Procuring Medical Devices: The Price Effect of Mergers among Orthopedic Prostheses Producers

Vincenzo Atella[†], Nicola Ceschin[§] and Francesco Decarolis[§],

[†]*CEIS Tor Vergata and DEF, University of Rome Tor Vergata*

[§]*Department of Economics and Baffi-Carefin, Bocconi University* [*]

October 2, 2021

Abstract

The efficacy of public procurement depends on the market structure. In this study, we exploit a change in market structure arising from a merger of two, large US-based manufacturers of orthopedic prostheses to study how this affected the prices paid by foreign public buyers. Using data on all public procurement events in Italy over the period 2012 to 2019, we set up a difference-in-difference model comparing the winning prices for orthopedic prostheses with those for other devices not subject to changes in seller concentration. The findings show that the merger caused an increase in prices of the order of 7%, corresponding to €190 million over the period 2015-2019.

JEL classification: I18, J18, C21

Keywords: procurement, medical devices, orthopedic prosthesis, consolidation.

[*]Vincenzo Atella, CEIS Tor Vergata and DEF, University of Rome Tor Vergata, Via Columbia, 2, 00133 Roma, e-mail: atella@uniroma2.it. Nicola Ceschin, Bocconi University, Economics Department, Via Roentgen 1, 20100 Milan, e-mail: nicola.ceschin@studbocconi.it. Fancesco Decarolis, Bocconi University, Economics Department, Via Roentgen 1, 20100 Milan, e-mail: francesco.decarolis@unibocconi.it. We would like to thank Chiara Motta, Andrey Tkachenko and Fabiola Zarrelli for skillful research assistance. We also thankfully acknowledge financial support from both the European Research Council (ERC Starting Grant 679217, Decarolis). The views expressed in this paper are those of the authors and do not imply any responsibility of their institutions.

Contents

1 Introduction

Governments spend roughly a third of total general government spending on public procurement and, among OECD countries, health is the second largest spending area (over 9% of GDP in 2019). The ongoing pandemic is bound to further increase this figure and, therefore, understanding the drivers and consequences of healthcare procurement is fundamental for the design of sound procurement practices. Industry concentration has led to important changes in how the public procurement system works. In the case of military procurement by the US Department of Defense, Carril and Duggan (2020) show that the sharp increase in concentration among defense contractors in the 1990s led to the adoption of more flexible contract arrangements and procurement practices that included awarding contracts without competition, or via single-bid solicitations. Thanks to these flexible practices, industry consolidation did not drive higher procurement costs.

But what is the expected outcome when public procurement follows rigid regulations that don't allow the introduction of flexible practices? This is an important question because, according to recent data by the World Bank, most countries have more rigid procurement laws than the US (Bosio et al., 2020). This study seeks to analyze how such a rigid procurement systems interact with industry concentration. To achieve this goal, we focus on the case of Italy, which has one of the most rigid procurement regulations according to (Bosio et al., 2020),[1] and exploit for estimation an exogenous change in concentration affecting the orthopedic prostheses market.

Medical devices (MDs) represent a large share of public health expenditure and orthopedic prostheses represent in Italy by far the largest item.[2] As reported by the Italian Ministry of Health (2018), in 2017 the National Health System (NHS) spent €427 million to satisfy the national demand for orthopedic implants, 27% more than in 2013 and almost €100 million more than the second expenditure item in the ranking, cardiac devices. Thus, their weight on Italian public accounts is sizeable. It is also representative of several trends happening in the healthcare sector and affecting suppliers, public buyers and final patients.

Starting from the latter, the number of patients that underwent these procedures has increased steadily over the last decade.[3] Phenomena such as population aging and growing obesity bolster this trend and, indeed, recent projections (Romanini et al., 2019) forecast an increase in the incidence rate of total knee arthroplasty (the second most common orthopedic procedure) of 45% by 2050 compared to 2017. On the public buyer side, orthopedic supplies have been at the heart of several policy reforms that attempted to con-

[1] For the index summarizing the extent of the regulations and varying between 0 and 4, Bosio et al. (2020) report that the countries with more regulation are Greece, Rwanda, Italy and Latvia (all with scores ranging between 3 and 2.92), while Israel is at the opposite end of the spectrum with a score of 0.31. The US is roughly halfway between these two polar extremes.

[2] At the European level, medical device expenses achieved 0.6% of GDP in 2018 (MedTech Europe, 2020). Over the last decade, the European medical device market has been growing on average by 4.2% per year, reaching a size of almost €110 billion. In the next years, this industry is forecast to continue to grow at a sustained rate of more than 5%. Italy appears as the fourth largest market in Europe, accounting for approximately 10% of total sales, behind Germany, France and UK.

[3] For hip replacement, the most common surgery in this category, the number of people hospitalized each year to receive these implants has grown by approximately 2.5% per year, reaching more than 100,000 individuals in 2018. For arthroplasties linked to partial or total knee and shoulder replacement, the average growth rate over this period was even higher, 4% and 11% respectively (AGENAS, 2018).

tain health expenditure growth. This process culminated in 2016 with the introduction of hip prostheses in the list of medical devices that could have no longer been purchased at the local level by hospitals and Local Health Authorities (LHAs). Since then the only channel allowed to procure these items were tenders issued by national or regional Central Purchasing Bodies (CPBs). At the same time, in the last decade the focus of public procurement all around Europe has shifted from *cost-saving* to the so-called *value-based procurement*, with the focus of the procurement process shifting from price to quality. Lastly, on the private supplier side, in the last decade the industry has witnessed an unprecedented wave of consolidation. Among the major events, there were the acquisition of Synthes Inc. by Johnson & Johnson in 2012, the acquisition of Biomet Inc. by Zimmer Holdings Inc. in 2015, and, lastly, the acquisition of Wright by Stryker Corp. in 2020. Although all of these firms are US based and operated, they are global leaders and these mergers triggered an increase in concentration in many markets across the globe.

The main goal of this study is to exploit the variation in industry concentration triggered by the largest of the three mergers, that between Biomet and Zimmer, to study how it impacted procurement costs for the Italian public buyers. This acquisition is particularly interesting for at least four reasons. First, it was one of the biggest deal in terms of the size of parties involved. Second, the leadership of Zimmer and Biomet was not limited to one or few kinds of orthopedic prostheses: in 2014, these two multinational corporations were among the five largest suppliers in virtually all of the prostheses markets. Third, this procedure has been challenged by European authorities and eventually approved with remedies. As noted by Ashenfelter et al. (2013), being on the edge between approval and rejection, the evaluation of such a "marginal merger" is of particular interest as it would highlight whether European competition authorities set their approval bar correctly or were too permissive. Fourth, detailed data are available for the period surrounding this merger. In particular, from 2013 onward all the purchases of MD by hospitals and LHA were recorded and periodically publicly released. This allows us, for example, to track closely the evolution of the market shares of each producer.

The two fundamental ingredients of the empirical analysis are a novel dataset and a causal identification strategy. The former is the result of the combination of three main sources of data that we use in order to measure procurement prices at a granular (i.e., contract) level. The latter is based on a differences-in-differences strategy that accounts for both the very peculiar distribution of the outcome variable and the availability of data for multiple medical devices to determine an adequate control group.

The main results of the analysis suggest that the merger caused a decrease in auction rebates (thus, an increase in prices). This price effect is of the order of 7%, corresponding to €190 million over the period 2015-2019. Although the price increase occurred for all of the types of prostheses considered, the effect was heterogeneous across product markets ranging from 7% for trauma implants to 16% for spine implants. The results are robust to a wide range of model specifications and sample definitions.

These results underscore that while both the public debate and the academic research have often focused on issues such as corruption, cartels and incomplete contracts, the role of supplier market power in public procurement markets is highly relevant. Moreover, our results call for particular attention in the evaluation by antitrust authorities of mergers involving public procurement markets, especially when the rigidity of the regulations limit the flexibility of the public buyers to respond effectively to supplier concentration. In the conclusions, we return to these policy considerations and elaborate on the need to better coordinate public procurement and antitrust policies.

Literature - There are three main streams in the literature to which this paper contributes. The first is that on procurement outcomes and, specifically, how they depend on market structure. The second involves the evaluation of consummated mergers. The last is the more specific analysis of competition in the medical device industry.

