only the formation of ABC and DE. Then we should be penalizing (i) the model's prediction of AB, (ii) the model's prediction of ABC, (iii) the model's prediction of DE, (iv) the model's inability to predict BC, and (v) the model's inability to predict BCD. Note however that the model did successfully pair B and C together when it formed ABC and so we should reward it for that.

Formalizing this logic, we solve the following problem:

$$\min_{\substack{\alpha \in [0,1] \\ \lambda > 0}} \left\{ F(\alpha, \lambda) = \sum_{J \in \mathcal{J}_1} \left(\sum_{c \in \sigma(J)} (NG_c(\alpha, \lambda))^2 - \sum_{c \in \mu(J)} (NG_c(\alpha, \lambda))^2 \right) + \sum_{J \in \mathcal{J}_2} \sum_{c \in \sigma(J)} (NG_c(\alpha, \lambda))^2 \right\},$$

where $NG_{\cdot}$ are the buyer's net gains after plugging in its optimal bid. c is a path (not necessarily at the end of its construction) from an element of the set of sequences $\sigma(\cdot)$ or $\mu(\cdot)$ that are defined below. The set $\mathcal{J}_1$ contains the final conglomerates predicted by the merger algorithm that are not observed in the data, and that may or may not contain a strict subset (in any order) that matches one of the conglomerates observed in the data. Returning to the example above, $\mathcal{J}_1$ contains ABC and DE. The set $\mathcal{J}_2$ contains the final conglomerates observed in the data, but not predicted by the algorithm. So $\mathcal{J}_2$ contains BCD. $\sigma(\cdot)$ is a function that takes an ordered list of firms and outputs all the sequences formed by adding, in the same order as the list, each of the members of the list. So $\sigma(\{A, B, C\}) = \{\{A, B\}, \{A, B, C\}\}$ and $\sigma(\{D, E\}) = \{D, E\}$. $\mu(\cdot)$ is defined similarly but it only acts over the members of its argument that belong to observed mergers: $\mu(\{A, B, C\}) = \{B, C\}$. If there are not at least two members of $J \in \mathcal{J}_1$ that belong to any observed merger, then

$\mu(J) = \varnothing$ and its corresponding term in the expression above is set to 0. The negative sign in front of the term that uses $\mu(\cdot)$ reflects the fact that these combinations should be rewarded.

6 Results, Goodness of Fit, and Counterfactuals

In this section we present our results. We first discuss results from the estimation of the average cost curve, which are a necessary input to the algorithm described in section 4. Next we present calibration results and characterize the fit of the model. Finally, we present counterfactual results from different policy experiments.

6.1 Average cost curve estimation

Using the estimation strategy described in section 3, we estimate a stochastic frontier model. As explained in section 2, we focus on the period 2003-2016.[23] Our estimates for the industry average cost curve are displayed in Table 4. Column (1) includes only total electricity output and line density (defined as the number of customers per kilometer of line), while column (2) adds a control for the price of capital and column (3) includes this variable along with controls for the fraction of urban area covered and the fraction of overhead lines. The table also presents estimates for σ_ξ^2 and for the inverse logit parameter $\gamma = \sigma_\epsilon^2/(\sigma_\epsilon^2 + \sigma_\xi^2)$.

We find that electricity output is positively correlated with costs and that line density has a negative correlation with costs. Columns (2) and (3) show that the estimated coefficients are very similar even when including additional controls, and so we are comfortable using only the estimates from model (1) for our merger simulations.

Figure 5 plots the inefficiency estimates for each of the LDCs, both for all years combined (in blue) and separately for 2003 (unshaded), since that is the starting year for the simulations. There is almost no difference in the distribution of the scores for the two distributions. The average inefficiency level in the full sample is 2.92, with a standard deviation of 0.89. Although there are some LDCs operating close to the minimum (recall that inefficiency scores

[23] We use the "time invariant" form of the stochastic frontier model, which consists of assuming that $\xi_{it} = \xi_i$ for all i. For further details and derivations see Kumbhakar and Lovell [2003].

Table 4: Stochastic frontier analysis. 2003-2016.

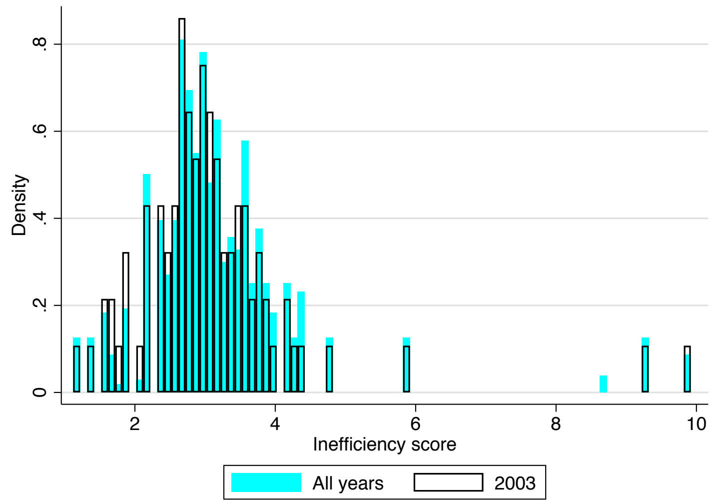
	(1)	(2)	(3)
log q	0.839***	0.838***	0.847***
	(0.0352)	(0.0256)	(0.021)
log density	−0.120**	−0.343***	−0.370***
	(0.0591)	(0.0577)	(0.0547)
log price capital		0.236***	0.244***
		(0.0332)	(0.0347)
log frac. urban area			−0.0210
			(0.0150)
log frac. overhead lines			−0.0571
			(0.0397)
constant	4.082***	2.362***	2.282***
	(0.408)	(0.440)	(0.446)
log σ_{ξ}^2	−1.885***	−2.079***	−2.193***
	(0.232)	(0.177)	(0.143)
inv. logit γ	1.230***	1.009**	0.795***
	(0.332)	(0.267)	(0.228)
μ	1.092***	0.993***	0.930***
	(0.210)	(0.225)	(0.251)
N	1,032	1,032	944

Notes: Dependent variable: log cost. We use the results from specification (1) for our merger simulations. Year 2010 dropped due to inconsistencies in the raw data. Estimation includes Hydro One Networks Inc. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

cannot be less than one), there are also some with inefficiency levels more than two standard deviations above the mean. These results can be compared with those from other studies on electricity distribution. Hess [2007] analyzes cost efficiencies for U.S. electricity distribution companies. He finds inefficiency scores ranging up to 3.3 (81% of our estimated scores are below that number), and his mean inefficiency level is a bit lower than ours at 1.13.