Regarding the contribution to the literature on public procurement and market structures, we have already explained how our work relates to the recent study of Carril and Duggan (2020) on US military procurement. In recent years, the interest in industry concentration and its effects has been mounting (see Benkard et al. (2021) for a recent review of the literature). A main obstacle to the understanding of the effects of concentration on prices and other economic outcomes is the non-random variation in market structure. In our study, however, we do not face this problem and we can safely claim that the merger Zimmer-Biomet was fully exogenous relative to the conditions of the Italian procurement market. Namely, the merger was not motivated by the specific features of the Italian prostheses market, as Italy contributes only a tiny fraction of the total revenues of these companies.[4]

Regarding the literature on consummated mergers, our usage of a DID method is fairly standard (see, among others, Ashenfelter and Hosken (2010) and Ashenfelter et al. (2013)). The key challenge that we face is to define a control group that is comparable to the treated one, and contemporaneously is not directly affected by the merger. This second requirement prevents us from using as controls the data from products substitutable with those of the merging parties since, at equilibrium, their price should respond to the merger (Deneckere and Davidson, 1985). Since the geographical market for orthopedic prostheses is national, we also cannot exploit spatial variation. Therefore, we rely on a comparison group containing a different kind of medical devices in the same national market. In particular, after a thorough analysis as outlined in section 4.2, the most comparable products to orthopedic supplies turned out to be pacemakers and implantable defibrillators.

Finally, regarding competition in the medical device industry an interesting analysis of how market dynamics influence the prices of these items was carried out by Grennan (2013) who investigated the effect of price discrimination and bargaining ability in a business-to-business market where healthcare providers directly buy MD from manufacturers. Closer to our case, Bucciol et al. (2020) analyzed the role of buyer characteristics in the Italian procurement market for MDs, where public administrations purchases goods and services through auctions on the behalf of healthcare providers. Using unit prices for a set of different MDs, they estimated a structural model which revealed that much of the variation in prices at which the public sector buys these products is due to the level of "competence" in procuring of the purchasing administration. This in turn is related with its institutional characteristics and size (in terms of procured turnover). We will incorporate their results in our identification strategy, controlling for buyer competence using a set of covariates that is very similar to what they used.

Other related studies include Rizzo et al. (2006) on the relationship between procurement prices and supplier concentration. Using data on tenders for medical devices in Italy for the period 1995-2005, they proxied market concentration with the total number of registered suppliers in different classes of MDs and with the actual number of producers participating in a smaller set of auctions. They estimated that a larger number of

[4]In the year before the merger (2014) France, Germany, Italy, Spain, Switzerland and the United Kingdom collectively accounted for only 17% of total world sales of Zimmer, by far the largest of the two merging parties (Zimmer Inc., 2015).

firms participating in a procurement auction is generally associated with lower awarding prices. Vellez (2012) improved on these estimates by controlling for more covariates at the contract and the procurement agency level and adopting an IV strategy to account for endogeneity in firms' participation. Although our research question is similar to that in these two studies, compared to their work our study benefits from a more granular data and a causal identification strategy exploiting the merger event.

2 The institutional setting and the merger

2.1 Public procurement of MDs in Italy

The Italian National Healthcare System (NHS, hereafter) is a tax-funded system, free of charge for everyone, regardless of income. As laid down in the D. Lgs. 502/92, healthcare is a policy area managed at the local level, by regions. The provision of health services is granted by Local Health Authorities (LHAs), regional institutions with an independent legal status, and by other entities that may control one or more Hospitals (Presidio Ospedaliero), whose accounts are included in the LHAs' balance sheets. Beside the LHAs, there are other entities involved in the supply of health services that retain independence in their management, balance sheets and procurement process, which we will define "Local Purchasing Authorities" (LPAs), since they are all subjects entitled to launch calls for tenders to buy the health supplies they need.[5]

This decentralized system entails potential inefficiency, as the small size of the LPAs does not allow exploitation of, for example, economies of scale and scope. For these reasons, beginning in the early 2000s regional and national governments have increasingly promoted initiatives to foster the centralization of procurement. One of the first was the introduction by Law 311/2004 of the "Piani di Rientro", national programs to restore budget balance in targeted regions, which included specific provisions on centralization. Moreover, by adopting the EU Directive 18/2004, implemented in Italy by the Legislative Decree n. 163/2006, regions were forced to adopt "Central Purchasing Bodies" (CPBs), regional agencies with the aim of procurement on the behalf of all local public administrations, including LPAs. However, in practice the responses of regions were very heterogeneous and LPAs continued to play a key role in the acquisitions of health-related goods. Finally, the DPCM 24/12/2015 was the real turning point on the path to public procurement centralization, as it provided for the first time a list of items the acquisition of which was mandated to be via CPBs. The majority of them were MDs, chosen for their high budgetary impact and the relative low degree of complexity. The list included, among others, hip prostheses, pacemakers and defibrillators.[6] Thanks to the level of detail of our data, we will be able to control for these legislative changes in our analysis.

The principles through which procurement auctions are executed in Italy are laid down in the above-mentioned EU Directive 18/2004, lately modified by EU Directive

[5] Among these LPAs we have Hospital Enterprises (Azienda Ospedaliera), hospitals turned into independent units; the Institutes for Treatment and Research (Istituto di Ricovero e Cura a Carattere Scientifico), hospitals focused on specific treatments and on research; University Hospitals Enterprises (Azienda Ospedaliera Universitaria), which combine the activities of a hospital with the training of young physicians.

[6] The full list is available at the following web address: https://www.gazzettaufficiale.it/eli/id/2016/02/09/16A00583/sg.

25/2014 and transposed into the Italian Legislative Decree 50/2016, the "Code of Public contracts of works, services and supplies". According to these provisions, the main dimensions along which a public auction can vary are in the awarding procedure, the awarding criterion and the contractual form. The awarding procedure is the process through which the contracting authority organizes the call for tenders. Simplifying the process a bit, the main types of procedures are *i)* open, where all of the interested supplier are allowed to submit their bids, *ii)* restricted, where before presenting their offers, companies are checked to ensure they satisfy some prerequisites, *iii)* negotiated, in which the public administration can select specific companies with which negotiate directly the terms of the supply contract, *iv)* competitive dialogue, where suppliers are involved in the definition of the characteristics of the goods and services to be purchased. On the other hand, the awarding criterion is the method through which the public administration selects the winning bid. This can be either the *i)* lowest price, where the winning bid is the one offering the higher rebate on the reserve price, or the *ii)* most advantageous economic offer (MEAT), where the contract is awarded to the bid achieving the highest score, with this being a weighted average of the bid's rebate and of the quality of the goods offered. In particular, following the provisions laid down in the EU Directives mentioned above, from 2017 (D.Lgs. n. 56/2017) the MEAT criterion has become the default criterion in public procurement auctions and the ceiling value of the weight attached to price in the computations of a bid's score has been set at 30%.

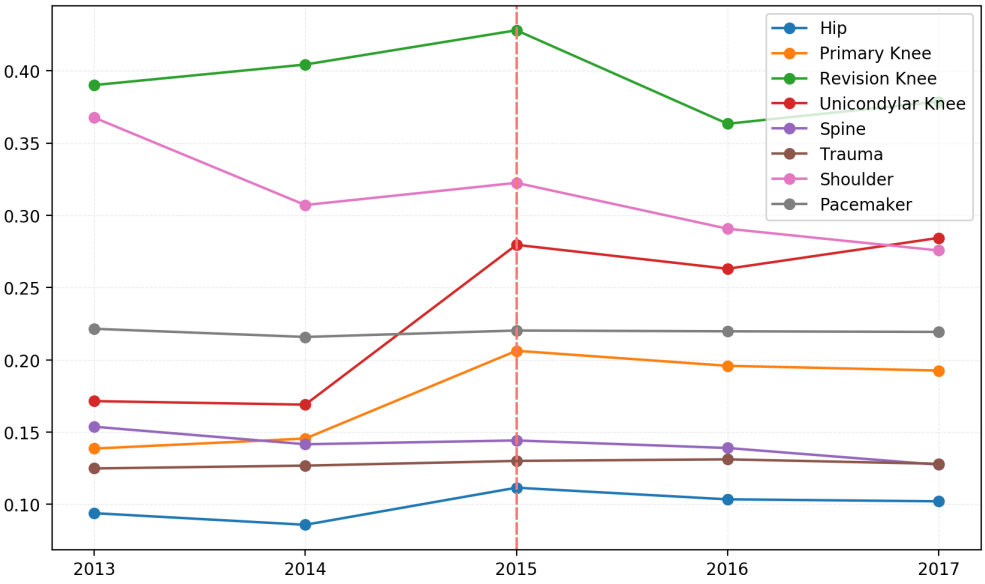
Following Decarolis and Giorgiantonio (2015), we group procurement auction formats into three, more general, categories, according to the combination of awarding criterion and procedure used. These are: *i)* First Price auctions (FP), where the lowest price criterion is combined with either open or restricted procedure; *ii)* Scoring Rule auctions (SR), where the MEAT criterion is combined with either open or restricted procedure, *iii)* Negotiations (N), that gathers all kinds of negotiated procedures and competitive dialogues, irrespective of the awarding criterion. Despite public administrators in charge of the call for tenders having some flexibility in choosing the auction format, this flexibility is very limited if compared to international standards.[7] In fact, in Italy, the main factors determining the structure of the call are the object of the purchase and the expected costs.

2.2 The merger

In April 2014, Zimmer agreed to buy Biomet for $13.4 billion to become the second-largest player in the rebounding market for treating muscle and orthopedic injuries. Zimmer was a US publicly owned company, global leader in the development and manufacturing of orthopedic devices. Biomet, relatively smaller in terms of annual turnover (€3.5 billion against €2.3 billion), was one of its closest competitors in this industry. These companies were among the market leaders both in the US and in the European market and the Federal Trade Commission (FTC) collaborated extensively with the EC in evaluating the impact of this merger on competition. In July 2014, the FTC requested additional information from the merging parties and eventually concluded that the deal would have raised substantially the market power in the markets for unicondylar knee implants, total elbow implants and bone cement. Eventually, the merger was approved in the summer of 2015 upon a package of divestments involving some of Zimmer and Biomet's most popu-

[7] As already mentioned in the Introduction, according to a survey conducted by the World Bank (Bosio et al., 2020) Italy ranks third worldwide in terms of the extent of its procurement regulations.

Figure 1: HHI of specific prostheses markets



Note: HHI built on the basis of expenditures data by Italian NHS, extracted from the *Consumptions Flows* (Flusso Informativo sui consumi) available on the website of the Italian Ministry of Health (see details in Appendix A). The CND codes (Classificazione Nazionale dei Dispositivi Medici) corresponding to the markets of implants represented are: P0908 for hip, P0901 for shoulder, P090903 for primary knee, P090904 for unicondylar knee, P090905 for revision knee, P0907 for spine, P0912 for trauma, J0105 and J010 for pacemakers and defibrillators (averaged and labelled as "pacemaker" in the plot). The red vertical dashed line indicates the year of the merger.

lar products in these three markets. In the EU, the approval procedure followed a similar path. Initially, the EC expressed its concerns and in October 2014 it opened a Phase II in-depth investigation.

The analysis of the EC revealed that the impact of the merger could potentially have been extremely heterogeneous depending on the specific market affected. While the relevant geographical market was assessed to be national, a total of 13 product markets were identified by the EC: elbow implants, hip implants, shoulder implants, bone cement, bone cement accessories, surgical tools (in particular, "pulsed lavage"), spine devices, trauma devices, dental implants and knee implants, which in turn were divided into 4 other specific product markets (patello-femoral knee implants, unicondylar knee implants, primary knee implants and revision knee implants). The concerns of the EC about an excessive dominant position of the new firm concentrate exactly around these latter markets. With respect to Italy, while the majority of markets were deemed to stay competitive, the Commission concluded that the merger would have "significantly impeded competition" in the market for unicondylar knee prostheses (European Commission, 2015).

To get a better understanding of the heterogeneous impact that the merger had on some of the markets involved, we can refer to Figure 1, which describes the evolution of concentration (measured via the Herfindahl-Hirschman Index, HHI) around the year of the merger (2015). As we see, the largest deviation in HHI corresponds to the market for unicondylar knee prostheses, consistent with the EC's concerns. This market shifts from being moderately concentrated (around 0.17) to being highly concentrated (more than 0.27), with the new entity gaining a market share of almost 50%. Other markets

witness a relatively smaller growth in HHI (hip and primary knee prostheses); some others are almost untouched (trauma and spine devices); finally, two markets (revision knee and shoulder implants) are only marginally affected by the merger, suffering just from a temporary deviation in a decreasing trend at extremely high levels of concentration.

The acquisition was eventually approved by the EC on February 9, 2015, conditional upon a commitments package submitted by Zimmer. In particular, Zimmer offered to divest a popular unicondylar knee implant and Biomet one of its elbow prostheses, both of them at the European level. Some other targeted divestments were required in the Danish and Swedish markets. Being cleared with remedies, we can claim this acquisition was on the edge of the approval bar for the EC (a "marginal merger", in the spirit of Ashenfelter et al. (2013)) and thus it represents a particularly interesting case to study *ex post*.

As a final note, it is important to remark that consolidation in the orthopedic prostheses industry is not an exception. The MD sector has indeed faced an outstanding wave of mergers and acquisitions in the last years (Crotti, 2017), the main examples of which are the $42 billion deal between Medtronic and Covidien in 2014 and the $25 billion acquisition of St. Jude Medical by Abbott in 2017. This brings more evidence to support the exogeneity of the Zimmer-Biomet merger to the Italian market, as consolidation looks like a phenomenon which affects virtually every sector in the MD industry. In this regard, the Italian markets for pacemakers and implantable defibrillators (pacemakers, hereafter) are outliers. As shown in Figure 1, these two markets have witnessed a substantial stability in concentration in the last decade, with no consolidation episodes involving close rivals and roughly constant market shares of the industry players. The gray line in Figure 1 plots the average HHI of these two markets. As will be better explained in section 4.2, this will be one of the criteria considered to pick these two markets as the preferred control group in our DID model.

3 Data

3.1 Main data sources and market definition

The main source of data used in this research is the National Database for Public Contracts ("Banca Dati Nazionale dei Contratti Pubblici", BDNCP), collected by the Italian National Anti-corruption Authority (ANAC). The dataset contains information on the universe of the calls for tenders issued by Italian public administrations whose reserve price exceeds the threshold of €40,000, for the period 2012-2019. Observation units are lots, which are tendered through auctions where one or more lots are procured. The lots included in our dataset were selected based on the Common Procurement Vocabulary (CPV) codes for orthopedic prostheses (treated) and cardiac stimulation devices (control).

Combining the CPV codes with the description of the lot, we were able to assign each observation to a specific product market, according to the market definition contained in the Zimmer-Biomet (European Commission, 2015) and in the Johnson&Johnson-Depuy (European Commission, 2012) EC rulings. In particular, the following seven markets were identified for the treatment group: hip prostheses; unicondylar, primary and revision knee prostheses; trauma devices; shoulder prostheses; spine devices. Other categories of prostheses and orthopedic accessories (patello-femoral knee implants, bone cement, elbow prostheses, dental prosthesis, etc.) were identified, but had too few observations. Therefore, we decided to group them in a residual market named "others". This residual

category also contains those lots for which the description was too generic, not allowing us to assign them to any specific market.

For the control group, we gathered all lots for pacemakers and of implantable defibrillators. These two items are often purchased jointly, making it unfeasible to assign them to separate product markets. This underscores the high degree of similarity between these two kind of devices, at least from the perspective of the contracting authority. Moreover, as mentioned above, the structure of the two markets is extremely similar, with the same 5 companies constantly accounting for about 95% of the total sales for both products (see Figure A.1 in the appendix). We thus treat pacemakers and implantable defibrillators jointly and refer to them as the pacemaker market.

3.2 The dependent variable

As firm price bids are expressed as percentage discounts over the reserve price, we use the winning discount on the lot reserve price as our dependent variable. We use it as a proxy of the final procurement price because, contrary, for instance, to infrastructure procurement, price renegotiation is rare in medical device procurement. Regarding the reserve price, this value is publicly known to the bidders as it is contained in the documents that publicize the call for tenders.