Recall that, as discussed in section 3, we use the predicted costs from this estimation to compute average costs. We can analyze what the distribution of LDC sizes tells us about the economies of scale of the LDCs pre- and post-mergers when using the average cost curve. This can be seen by the shape of the interpolated average cost surface shown in Figure 6. As consolidation occurs, output associated with the merged entities is larger than before the merger, which potentially places them in a region of the surface where average costs are

Figure 5: Inefficiency scores for the pooled sample and for year 2003 only

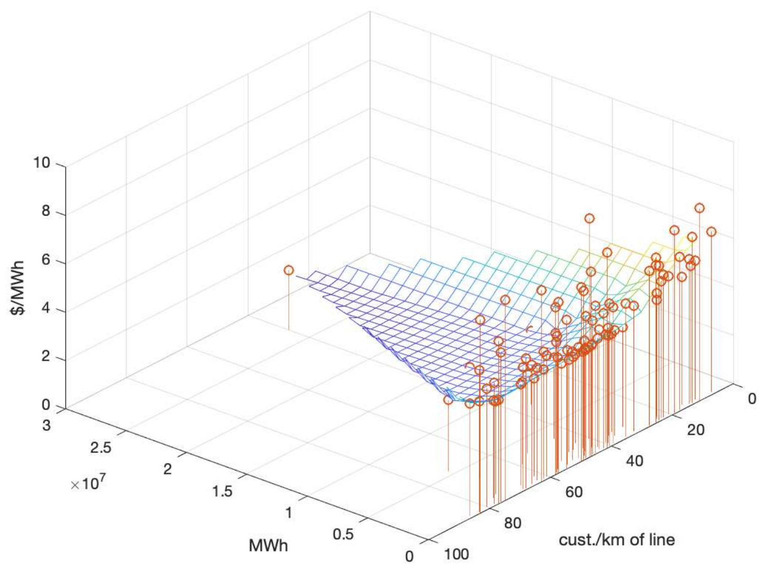


Notes: Inefficiency scores implied by the stochastic frontier analysis. The "All years" data consist of all the LDCs that ever appear in our dataset. The "2003" data show the distribution of the inefficient scores only for LDCs that existed in 2003.

lower. This is the economies of scale effect of the merger. Too much consolidation can push them further into diseconomies of scale. Our estimates from the stochastic frontier model also give us the ranking of firms with respect to their inefficiencies. These will be combined at each potential merger to determine how far from the surface is the new average cost.

We also explore the different components of total costs and the average costs we obtain with them. We can only do this analysis for the years 2009 - 2014 due to the unavailability of the cost categories decomposition for the rest of the years. Table A.1 in the Appendix repeats the regressions from the baseline but only using O&M (operating and maintenance) costs in columns 1-3 and only using administrative costs in columns 4-6. Both O&M and admin costs together account for roughly 60% of total costs during those years. The coefficients for O&M costs are similar to those from the pooled costs while the coefficients for the administrative costs specifications are not. These regressions suggest that the effect of output on total costs is the same if we only use O&M costs, but there is a lower effect if we use administrative costs only. However, density has a larger effect on O&M relative to the baseline, but a smaller effect on administrative costs. Using the results from columns 1 and 4 from Table A.1, we

Figure 6: Average cost curve as a function of q_i and line density



Notes: Data for 2003. Circles represent actual data points. Surface is the interpolated average cost curve.

obtain the average cost curves in Figure A.2 and Figure A.3, respectively.

When we compare these curves to Figure 6 we observe that the curvature in the latter is mostly due to the administrative costs component. The magnitude of the administrative average cost curve is larger than that of the O&M costs because the share of administrative costs is about twice that of the O&M costs. In 2009 the administrative costs accounted for 39% of total costs and O&M for 23%. In 2014, the last year for which we have these data by categories, the corresponding shares were 43% and 26%. In conclusion, even though the coefficients from the stochastic frontier regressions are smaller for administrative costs, the economies of scale seem to be arising from this category since the curvature is more pronounced when using these costs than when using only O&M costs.

6.2 Calibration results and goodness of fit

Table 5 displays the parameter values and standard deviations found using our merging algorithm and the grid search over α and λ. According to the calibration results α is 0.41, which implies that the buyer's influence on the resulting efficiency for the merged entity is slightly smaller than that of the seller. λ is found to be 3.047 $\times 10^6$. Table 6 shows that our model predicts a survival ratio of 0.67 (i.e. 67% of the original 92 LDCs survive until 2016). This is almost exactly equal to the survival ratio observed in the data. Our survival ratio is a bit lower because we predict slightly more conglomerates than in the data (25.39 vs 19), and these are slightly smaller in terms of number of LDCs than those observed (2.23 vs 2.26). The remaining LDCs in 2016 are a bit larger than in the data (54,200 customers vs 47,970) and they have slightly lower average cost (13.53 vs 15.95 \$/MWh). The model's ability to correctly predict the size distribution is confirmed in Figure 7, which shows that the size distribution of LDCs is similar in the data and the baseline model results. The only difference is directly related to the additional conglomerates predicted by the model. Overall we think that our merger algorithm does a good job of matching the data.