To assess how industry concentration affects the winning discount, it is necessary that the way in which the reserve price is set does not itself respond to the market structure.[8] Procurement regulations detail a set of tools that a purchasing authority can use to determine the reserve price. These include reference prices set by the national regulator, market surveys and regional prices lists, but they also leave a certain degree of discretion.[9] Therefore, to understand how the reserve price is determined in practice we interviewed the researchers of the "Observatory on Management of Public Procurement and Contracts in Health Care" (MASAN) of CERGAS – SDA Bocconi, a research center devoted to the study of the healthcare sector in Italy. They reported that the usual way that purchasing administrations proceed is to refer to a set of sources that they consider in a hierarchical order. First, they look at the outcomes of the past auctions they issued for the same item. If a certain item is purchased regularly, this is exactly what art. 35(12) of D.Lgs. 50/2016 prescribes. If this information is missing (this happens for example when it is the first time a particular item is purchased), the calls for tenders for similar items issued by different contracting authorities are considered. If this information is not available, the public procurer refers directly to the price lists of the suppliers. On the basis of this evidence, the purchasing administration marginally adjusts the reserve price with

[8] Indeed, suppose that two LPAs purchase the same item for an identical price, but one of them had set a higher reserve price than the other. This behavior would mean that the resulting winning rebates in the two auctions are not directly comparable.

[9] The regulations can be summarized as follows. Before the reform of the laws governing public procurement in 2016, the norm dealing with this task was art. 89 of the D.Lgs. 163/06. This established a set of tools that a purchasing authority had at its disposal to determine the reserve price, being the reference prices set by the AVCP (the predecessor of ANAC), generic market surveys or its own prices lists. A reference price scheme for medical devices was attempted by D.L. 98/2011, but the Regional Administrative Court (TAR) of Lazio declared it unlawful the following year. A new reference price list was set up by ANAC in 2014, but it regarded a tiny fraction of MDs. With the reform of the Code of Public Contracts in 2016, art. 89 was abrogated and substituted by art. 35 D.Lgs. 50/2016. The provisions contained in this norm are rather generic, specifying how to compute the reserve price only in few specific cases.

the aim of achieving a value that is closer to the true value an item has for its producer, in order to reduce progressively the rents of bidding firms.

At the same time, a reserve price that is too low would make the auction unattractive and put at great risk the timely supply of the MD. To minimize this risk, granting a "fair" amount of profit should be considered when setting the reserve price. Furthermore, other factors that prevents the reserve price falling close to the marginal cost of production of the most efficient firms are quality and innovation. Indeed, as explained in section 2, the assessment of quality has an important weight in the evaluation of bids.[10] According to the researchers interviewed, this trade-off results in a shrinkage of the average winning rebates over time. This is consistent with the trend of the average rebate over time observed in the data. But if this were the case, taking a simple difference of the rebate before and after the merger would clearly be misleading in assessing the treatment effect. A DID design naturally fixes this problem. However, for this identification strategy to hold in this case, we have to assume that, other things being equal, purchasing administrations set the reserve price in a similar "fashion" between the treatment and the control group. This condition seems realistic as the market of orthopedic prostheses and that of pacemakers are considered comparable from the point of view of the procurement process, as their inclusion in the list of DPCM 24/12/2015 (which involved only other three other kinds of MD) testifies. Indeed, these markets are both very mature and historically these MD are among those purchased with the highest frequency by LPAs.

3.3 Control variables

A series of factors must be taken into account to assess the relationship between concentration and rebates. A potentially relevant factor is the frequency with which an LPA purchases a class of MD: an LPA procuring more often a MD might gain a higher "expertise" in setting the reserve price and designing this procedure, which in turn might yield lower rebates over time. In order to take this into account, we built the variable *maturity*, which indicates for each observation the number of auctions that an LPA issued on that specific product from 2007 (the first year for which data are available).

Another key factor in determining MD prices might be the general degree of "ability" of the public procurer. Bucciol et al. (2020) showed that fixed effects for LPAs contribute to explaining much of the variation of the prices at which the NHS buys MD in Italy. Moreover, they demonstrated that these correlate well with variables denoting the competence of public procurers. In order to incorporate their conclusions in this study, we built a set of control variables at the LPA-year level which track the effect of LPA competence on rebates. For this purpose, we use data from the balance sheets of hospitals, available on the website of the Italian Ministry of Health. Following Bucciol et al. (2020), we consider a variable to account for potential scale economies in purchasing (the logarithm of total personnel cost) and two variables referring to the distribution of costs (the fraction of non-health personnel over total personnel costs and the fraction of medical devices purchases over total healthcare material purchases), which should serve as a measure of the "intensity" of the procurement activity in that LPA.

Balance sheet data, however, are not available for CPBs. Moreover, these data would require a different interpretation between hospitals and CPBs since the latter are public agencies exclusively devoted to procurement (i.e. they do not provide any health service).

[10] In Section B of the Appendix, we provide some examples that contribute to validate the description offered here on how the reserve price is set.

Table 1: Descriptive statistics of the two samples used

	LHA sample *(n=5,690)*				LHA + CPB sample *(n=7,490)*			
	Mean	St. Dev.	Max.	Min.	Mean	St. Dev.	Max.	Min.
rebates	0.17	0.24	0.99	0.00	0.18	0.24	0.99	0.00
Log Reserve Price	12.22	1.19	17.03	0.07	12.36	1.32	17.69	3.80
Log Maturity	1.30	1.07	4.23	0.00	1.25	1.07	4.23	0.00
Log Total Personnel	18.99	0.64	20.52	16.30				
Non-Health Pers.ratio	0.18	0.04	0.38	0.07				
MD ratio	0.36	0.11	0.75	0.02				
Dummies								
Treated	0.72				0.73			
First Price	0.11				0.09			
Negotiation	0.45				0.41			
Scoring Rule	0.35				0.43			
Budg. Plan	0.31				0.27			
CPB	0.00				0.19			

Note: on the left side, the Main sample, includes all lots awarded between 2012 and 2018 by all public buyers, excluding CPBs; on the right side, the CPBs sample contains the same lots as the Main sample as well as those of CPBs and those awarded in 2019 by all types of buyers. The (lot) rebate is the winning rebate, expressed as a percentage over the reserve price. Log Reserve Price is the log of the reserve price, measured in euro. Log Maturity records the log of the number of auctions that an LPA issued on that same category of prostheses since 2007. Log Tot Personnel is the log of total personnel cost. Non-Health Pers.ratio is the fraction of non-health personnel over total personnel costs. MD ratio is the fraction of medical devices purchases over the total healthcare material purchases. The cost measures used in these last three variables are all expressed in euros. The set of dummy variables in the lower part of the table are equal to one if the lot is for a treated product market (Treated) or is awarded: via price-only auctions (First Price), via a negotiated procedure (Negotiation), via a scoring rule auction (Scoring Rule), by an administration located in a region subject to a budgetary plan to curb healthcare expenditures (Budg. Plan), by a CBP (CPB).

Therefore, in our analysis we consider two samples: the *Main sample*, where CPBs are excluded (5690 observations), and a larger sample that also includes CPBs (7490 observations). We consider the former as our baseline sample as the availability of balance sheet data enhances the possibility to control for potential confounders. Moreover, since CPBs are usually able to obtain higher rebates compared to hospitals (Ferraresi et al., 2020), by analyzing the two samples separately we can highlight the potentially heterogeneous effects of industry concentration across different types of buyers.

The inclusion of CPB purchases might also complicate the interpretation of the estimates due to the interplay with a reform mentioned earlier regarding the centralization of procurement of hip prostheses and pacemakers that was implemented in the year right after the merger (2016). However, despite this law, LPAs continued to issue auctions for these MD well after 2016. The reason is that, in practice, many CPBs were not ready for this change (Osservatorio Conti Pubblici Italiani, 2018) and they needed more time to start to fully comply with this new obligation. At the same time, some "forerunnner" CPBs were active in the procurement of these items even before 2016. Thus, CPB observa-

tions are spread over much of the sample period and this allows us to evaluate the effect of the mergers in those estimates involving lots awarded by both LPAs and CPBs. Additional controls included in the analysis are three dummies for the auction formats (first price, scoring rule and negotiations), as categorized by Atella and Decarolis (2019), and whether the lot was awarded in a region subject to a budgetary plan to curb expenditures ("Piano di Rientro," labeled as *Budg. Plan*). Finally, since larger lots might be associated with higher discounts, we control for the logarithm of the reserve price. Descriptive statistics for the samples with and without the CPBs are presented in Table 1.[11]

4 Empirical analysis

4.1 The empirical model

The object of interest of our analysis is the Average Treatment Effect on the Treated (ATT), namely the price effect of the merger for the orthopedic prostheses markets. Two features of our data are particularly important in determining how we implement our DID strategy. The first is that the conditional distribution of the errors of our outcome variable, the winning rebates, is not normal as rebate is expressed in percentage terms, thus being defined only on the interval $[0,1)$. To deal with this "fractional outcome," we assume a beta distribution for the response variable.[12] In our case, the flexibility of the beta distribution to take on different shapes depending on its location and dispersion parameters is also important as the distribution of rebates is distinctly right-skewed, with most of the observations concentrated next to 0. Looking at Panel a) of Figure 2 reveals this by displaying the box-plots of the sample outcome variable over the years.

The second key challenge that the data present is a mass point at zero in the winning rebate distribution. The graph in panel b) of Figure 2 illustrates this fact: if we consider a histogram of the rebate values to approximate the empirical density function, we observe a large probability mass at the value 0 (indicated by the vertical black line). This conflicts with beta distribution not being defined at the boundary values 0 and 1. Our solution to cope with these problems is to follow the method by Ospina and Ferrari (2012), which entails adapting the beta distribution to allow for the value 0.

According to Ospina and Ferrari (2012) the random variable Y follows a beta distribution with parameters μ and ϕ ($0 < \mu < 1, \phi > 0$) if its density function is:

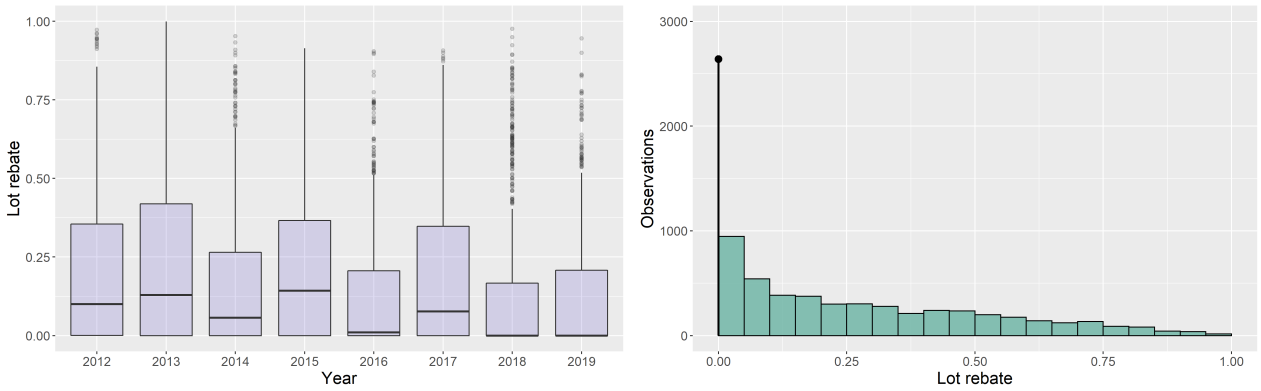
$$f(y;\mu,\phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \qquad y \in (0,1) \tag{1}$$

where $\Gamma(\cdot)$ is the gamma function. If $Y \sim \mathcal{B}(\mu,\phi)$ then $E(Y) = \mu$ and $Var(Y) = \mu(1-\mu)/(\phi+1)$, hence μ is the distribution mean and ϕ plays the role of a precision parameter. Then, to allow for the value zero, we introduce a mixed continuous-discrete distribution that we

[11] Relative to the LHA sample, the CBP sample covers all CPB-issued auctions, as well as the 2019 purchases. The 2019 data are not in the LHA sample because balance sheets data are not available after 2018.

[12] The literature offers various approaches to deal with this type of "fractional outcome," ranging from simply ignoring this feature to sophisticated methods like that in Papke and Wooldridge (1996). Tobit models have been used, as well as models which link the predictors to the conditional expectations of the outcome through functions mapping in the unit interval, as the logistic function (for an overview see Ramalho et al. (2011)). The beta regression approach that we follow belongs to this latter group.

Figure 2: Empirical distribution of the outcome variable



(a) Boxplot of rebates over the years

(b) Histogram of rebates

call a *zero-inflated beta distribution* (ZIBD), with density:

$$ZIBD(y;\alpha,\mu,\phi) = \begin{cases} \alpha & y = 0 \\ (1-\alpha)f(y;\mu,\phi) & y \in (0,1), \end{cases} \tag{2}$$

where $f(y;\mu,\phi)$ is the beta density as parametrized above and α is the probability mass at 0. Essentially, the ZIBD is a mixture distribution of a beta and a degenerate distribution at 0. The mean and the variance are respectively:

$$E(y) = (1-\alpha)\mu, \tag{3}$$

$$Var(y) = (1-\alpha)\frac{\mu(1-\mu)}{\phi+1} + \alpha(1-\alpha)(-\mu)^2. \tag{4}$$

The expected value is thus the weighted average of 0 and μ. We are now ready to define our regression model, the *zero-inflated beta regression* (ZIBR). Given our matrix of data in which we have n independent realizations of $(y_i, \mathbf{x}_i)$, where y_i is the outcome variable for observation and $\mathbf{x_i} = (x_{1i},...,x_{ki})^\intercal$ its vector of covariates, we assume that:

$$y_i|\mathbf{x}_i \sim ZIBD(\mu_i,\phi_i,\alpha_i), \tag{5}$$

with $g_1(\mu_i) = \mathbf{x_i}^\intercal\beta$ and $\beta = (\beta_i,...,\beta_k)^\intercal$ the vector of unknown regression parameters. This implies that:

$$E(y_i|\mathbf{x}_i) = \mu_i = g_1^{-1}(\mathbf{x_i}^\intercal\beta). \tag{6}$$

We further assume that also the precision parameter (ϕ_i) and the probability at 0 (α_i) are functions of our linear predictors:

$$g_2(\phi_i) = \mathbf{x_i}^\intercal\gamma, \tag{7}$$

$$g_3(\alpha_i) = \mathbf{x_i}^\intercal\rho, \tag{8}$$

where again γ and ρ are vectors of unknown regression parameters to be estimated. In this

setup, g_1, g_2, g_3 are called *link functions* and we assume that $g_1 : (0,1) \to \mathbb{R}$, $g_2 : (0,+\infty) \to \mathbb{R}$ and $g_3 : (0,1) \to \mathbb{R}$ and that they are strictly monotonic and twice differentiable (thus, invertible). There are several alternatives available for these functions. We use the standard logit link for g_1 and g_3 (e.g. $g_1(\mu) = \log[\mu/(1-\mu)]$) and the log link for g_2 ($g_2(\phi) = \log(\phi)$). This model is estimated by maximum likelihood, which under the usual regularity conditions allows us to obtain consistent estimates of (β, γ, ρ).[13]

Our DID identification strategy is embedded into the $g_1(\mu)$ link function, where the set of linear predictors is then defined as:

$$g_1(\mu_i) = X_{i,t}\beta = \beta_0 + \beta_1 Treat_i + \beta_2 Treat_i \text{ x } Post_t + \beta_3 Z_{i,t} + \nu_t \tag{9}$$

and where subscript i indicates the lot, t the year, $Treat_i$ is a dummy that takes value 1 if the observation belongs to the treatment group, $Post_t$ is another dummy equal to one for lots issued from 2015 onward, $Z_{i,t}$ is the matrix of controls listed in Table 1 which help to make our identification strategy more plausible (Lee and Kang, 2006) and, finally, ν_t are year-fixed effects.

4.2 The control group

From an empirical point of view, among the several challenges embedded in a DID analysis the key challenge for our identification strategy is represented by the selection of an adequate control group. Ideally, this should be a similar product market, hit by the same supply and demand shocks as the treatment group, but different enough so as not to be affected directly by the treatment. In section 2.2, we highlighted how pervasive the phenomenon of consolidation is across MD manufacturers. Thus, it is not easy to find a MD market where no mergers between rivals occurred within the window of time of our analysis.

The market of pacemakers, however, serves this purpose well. As discussed in Cergas (2017), pacemakers are the only MD for which concentration was stable over the period 2014-2017. At the same time, as is evident from Figure 1, this level of concentration stands between the treated markets with the highest and lowest HHI, representing somehow a sample average of concentration for treated markets.[14] Finally, and most crucially, Figure 3 supports our choice of the control group by showing that the pre-treatment trends of the average rebates of these two markets co-moves pretty well.

[13] Finite-sample properties are studied by Ospina and Ferrari (2012) through a Monte-Carlo simulation and confirm the ability of the estimation algorithm to provide unbiased estimates when the model is valid for the underlying data.