From the model we can also see that buyers are more inefficient than sellers and that the mean bid is -\$10.44 million. That the mean bid is negative may seem surprising at first, but makes sense because seller profits (π_j) are, on average, negative, led by a few LDCs with very negative profits (one with $\pi_j = -346.22$ million). The mean seller profit is $-\$10.84$ million, but the median is \$0.03 million. We also present results for the median bid and note that it is much closer to 0. Buyers seem to be targeting LDCs that are experiencing some financial difficulty. Note that the mean of bids becomes less negative as η decreases consistent with the idea that at lower η the seller has more bargaining power. On the other hand, the median of bids becomes more negative as η decreases. This is because the minimum of the bids gets higher as η decreases. In general the dispersion of the bids gets smaller as η decreases as measured by the standard deviation. Furthermore, as η decreases, both the buyer and the seller become smaller, and the ratio of buyer to seller size goes from around 1.4 when $\eta = 1$ or 0.25 to around 1.2 when $\eta = 0.50$ or 0.75. Therefore, as η decreases there is more activity among smaller LDCs but the ratio of buyer/seller sizes is similar in the extremes of too much weight on the seller or on the buyer.

In addition to results for the baseline values of the parameters in Table 5 we also present results for a number of other specifications in order to test the robustness of our model to

Table 5: Values of α and λ from grid search

Specification	α	λ
Baseline	0.4057	3.047×10^6
	(0.1910)	(1.445×10^6)
NB weight		
$\eta = 0.25$	0.4606	4.031×10^6
	(0.2197)	(1.360×10^6)
$\eta = 0.75$	0.4001	2.944×10^6
	(0.2054)	(1.441×10^6)
$\eta = 1.00$ (TIOLI)	0.4068	2.453×10^6
	(0.2067)	(1.336×10^6)
Sortings		
Largest to smallest	0.7050	1.333×10^6
First 5 LDCs account	0.5277	3.619×10^6
for 50% of customers	(0.1413)	(1.206×10^6)
First 15 LDCs account	0.3719	2.419×10^6
for 75% of customers	(0.2480)	(1.551×10^6)

Notes: Standard deviations in parentheses. The values for the "Largest to smallest" case do not have standard deviations because only one sorted list was used.

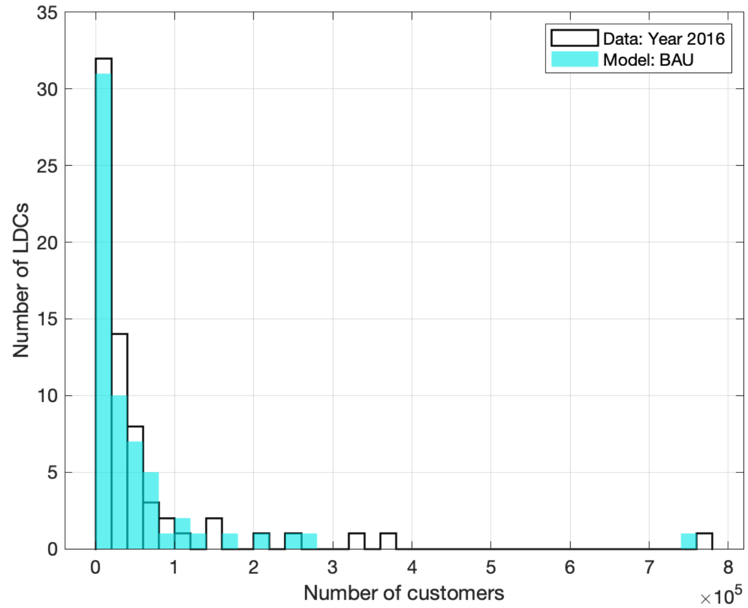
some of our assumptions. We test the robustness of our estimates to alternative values of the Nash Bargaining weight η. We allow it to vary from 0.25 to 1. At 0.25 the buyer has less bargaining power, while at 0.75 it has more. We also allow the buyer to have all the bargaining power and to make a take-it-or-leave-it (TIOLI) offer. We also consider the impact of alternative sortings. We vary the concentration of customers that must be covered by the first N movers in the sequence. Specifically, we consider three alternative sortings. In the first, we force the first 5 LDCs to account for 50% of customers, and in the second the first 15 LDCs must account for 75%. Finally, we also consider a sorting from largest to smallest LDC, as in Gowrisankaran [1999] and Jeziorski [2014]. As in the baseline case, in most of the robustness checks the buyer's influence on the resulting efficiency for the merged entity is slightly smaller than that of the seller. The value of λ is also quite robust to these changes.

Table 6: Mergers under different specifications for the BAU scenario

	Data 2016	Baseline	Alternative specifications $\eta =$ 0.25	$\eta =$ 0.75	$\eta =$ 1.00	Largest to smallest	First 5 50% cust.	First 15 75% cust.
Survival ratio	0.70	0.67	0.63	0.77	0.78	0.59	0.70	0.63
Nbr conglomerates	19	25.39	25.46	18.33	15.17	25.6	24.59	24.81
Nbr merged firms (rel. to baseline)			3.68	-9.2	-10.12	7.36	-2.76	3.68
Nbr conglom. (rel. to baseline)			0.07	-7.06	-10.22	0.21	-0.80	-0.58
Avg. Nbr LDCs/conglom.	2.26	2.23	2.36	2.11	2.34	2.49	2.11	2.39
Avg. size (1000 Cust)	47.97	54.2	56.8	46.9	47.6	61	51.2	57.5
S.D. size (1000 Cust)	105.75	113.4	115.9	105.3	105.9	111.7	112	113.3
Avg. size buyer (1000 Cust)		56.6	56.5	77	90.8	68.5	49.2	60.3
Avg. size seller (1000 Cust)		44.5	40.1	62.9	64.8	20.3	54.7	42
Avg. AC ($/MWh)	15.95	13.53	13.47	13.44	13.48	13.26	13.44	13.67
S.D. AC ($/MWh)	7.83	6.53	6.64	6.26	6.23	6.52	6.44	6.62
Avg. inefficiency ξ		3.02	3.03	3	3.02	3.14	3.01	3.03
S.D. inefficiency ξ		1.12	1.14	1.07	1.07	1.17	1.1	1.15
Avg. ineff. Buyer		3.11	3.09	3.17	3.17	3.16	3.05	3.08
Avg. ineff. Seller		2.84	2.86	2.77	2.71	2.72	3.02	2.79
Mean bid (mill. $)		-10.44	-11.44	-17.92	-19.75	-1.2	-9.46	-9.54
Median bid (mill. $)		-2.05	-5.26	0.22	0	-2.16	-2.09	-1.95
S.D. bid (mill. $)		56.06	46.37	76.19	78.93	5.07	49.26	54.81
Mean net transfer (mill. $)		0.02	0.05	0.01	0	0.02	0.01	0.02
S.D. net transfer (mill. $)		0.08	0.27	0.03	0	0.1	0.05	0.09
Combinations tested		7,637	7,672	7,529	7,563	7,823	7,605	7,813
Nbr remaining LDCs	64	61.64	57.96	70.84	71.76	54.28	64.4	57.96

Notes: Averages are non-weighted. We used 2003-2016 for estimation of AC. All years pooled for calibration. BAU is public to private taxed at 22% and public to public at 0%. The last two columns correspond to baseline scenarios where the first n LDCs in the sorted list account for X% of total customers, with $n = 5, 15$ and $X = 50, 75$.

Figure 7: Data vs. BAU distribution



Notes: Comparison of the distribution of LDC sizes between the model BAU's predictions and the data for a representative initial sequence. We set $s_{ij} = 0$. The results in the BAU column of Table 7 are the averages of 100 of the distributions shown in these graphs, each with a different sequence and with random draws for s_{ij}.

Table 6 shows the main outcomes of our model for each of the different specifications of Table 5. Again, we see fairly robust survival ratios across specifications, the lowest being 0.59 when we order largest to smallest and the biggest being 0.78 when we assume TIOLI offers. Overall we think that the model is quite robust to these different possible changes. Counterfactual simulation results derived using these alternative values are also similar to those presented below, but we do not present them here. They are however are available upon request.

6.3 Evaluation of merger-incentivizing policies

In this section we use our model to evaluate the proposed tax incentive scheme to see whether it would achieve the desired reduction in the number of LDCs. Recall that under the BAU there are no transfer taxes on public to public transactions, while transactions involving private firms are taxed at 22%. The proposed tax incentive would eliminate the tax on transactions involving private firms. To preview a bit our results, we will find, not surprisingly given the small number of private LDCs, that this incentive has almost no impact on the extent of consolidation and so we also consider the possibility that the government would subsidize transactions involving also public LDCs. We implement this experiment by allowing the transfer tax on all mergers to be negative. We consider counterfactual subsidy-scenarios the value of which varies from 25% to 200%. To get a better sense of the role of taxation, we also consider a counterfactual scenario in which all transactions, including public to public are taxed at 25%. Results are presented in Table 7.

In our discussion of the counterfactual results, we compare everything to our benchmark scenario, BAU (column (2)), which, as mentioned above, is the resulting industry configuration when using our model outcomes at the current tax regime. Our findings suggest that the Proposed tax reduction policy would have no effect on the industry configuration. This is not surprising, since presently there are few private players in the industry. For this reason we consider negative taxes, or subsidies, of various sizes. A subsidy of 25% results in the formation of just over two fewer conglomerates (22.65 vs 25.39), but these conglomerates are a bit larger (2.52 LDCs vs 2.23) and so the total number of LDCs in the industry falls to 58.88 in 2016 instead of 61.64. Increasing the subsidy to 50% results in approximately one fewer LDC in 2016. Further amalgamation is only generated with much larger subsidies. At 100% we end up with 53.36 LDCs, and at 200% just 49.68. In no reasonable case can we come anywhere close to achieving the target number of LDCs.

The consolidation generated by the experimental policies progressively increases the average size of LDCs. The increase ranges from just under three thousand customers with a 25% subsidy, to almost thirteen thousand customers for 200%. Inefficiency levels at 25% are slightly lower than the baseline, and decrease slightly with subsidy size. We also find that acquirers are on average more inefficient than acquirees in the baseline and when higher taxes are imposed on transfers. It remains true at very low subsidies (for instance at a subsidy of 10% buyer inefficiency is 3.11 vs 3.04 for sellers), but flips as the subsidy increases. At a subsidy of 25% buyer inefficiency is 3.1 compared to 3.16 for sellers. The results with subsidies