[14] Notice also that the "technological complexity" of pacemakers is comparable to that of the treated product markets if compared to other frequently purchased MD such as syringes and needles, or cotton gauze. Another hint pushing in favor of the comparability of the markets of pacemakers and prostheses (hip prostheses in particular, the most numerous category) is their inclusion in the list of DPCM 24/12/2015 on procurement centralization together with just three other kinds of MD. This suggests that the public administration perceives these two markets as similar from the point of view of their maturity, of the degree of standardization and of their budgetary impact. Therefore, pacemakers will be our control market.

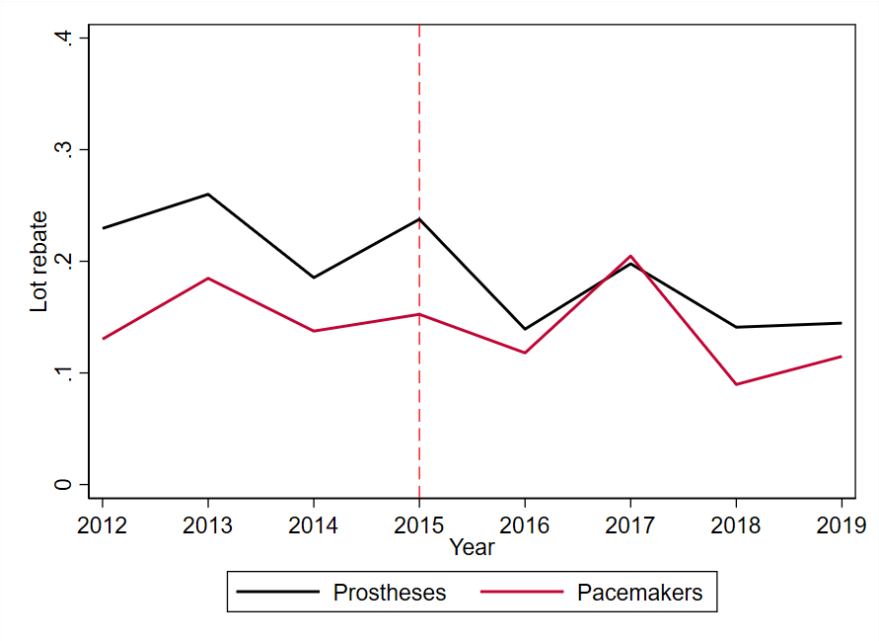


Figure 3: Average rebates over time across treatment and control groups

5 Results

In this section we present the results of our empirical analysis. We start by discussing the evidence obtained from the sample of tenders in the LHAs. We then expand the analysis to look at the extended sample of LHAs and CPBs. As a further refinement, we investigate the presence of heterogeneity across prostheses markets. Lastly, we conclude with a discussion of some robustness checks.

5.1 LHA sample

Table 2 presents the results of the estimation of ZIBR model for the pooled sample of markets of orthopedic prostheses for four specifications with an increasing number of covariates. Estimates are reported as average marginal effects. Recall that the dependent variable is the (lot) rebate, expressed as a percentage of the reserve price, thus positive coefficients in Table 2 indicate higher discounts (i.e., lower procurement prices). The table reveals that the basic DID model without any additional controls, column (1), yields a negative coefficient of -0.024 which is only weakly statistically significant. In the next columns, as we introduce controls to differentiate between different types of public buyers and different award procedures, both the magnitude and the significance of the coefficient on the treatment effect increase. In particular, this coefficient is -0.076 and highly statistically significant in column (2), where we control for the different auction formats – first price, scoring rule and negotiated procedures. This difference between the estimates in column (1) and (2) can be explained considering that negotiated procedures are more frequent in the procurement of prostheses than in that of pacemakers, and that these procedures have systematically lower rebates relative to the other auction formats. In general, controlling for features of the auction formats appears highly relevant as, in fact, the co-

efficients on both first price and negotiated procedures are both large in magnitude and statistically significant.

The following two columns of the table confirm the evidence in column (2): the effect of the merger is highly statistically significant and its magnitude corresponds to a drop in the winning discount of between 8.5 and 8.8 percentage points when we further control for features of the contract (the log of the reserve price) and the public buyer, in terms of both experience with the procurement of the same category of prostheses (*Maturity*), organizational structure and competence level (the latter three variables in column (4) of Table 2).

Table 2: Regression results from the ZIBR model for the LHA sample.

Dependent var: Lot Rebate	(1)	(2)	(3)	(4)
Treated	0.058***	0.079***	0.081***	0.076***
	(0.009)	(0.008)	(0.009)	(0.009)
Treated x Post	-0.024*	-0.076***	-0.085***	-0.088***
	(0.013)	(0.012)	(0.012)	(0.012)
First Price		0.086***	0.043**	0.035*
		(0.018)	(0.018)	(0.018)
Negotiation		-0.180***	-0.178***	-0.177***
		(0.015)	(0.015)	(0.015)
Scoring Rule		0.022	0.010	0.013
		(0.015)	(0.016)	(0.015)
Maturity			0.003	0.003
			(0.003)	(0.003)
Log Reserve Price			0.010***	0.010***
			(0.002)	(0.002)
Budg. Plan			0.051***	0.041***
			(0.007)	(0.007)
Log Personnel				0.010**
				(0.004)
Non-Health Pers. Ratio				-0.362***
				(0.075)
MD Ratio				-0.065**
				(0.029)
Observations	5,690	5,690	5,690	5,690

Note: Robust standard errors in parenthesis.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Regarding the signs and magnitudes of the coefficients other than the merger, they all appear reasonable. Other things being equal, larger auctions, as measured by their reserve price, are associated with higher discounts, consistent with the presence of economies of scale.[15] Finally, other factors that lead to larger rebates are: being in a region under a

[15]Controlling for the (log of) the reserve price helps to proxy for the economic value and, potentially, the complexity of the procurement event. But it is also important because different regulations apply for contracts crossing certain values of the reserve price. Our results are qualitatively the same if, instead of a continuous measure as in Table 2, we control for dummy variables grouping lots depending on the threshold values of the reserve price that they cross.

budgetary program (*Piano di Rientro*), which is consistent with the emphasis public buyers in these regions must put on curtailing healthcare costs; being awarded by an LPA with a high share of administrative personnel relative to medical personnel, which is consistent with the existing results in the literature discussed earlier.

Overall, the estimates above indicate that the merger caused a reduction in the winning rebates between 7 and 9 percentage points on the pooled markets for orthopedic prostheses. This supports the hypothesis that consolidation caused a price increase. To facilitate the economic interpretation of these estimates notice the following. The magnitude of this effect appears substantial if we consider that the average rebate for the treatment group after 2014 was 18% (slightly above the average of the whole sample, 17%, as reported in Table 1). According to these results, the rebate for the treated observations should have been on average at least 7 percentage points higher absent the merger. To get a sense of the economic magnitudes involved consider the following: the Italian Ministry of Health (2018) reports that in 2017 the NHS spent €427 million on orthopedic devices and that the lots were awarded at an average rebate of 20% in that specific year, thus, assuming a counterfactual rebate of 27% (i.e., the baseline 20% percent discount increased by the most conservative estimate of 7% percent), we see that the final expenditure absent the merger would have been about €389 million, or €38 million less than the actual. If we considered the whole period after the merger (2015-2019) and we repeated this calculation for every year, the impact of the merger on the Italian State's balance sheet would be equal to €190 million.

5.2 Full sample (LHA + CPB)

Table 3 presents the results from the estimation of the ZIBR model on the larger LHA and CPB sample.[16] The coefficients are close to those estimated in the previous sample in terms of both magnitude and significance. Its magnitude becomes stable after the inclusion of the auction format (from column (2) onward), ranging around to 6 percentage points. Interestingly, the lots awarded through a CPB show just a tiny increase in the rebate (1.5%) compared to those awarded by LPAs, suggesting that the impact of purchasing through centralized bodies is still limited for the goods we are considering. Approximating public expenditure as we did at the end of the previous paragraph, under the counterfactual scenario of an increase of the average discount by 6 percentage points for treated observations, we would obtain €161 million of total savings for Italy's accounts over the period 2015-2019.

[16] Recall that there are three main reasons why we deem our previous set of results as the most appropriate to capture the ATT. First, by restricting the sample to LPAs, the comparability among observations is enhanced. Second, this allows us to control for the potentially confounding effect of LPA "competence" in procurement on the outcome of the auctions by including in the regressor set the balance sheets data, which have been found by the literature to be particularly relevant (Bucciol et al., 2020). Third, the DPCM 24/12/2015 was enforced right in the year after the merger, thus, by excluding the auctions awarded through CPBs, we avoid biases due to procurement centralization. However, in practice these problems are less severe than in theory: indeed, we can partially control for the effect of CPBs by flagging these observations with a dummy; moreover, the enforcement of the DPCM was not strict as many CPBs were not yet functioning at that moment, while at the same time other CPBs were purchasing MD listed in this provision even before 2016. Furthermore, by including lots awarded through CPBs, the sample gets considerably larger and the sample period extends by one year (2019).