Table 7: Mergers under different policy environments

			Counterfactuals					
	Data 2016	BAU	Tax 25%	Prop. tax	Subs. 25%	Subs. 50%	Subs. 100%	Subs. 200%
Survival ratio	0.70	0.67	0.76	0.67	0.64	0.62	0.58	0.54
Nbr conglomerates	19	25.39	19.91	25.39	22.65	21.23	19.26	16.92
Nbr merged firms (rel. to bench.)			-8.28	0	2.76	4.6	8.28	11.96
Nbr conglom. (rel. to bench.)			-5.48	0	-2.74	-4.16	-6.13	-8.47
Avg. Nbr LDCs/conglom.	2.26	2.23	2.1	2.23	2.52	2.74	3.1	3.63
Avg. size (1000 Cust)	47.97	54.2	47.6	54.2	56.9	58.9	62.3	66.9
S.D. size (1000 Cust)	105.75	113.4	91.4	113.4	135.8	145.6	159	177.2
Avg. size buyer (1000 Cust)		56.6	70.9	56.6	72.6	82.5	101.1	131.6
Avg. size seller (1000 Cust)		44.5	5.4	44.5	60.6	59.8	57.1	53.8
Avg. AC ($/MWh)	15.95	13.53	12.89	13.53	13.87	13.91	13.98	14.07
S.D. AC ($/MWh)	7.83	6.53	6.1	6.53	6.6	6.68	6.8	6.97
Avg. inefficiency ξ		3.02	3.06	3.02	2.94	2.93	2.92	2.92
S.D. inefficiency ξ		1.12	1.1	1.12	1.16	1.18	1.21	1.25
Avg. ineff. Buyer		3.11	3.26	3.11	3.1	3.1	3.09	3.08
Avg. ineff. Seller		2.84	2.33	2.84	3.16	3.21	3.22	3.21
Mean bid (mill. $)		-10.44	-1.2	-10.44	-1.19	6	18.21	40.07
Median bid (mill. $)		-2.05	-2.46	-2.05	0.64	2.63	6.42	12.69
S.D. bid (mill. $)		56.06	5.34	56.06	39.06	26.51	30.42	86.5
Mean net transfer (mill. $)		0.02	0.97	0.02	-15.25	-29.81	-56.36	-105.21
S.D. net transfer (mill. $)		0.08	1.34	0.08	35.35	68.46	131.04	250.28
Combinations tested		7,637	7,570	7,637	7,676	7,731	7,825	7,919
Nbr remaining LDCs	64	61.64	69.92	61.64	58.88	57.04	53.36	49.68

Notes: Averages are non-weighted. We used 2003-2016 for estimation of AC. All years pooled for calibration. BAU is public to private taxed at 22% and public to public at 0%. Prop. tax is an elimination of the 22% tax levied in 2016 between public and private LDCs. Subsidy X% is a negative transfer tax of X% (X=25, 50, 100, and 200). Tax 25% imposes a transfer tax of 25%.

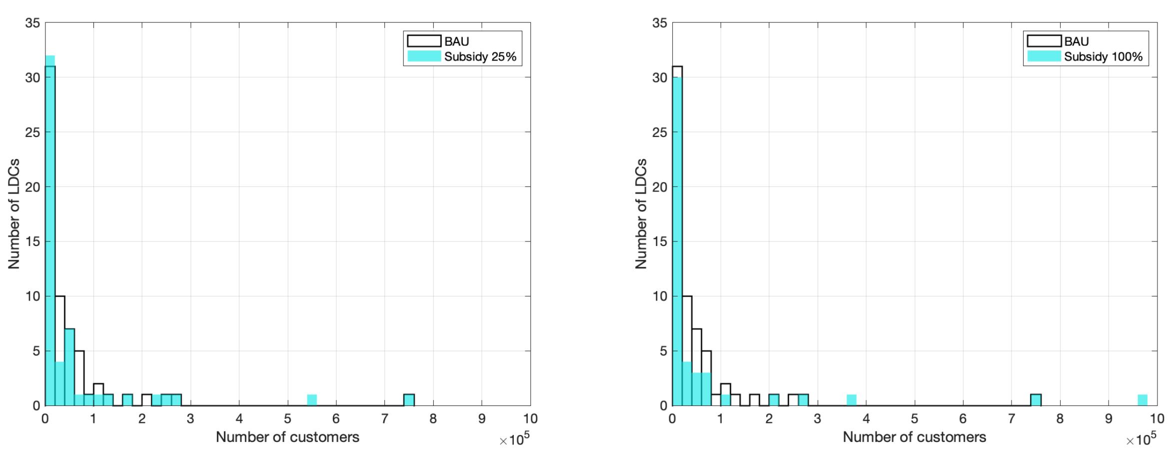
are comparable to the result in Kwoka and Pollitt [2010] that shows that inefficient LDCs buy more efficient ones. Our model, however, allows for varying returns to scale by allowing different levels of curvature in the average cost curve. In Kwoka and Pollitt [2010], their methodology implicitly assumes constant returns to scale, which eliminates any incentives to limit growth in size.

Recall from above that bids are negative in the BAU reflecting the fact that many of the sellers have negative profits, some of them significantly so. Not surprisingly, as the transfer tax decreases and the merger subsidy increases, mean and median bids become positive and progressively larger.

Finally, acquisitions occurred at transaction prices of -\$10.44 million in the BAU. This is because π_{ij} is small compared to the draw of s_{ij}. This negative transfer falls to -\$1.19 million for a 25% subsidy, and then becomes positive for bigger subsidies.

Figure 8: Distribution of LDC sizes under BAU and Subsidies 25% and 100%

(a) Subsidy 25% (b) Subsidy 100%



Notes: Comparison of the LDC sizes distributions between the model BAU's predictions and the outcome from two different subsidy levels.

Next we look more carefully at two of the experiments to understand more about the acquisition patterns. We focus on Subsidy 25% and Subsidy 100%, but results from the other experiments are available upon request. In Figure 8 we present histograms of the distribution of the size of the LDCs at the baseline against the distribution under the counterfactual scenarios of 25% and 100% subsidies. Not surprisingly, since there are few conglomerates formed, there is very little change in the distributions. We see more impact at 100%, but even there we see just a slight shift towards larger LDCs.

Overall the findings from our experiments suggest that it would require a huge subsidy to generate anywhere near the sort of consolidation hoped for by the government panel. The fact that there are not more mergers can be explained by the shape of the AC curve. The AC curve would have to have been much more convex in order for larger economies of scale to be achieved through consolidation.

7 Conclusions

This paper proposes a method to simulate mergers in the electricity distribution industry that endogenizes the merger process with a sequential set of offers determined via Nash bargaining. The method is easily computable, even when the number of firms is large, and it could be used to guide policy recommendations regarding industry consolidation initiatives advanced by political authorities, not just for the electricity distribution industry, but other industries as well where similar forces are at play.

We apply our method to analyze the proposed restructuring of Ontario's electricity distribution market. Our findings suggest that the proposed tax incentives meant to encourage consolidation are insufficient to achieve the desired outcomes. In fact, even large subsidies would have little effect on industry structure. Moreover, our results suggest that consolidation would not have the hoped-for impact of lowering average cost of production across the sector.

References

Ashenfelter, O., Hosken, D., and Weinberg, M. (2015). Efficiencies brewed: Pricing and consolidation in the U.S. beer industry. *RAND Journal of Economics*, 46:328–361.

Becker-Blease, J., Goldberg, L., and Kaen, F. (2008). Mergers and acquisitions as a response to the deregulation of the electric power industry: Value creation or value destruction? *Journal of Regulatory Economics*, 33(1):21–53.

Bloom, N., Propper, C., Seiler, S., and Reenen, J. V. (2015). The impact of competition on management quality: Evidence from public hospitals. *Review of Economic Studies*, 82:457–489.

Brasington, D. (1999). Joint provision of public goods: The consolidation of school districts. *Journal of Public Economics*, 73:373–393.

Brown, D. and Sappington, D. E. (2017). Optimal policies to promote efficient distributed generation of electricity. *Journal of Regulatory Economics*, 52:159–188.

Collard-Wexler, A. (2014). Mergers and sunk costs: An application to the ready-mix concrete industry. *American Economic Journal: Micro*, 6(407-447).

Deneckere, R. and Davidison, C. (1985). Incentives to form coalitions with bertrand competition. *RAND Journal of Economics*, 16:473–486.

Dimitropoulo, D. and Yatchew, A. (2017). Is productivity growth in electricity distribution negative? An empirical analysis using ontario data. *The Energy Journal*, 38:175–200.

Einav, L. and Levin, J. (2010). Empirical industrial 0rganization: A progress report. *Journal of Economics Perspectives*.

Eliason, P., Heebsh, B., McDevitt, R., and Roberts, J. (2020). How acquisitions affect firm behavior and performance: Evidence from the dialysis industry. *Quarterly Journal of Economics*, 135:221–267.

Elston, M., Laughren, F., and McFadden, D. (2012). Renewing Ontario's electricity distribution sector: Putting the consumer first. http://www.energy.gov.on.ca/en/ldc-panel/. last accessed January 31, 2013.

Fan, Y. and Yang, C. (2020). Competition, product proliferation and welfare: A study of the U.S. smartphone market. *American Economic Journal: Microeconomics*, 12:99–134.

Fee, C. and Thomas, S. (2004). Sources of gains in horizontal mergers: Evidence from customer, supplier, and rival firms. *Journal of Financial Economics*, 74:423–460.

Focarelli, D. and Panetta, F. (2003). Are mergers beneficial to consumers? Evidence from the market for bank deposits. *American Economic Review*, 93:1152–1172.

Fyfe, S., Garner, M., and Vegh, G. (2013). Mergers by choice, not edict: Reforming Ontario's electricity distribution policy. http://www.cdhowe.org. last accessed March, 2013.

Gaynor, M., Laudicella, M., and Propper, C. (2012). Can governments do it better? Merger mania and hospital outcomes in the English NHS. *Journal of Health Economics*, 31:528–543.

Gordon, N. and Knight, B. (2009). A spatial merger estimator with an application to school district consolidation. *Journal of Public Economics*, 93:752–765.

Gowrisankaran, G. (1999). A dynamic model of endogenous horizontal mergers. *RAND Journal of Economics*, 30:56–83.

Gowrisankaran, G., Reynolds, S., and Samano, M. (2016). Intermittency and the value of renewable energy. *Journal of Political Economy*, 124:1187–1234.

Greene, W. H. (2008). The econometric approach to efficiency analysis. *The measurement of productive efficiency and productivity growth*, 1(1):92–250.

Harford, J. (2005). What drives merger waves? *Journal of Financial Economics*, 77:529–560.

Harman, G. and Harman, K. (2003). Institutional mergers in higher education: Lessons from international experience institutional mergers in higher education: Lessons from international experience. *Tertiary Education and Management*, 9:29–44.

Harrison, T. (2006). Hospital mergers: Who merges with whom? *Applied Economics*, 38:637–647.

Hess, B. (2007). Do mergers really increase efficiency? A cost efficiency analysis of electricity distributors in the US.

Hollenbeck, B. (2020). Horizontal mergers and innovation in concentrated industries.

Hosken, D. and Weinberg, M. (2013). Evidence on the accuracy of merger simulations. *Review of Economics and Statistics*, 95:1584–1600.

Igami, M. and Uetake, K. (2019). Mergers, innovation, and entry-exit dynamics: Consolidation of the hard disk drive industry, 1996-2016. forthcoming *Review of Economic Studies*.

Jeziorski, P. (2014). Estimation of cost efficiencies from mergers: Application to U.S. radio. *RAND Journal of Economics*, 45:816–846.

Joskow, P. (2000). Deregulation and regulatory reform in the us electric power sector. In Peltzman, S. and Winston, C., editors, *Deregulation of Network Industries*. Brookings, Washington, DC.

Kamien, M. and Zang, I. (1990). The limits of monopolization through acquisition. *Quarterly Journal of Economics*, 105:465–499.

Knittel, C. R. (2002). Alternative regulatory methods and firm efficiency: Stochastic frontier evidence from the U.S. electricity industry. *The Review of Economics and Statistics*, 84(3):530–540.

Kumbhakar, S. C. and Lovell, C. K. (2003). *Stochastic frontier analysis*. Cambridge University Press.

Kwoka, J. (2005). Electric power distribution: economies of scale, mergers, and restructuring. *Applied Economics*, 37:2373–2386.