Table 3: Regression results of ZIBR model on the LHA + CPB sample

Dependent var: Lot Rebate	(1)	(2)	(3)	(4)
Treated	0.059***	0.083***	0.088***	0.084***
	(0.010)	(0.009)	(0.009)	(0.009)
Treated x Post	-0.021*	-0.061***	-0.062***	-0.058***
	(0.012)	(0.012)	(0.012)	(0.012)
First Price		0.087***	0.076***	0.071***
		(0.018)	(0.019)	(0.019)
Negotiation		-0.177***	-0.179***	-0.184***
		(0.014)	(0.015)	(0.015)
Scoring Rule		0.042***	0.032**	0.021
		(0.015)	(0.015)	(0.015)
Maturity			0.004	0.005*
			(0.003)	(0.003)
Log Rsv. Price			0.009***	0.007***
			(0.002)	(0.002)
Budg. Plan			0.008	0.008
			(0.006)	(0.007)
CPB				0.028***
				(0.007)
Observations	7,490	7,490	7,490	7,490

Note: Robust standard errors in parenthesis.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

5.3 Product-specific estimates

In order to understand if there are heterogeneous effects within this market, we replicate the previous analyses on sub-samples made by each specific market for orthopedic prostheses. In the top panel of Table 4, we report the estimates for the LHA sample. A first remark is that the treatment effect coefficient is still negative and varies across markets, ranging in an interval from a low of 7.7 percentage points (for trauma) to a high of 16.4 percentage points (for spine). The estimates for knee and shoulder devices are similar at 14 and 15.7 percentage points respectively, while that for hips is lower at 11.8 percentage points. However, we remark that the estimates for knee must be considered cautiously, as they pool together unicondylar, revision and primary implants because only a few dozen lots are observed for each one of these products taken individually and, moreover, lots often involve the joint procurement of different types of knee devices. The level of significance of all coefficients is rather homogeneous, all being statistically significant at the 1% level.

The finding is qualitatively similar for the full sample (LHA+CPB) in the bottom part of the table. However, like what we found by comparing the estimates in Table 2 and Table 3, for the product-specific estimates we observe again that all estimates are attenuated in the full sample: all estimates are roughly halved in magnitude, except for those of spine devices whose decline is more limited across the two samples. Regarding the rather limited merger effect for trauma devices, this is consistent with our earlier discussion about this market which does not show any major change in concentration.

Table 4: Regression results of the ZIBR model for specific markets.

	(1) Hip	(2) Knee	(3) Shoulder	(4) Spine	(5) Trauma
LHA sample:					
Treated x Post	-.118***	-.140***	-.157***	-.164***	-.077***
	(.018)	(.028)	(.033)	(.021)	(.014)
Observations	2,162	1,753	1,697	2,188	3,417
Treat. Obs.	586	177	121	612	1841
LHA+CPB sample:					
Treated x Post	-.069***	-.070***	-.070***	-.145***	-.045***
	(.016)	(.023)	(.027)	(.018)	(.013)
Observations	2,890	2,311	2,214	2,800	4,503
Treat. Obs.	685	269	172	758	1427

Note: Robust standard errors in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

5.4 Robustness checks

We conclude this section with an assessment of the robustness of our baseline estimates to two main modifications: *i)* different time windows for the analysis sample and *ii)* different control groups.

The first exercise is particularly useful to explore how the price effect of the merger evolves over time. By including an increasing number of years post-treatment, we can assess whether the treatment effect stays constant, increases or fades out. Table 5 reports the results from the estimation of the ZIBR model under different time windows for the LHA+CPB sample.

Table 5: ZIBR regression results for the CPB sample under different time windows.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Periods	12-19	12-18	12-17	12-16	13-19	13-18	13-17	13-16
Treated x Post	-.058***	-.062***	-.067***	-.042***	-.037***	-.040***	-.041***	-.019
	(.012)	(.012)	(.013)	(.014)	(.014)	(.014)	(.015)	(.016)
Observations	7,490	6,997	6,099	4,960	6,685	6,192	5,294	4,155

Note: Robust standard errors in parenthesis. The header of the columns indicates the period under which the ZIBR is estimated, reporting just the last two digits of the year (i.e. *12-16* means *2012-2016*).
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

The first three columns keep the usual starting year while shrinking the post-treatment period. The first thing that we notice is that the coefficient is stable across the columns, with the peak occurring when excluding the last year of the sample from the estimation (2019, column (2)). At the same time, if we keep just the first year post merger (2016, column (4)) we get the lowest treatment effect, which differs from the other estimates by 1-2 percentage points. This suggests that the merger had an immediate effect on prices

after 2015 and that this does not fade out over time, but stays rather constant or even reinforces in the following years. Columns (5) to (8) repeat the same exercise setting 2013 as the starting year of the sample. We notice that the treatment effect is extremely stable to this shrinkage of the pretreatment period, both in magnitude and in significance. Column (8) displays the model estimated on the shortest sample (2013-2016), which exhibits the smallest coefficient. However, even in this case we do not acknowledge any significant deviation from the ATT in the other columns, differing only 0.006 from the coefficient under the full pre-treatment period (column (4)).

The second set of robustness checks involve the use of alternative control groups. As discussed earlier, it is difficult to find a control group which is both not directly affected by the merger, and subject to the same demand and supply shocks as the treatment. In our case, by considering pacemakers a good benchmark, we privileged the first characteristic, while we have no strong evidence on the validity of the second condition. Finding confirmation of our results from another control group which tries to account for the idiosyncratic developments of the orthopedic prostheses markets would make previous results more robust. This is indeed the strategy proposed by Ashenfelter and Hosken (2010) as a particularly demanding test of the DID model. Following ideas put forward in Ashenfelter and Hosken (2010), we implement this approach by using the trauma market as control.[17] As seen in Figure 1, concentration is relatively low in this market and did not change after the merger.[18]

The extreme fragmentation of this market and the relatively low market power of the merging firms suggests *a priori* little to no effect of the merger on prices. Also, this was the same conclusion as the European Commission, which dedicates to this sector very little attention in the Zimmer-Biomet final decision, as the evidence on the high competitiveness of the market was overwhelming. The DID model we estimated before revealed that a significant reduction in rebates did occur after 2015, but it was lowest among the markets in which Zimmer and Biomet were present (around 5%). At the same time, we can safely claim that trauma devices are the MDs that are most likely to suffer from the same supply and demand shock as orthopedic prostheses among those available. Therefore, we can conclude that this control is complementary in terms of its properties compared to pacemakers.

Using the full sample (LHA+CPB), in Table 6 we report the estimates with trauma as control group. As expected, the magnitude and the significance of the ATT using this closer control are in general much lower than the baseline ones. Nonetheless, the fact that they remain significant and still economically meaningful with an estimated effect of 3 percentage points is highly indicative of the robustness of our baseline estimates.

Finally, in the appendix, we also explore further margins of robustness involving different econometric models (linear and hurdle models) and different model specifications, as well as additional tests of the benchmark DID model. Overall, the robustness checks broadly confirm the conclusions from our baseline analysis.

[17] The market for trauma devices includes all the tools used in osteosynthesis, that is the treatment of bone fractures through the surgical implementation of an implant. They include, for example, plate, screws, pins, intramedullary nails and systems for external fixation. The most attractive feature of this market is that it is extremely competitive.

[18] Indeed, in the year before the merger, Zimmer was the third largest producer by sales value in this market, but had a market share of just 7%. Biomet was the tenth, with only 1,7%. In that year, the number of firms involved in supplying the NHS with these devices was more than 150.

Table 6: ZIBR regression results using trauma devices as control (LHA+CPB sample).