Kwoka, J. and Pollitt, M. (2010). Do mergers improve efficiency? Evidence from restructuring the U.S. electric power sector. *International Journal of Industrial Organization*, 28:645–656.

Lim, C. and Yurukoglu, A. (2016). Dynamic natural monopoly regulation: Time inconsistency, moral hazard, and political environments. forthcoming, *Journal of Political Economy*.

Maksimovic, V., Phillips, G., and Prabhala, N. (2011). Post-merger restructuring and the boundaries of the firm. *Journal of Financial Economics*, 102:317–343.

Mermelstein, B., Nocke, V., Satterthwaite, M., and Whinston, M. (2020). Internal vs. external growth in industries with scale economies: A computational model of optimal merger policy. *Jounal of Political Economy*, 128:301–341.

Mitchell, M. and Mulherin, J. H. (1996). The impact of industry shocks on takeover and restructuring activity. *Journal of Financial Economics*, 41:193–229.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *RAND Journal of Economics*, 49:627–649.

Okazaki, T., Onishi, K., and Wakamori, N. (2019). Compatible Mergers: Assets, Service Areas, and Market Power. CIRJE F-Series CIRJE-F-1134, CIRJE, Faculty of Economics, University of Tokyo.

Ontario Energy Board (2015). Report of the board: Rate-making associated with distributor consolidation. Technical report.

Ontario Energy Board (2016). Handbook to electricity distributor and transmitter consolidations. Technical report.

Park, M. and Town, R. (2014). Industry shock expectations, interindustry linkages, and merger waves: Evidence from the hospital industry. *Journal of Economics and Management Strategy*, 23:548–567.

Perry, M. and Porter, R. (1985). Oligopoly and the incentive for horizontal merger. *American Economic Review*, 75:219–227.

Pesendorfer, M. (2003). Horizontal mergers in the paper industry. *RAND Journal of Economics*, 34:495–515.

Peters, C. (2006). Evaluating the performance of merger simulation: Evidence from the u.s. airline industry. *Journal of Law and Economics*, 49:627–649.

Qiu, L. and Zhou, W. (2007). Merger waves: A model of endogenous mergers. *RAND Journal of Economics*, 38:214–226.

Saarimaa, T. and Tukiainen, J. (2014). I don't care to belong to any club that will have me as a member: Empirical analysis of municipal mergers. *Political Science Research Methods*, 2:97–117.

Seim, K. and Waldfogel, J. (2013). Public monopoly and economic efficiency: Evidence from the pennsylvania liquor control board's entry decisions. *American Economic Review*, 103(2):831–62.

Stahl, J. (2016). Effects of deregulation and consolidation of the broadcast television industry. *American Economic Review*, 106:2185–2218.

Weese, E. (2015). Political mergers as coalition formation: An analysis of the heisei municipal amalgamations. *Quantitative Economics*, 6:257–307.

Yatchew, A. (2000). Scale economies in electrical distribution: A semiparametric analysis. *Journal of Applied Econometrics*, 15(2):187–210.

Appendix: Additional tables and figures

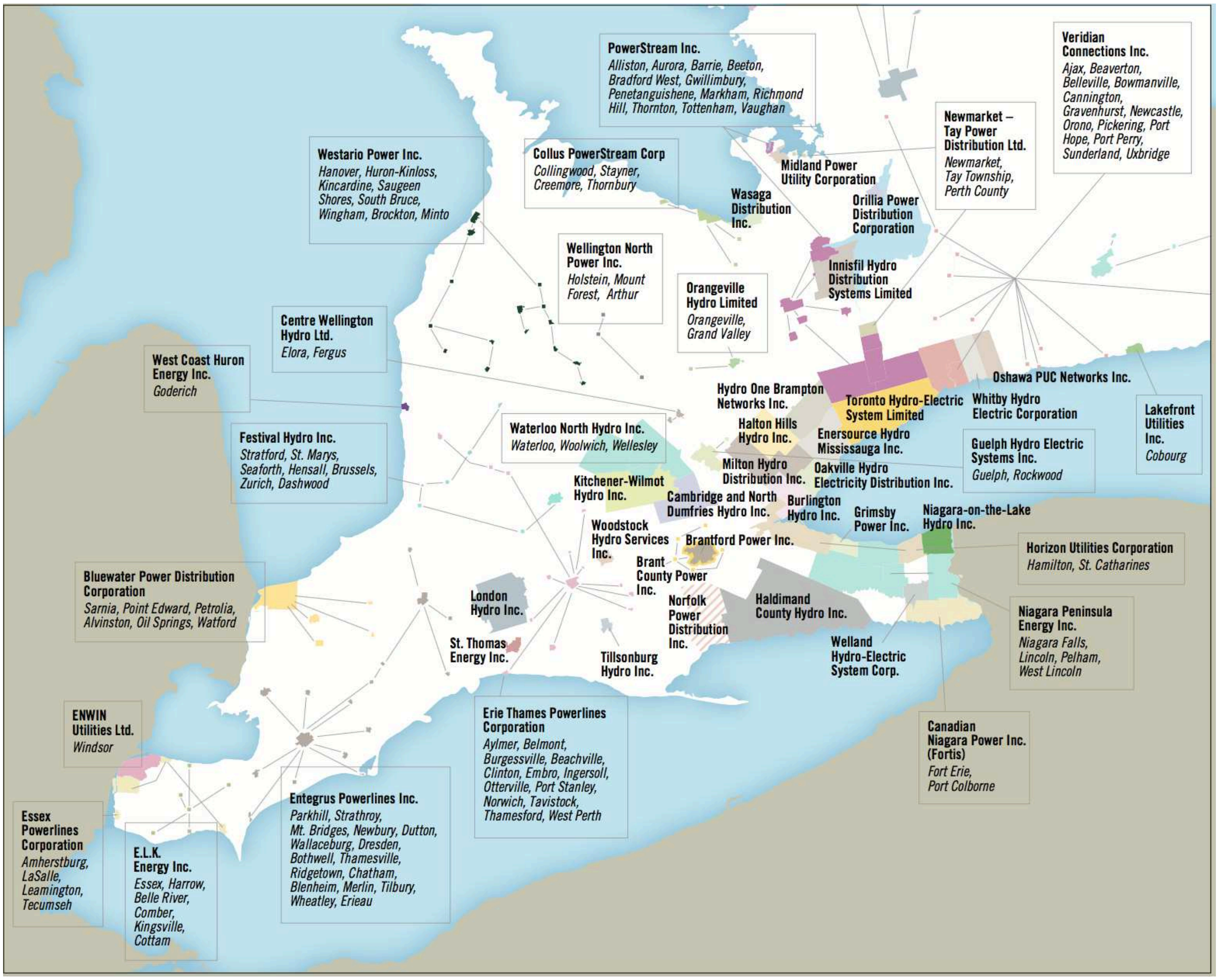
Table A.1: O&M and Administrative Costs SFA

	(1)	(2)	(3)	(4)	(5)	(6)
log q	0.948***	0.920***	0.908***	0.823***	0.786***	0.760***
	(0.0318)	(0.0314)	(0.0375)	(0.0213)	(0.0213)	(0.0263)
log density	-0.430***	-0.654***	-0.653***	-0.163**	-0.404***	-0.329***
	(0.0770)	(0.0865)	(0.102)	(0.0550)	(0.0619)	(0.0766)
log price capital		0.277***	0.269***		0.326***	0.307***
		(0.0547)	(0.0565)		(0.0429)	(0.0436)
log frac. urban area			-0.00379			-0.0685
			(0.0507)			(0.0368)
log frac. overhead lines			-0.101			-0.175
			(0.162)			(0.122)
constant	1.718	-0.0945	0.0891	4.011***	1.572***	1.805***
	(2.369)	(1.636)	(1.537)	(0.331)	(0.473)	(0.505)
log σ_ν^2	-1.589***	-1.647***	-1.623***	-2.381***	-2.421***	-2.345***
	(0.147)	(0.146)	(0.153)	(0.151)	(0.145)	(0.165)
inv. logit γ	1.818***	1.832***	1.870***	1.269***	1.448***	1.581***
	(0.193)	(0.191)	(0.202)	(0.215)	(0.201)	(0.224)
μ	1.845	1.716	1.720	0.725***	0.824***	0.747***
	(2.334)	(1.547)	(1.413)	(0.173)	(0.212)	(0.135)
N	369	369	369	370	370	370

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

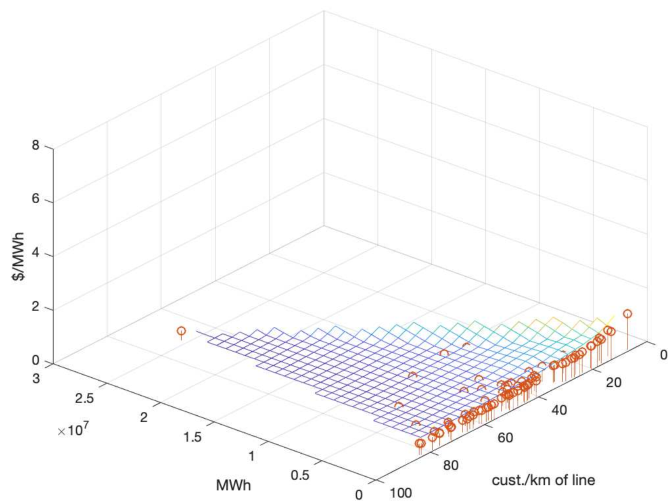
Notes: O&M costs and admin costs only. Due to the data availabilities, estimates for year 2009 and 2011 - 2014 only. Dependent variable for columns (1)-(3): log O&M costs. Dependent variable for columns (4)-(6): log administrative costs. Estimation includes Hydro One Networks Inc.

Figure A.1: Map of Ontario's LDCs (2014)



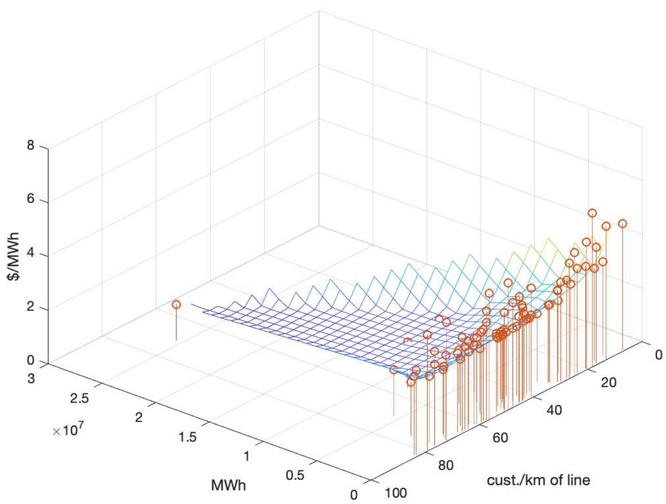
Source: IESO http://www.ieso.ca/-/media/files/ieso/document-library/maps/ontario-ldc-map.pdf.

Figure A.2: Average cost curve for O&M costs (2009)



Notes: Average cost curve for O&M costs as a function of q_i and line density. Data for 2009 since costs data are not available by categories before 2009. Circles represent actual data points. Surface is the interpolated average cost curve.

Figure A.3: Average cost curve for administrative costs (2009)



Notes: Average cost curve for administrative costs as a function of q_i and line density. Data for 2009 since costs data are not available by categories before 2009. Circles represent actual data points. Surface is the interpolated average cost curve.