Dependent var: Lot Rebate	(1)	(2)	(3)	(4)
Treated	-0.015	-0.016*	-0.020**	-0.016*
	(0.011)	(0.009)	(0.009)	(0.009)
Treated x Post	-0.006	-0.033***	-0.030**	-0.033***
	(0.013)	(0.012)	(0.012)	(0.012)
First Price		0.167***	0.163***	0.154***
		(0.020)	(0.021)	(0.021)
Negotiation		-0.168***	-0.168***	-0.172***
		(0.015)	(0.015)	(0.016)
Scoring Rule		0.073***	0.063***	0.051***
		(0.016)	(0.016)	(0.016)
Maturity			0.004	0.005
			(0.003)	(0.004)
Log Rsv. Price			0.013***	0.011***
			(0.002)	(0.002)
Budg. Plan			-0.006	-0.003
			(0.008)	(0.009)
CPB				0.032***
				(0.009)
Observations	5,448	5,448	5,448	5,448

Note: Robust standard errors in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

6 Conclusion

Public procurement has a crucial influence on national accounts. Healthcare procurement is both a major component of public procurement in developed countries and an important driver of how healthcare systems deliver value to citizens. While both the public debate and the academic research have often focused on issues such as corruption, cartels and incomplete contracts, the role of supplier market power in public procurement markets has had little attention. This study considers the case of orthopedic prostheses, which account by far for the largest amount of public expenditure among medical devices in Italy, and tries to establish whether a causal link between concentration and outcomes of procurement auctions exists. Our results indicate that the merger between two US companies that are among the global market leaders in this sector had a substantial effect on prices: the estimated effect is of the order of €175 million over the period 2015-2019.

The evidence provided in this study is also indicative of the need to better coordinate public procurement and antitrust policies. First, it is well known that the tools of the antitrust authorities to evaluate merger effects in procurement markets are not as sophisticated as those used in more conventional product markets. Part of the reason lies in the nature of public procurement where purchases are lumpy and the very definition of markets and market shares is problematic.[19] Second, our analysis reveals how the outcomes

[19]In footnote 6 at page 5 of the final ruling of the EC on the Zimmer - Biomet merger (European Commission, 2015), the EC describes bidding data (data on procurement auctions, often publicly available) as "unsuitable to conduct a meaningful analysis". A rigorous quantitative analysis of substitution patterns was

of national public procurement might be adversely affected by supply concentration decisions taken by foreign firms. In this case the existence of differences in the flexibility of the procurement system between the US and Italy might contribute to explaining why the latter is likely to face harsher price increases than the former.[20] The policy implication stemming from such a setup is to promote a local antitrust authority embedded with knowledge of the specificity of its national public procurement system. Thus, in the absence of reforms in the public procurement system aimed at increasing its flexibility, suppliers concentration might require a more stringent enforcement of antitrust regulations relative to that applied to evaluate horizontal mergers in typical markets.

Finally, despite the difficulties in conducting reliable cross-country studies due to the lack of comparable data, it would certainly be very valuable to assess how different procurement systems responded to concentration in the MD industry and which procurement reforms might be implemented to limit the potential problems caused by supplier concentration in terms of both short run price increases and longer run quality declines.

not possible in this case. To recover the market definition, the EC had to resort to the opininos of expert and market participants. Even the SSNIP test was performed by means of questions to market participants (see page 17 of the ruling). To cope with this obstacle, the EC tested the impact of the merger under different definitions of product markets. Regarding market shares, the only way for the EC to proxy them was to ask directly for volumes and sales data from producers and third party collectors.

[20]Furthermore, in the US a relevant component of medical device purchases is performed through bargaining by private hospitals, see Grennan 2015.

References

AGENAS (2018). Programma nazionale esiti – pne edizione 2019. https://pne.agenas.it/main/doc/Report_PNE_2019.pdf. [accessed 5-January-2021].

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Ashenfelter, O. and D. Hosken (2010). The effect of mergers on consumer prices: Evidence from five mergers on the enforcement margin. *The Journal of Law and Economics 53*(3), 417–466.

Ashenfelter, O. C., D. S. Hosken, and M. C. Weinberg (2013). The price effects of a large merger of manufacturers: A case study of maytag-whirlpool. *American Economic Journal: Economic Policy 5*(1), 239–61.

Atella, V. and F. Decarolis (2019). Procuring medical devices: Evidence from italian public tenders. *CEIS Working Paper 472*.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics 21*(1), 1–42.

Benkard, L., A. Yurukoglu, and A. L. Zhang (2021). Concentration in product markets.

Bosio, E., E. L. Djankov, Simeonand Glaeser, and A. Shleifer (2020). Price discrimination and bargaining: Empirical evidence from medical devices.

Bucciol, A., R. Camboni, and P. Valbonesi (2020). Purchasing medical devices: The role of buyer competence and discretion. *Journal of health economics 74*, 102370.

Carril, R. and M. Duggan (2020). The impact of industry consolidation on government procurement: Evidence from department of defense contracting. *Journal of Public Economics 184*, 104141.

Cergas, S. B. (2017). Osservatorio sulle aziende e sul sistema sanitario italiano. rapporto oasi 2020. (Capitolo 14).

Crotti, N. (2017, October). It's another big year for medtech M&A. https://www.medicaldesignandoutsourcing.com/ianother-big-year-medtech-ma/.

Decarolis, F. and C. Giorgiantonio (2015). Public procurement of healthcare in europe: The case of medical devices. *Rivista di Politica Economica 104*, 4.

Deneckere, R. and C. Davidson (1985). Incentives to form coalitions with bertrand competition. *The RAND Journal of economics 16*(4), 473–486.

Dunn, P. K. and G. K. Smyth (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics 5*(3), 236–244.

European Commission, D. C. (2012, February). Case no iv/m.1286 – johnson&johnson / depuy. *REGULATION (EEC) No 4064/89 MERGER PROCEDURE*. [accessed 9-January-2021].

European Commission, D. C. (2015, February). Case m.7265 – zimmer/biomet. *MERGER PROCEDURE, REGULATION (EC) 139/2004*. [accessed 9-January-2021].

Feng, C., A. Sadeghpour, and L. Li (2017). Randomized quantile residuals: an omnibus model diagnostic tool with unified reference distribution. *arXiv preprint arXiv:1708.08527 7*.

Ferraresi, M., G. Gucciardi, and L. Rizzo (2020). Savings from public procurement centralization in the healthcare system. *European Journal of Political Economy 66*(1), 1019–63.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review 103*(1), 145–77.

Italian Ministry of Health (2017). Consumption flows of medical devices. http://www.salute.gov.it/portale/temi/p2_6.jsp?lingua=italiano&id=4652&area=dispositivi-medici&menu=spesa. [accessed 15-December-2020].

Italian Ministry of Health (2018). Rapporto sulla spesa rilevata dalle strutture sanitarie pubbliche del ssn per l'acquisto di dispositivi medici - anno 2017. http://www.salute.gov.it/imgs/C_17_pubblicazioni_2806_allegato.pdf. [accessed 5-January-2021].

Lee, M.-j. and C. Kang (2006). Identification for difference in differences with cross-section and panel data. *Economics letters 92*(2), 270–276.

MedTech Europe (2020). https://www.medtecheurope.org/wp-content/uploads/2020/05/The-European-Medical-Technology-Industry-in-figures-2020.pdf. [accessed 25-January-2021].

Ospina, R. and S. L. Ferrari (2012). A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis 56*(6), 1609–1623.

Osservatorio Conti Pubblici Italiani (2018, April). Gli acquisti di beni e servizi della Pubblica Amministrazione: si sta risparmiando? https://osservatoriocpi.unicatt.it/. [accessed 10-January-2021].

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of applied econometrics 11*(6), 619–632.

Ramalho, E. A., J. J. Ramalho, and J. M. Murteira (2011). Alternative estimating and testing empirical strategies for fractional regression models. *Journal of Economic Surveys 25*(1), 19–68.

Rizzo, I., C. Guccio, and G. Pignataro (2006). Efficiency of procurement procedures for medical devices. *Rivista di politica economica 96*(1-4), 135.

Romanini, E., F. Decarolis, I. Luzi, G. Zanoli, M. Venosa, P. Laricchiuta, E. Carrani, and M. Torre (2019). Total knee arthroplasty in italy: reflections from the last fifteen years and projections for the next thirty. *International orthopaedics 43*(1), 133–138.

Smithson, M. and J. Verkuilen (2006). A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods 11*(1), 54.

Vellez, M. (2012). Auctions versus negotiations: Evidence from public procurement in the italian healthcare sector. *CEIS Working Paper* (235).

Zimmer Inc. (2015). 2014 annual report. https://investor.zimmerbiomet.com/~/media/Files/Z/ZimmerBiomet-IR/documents/annual-reports/zimmer-2014-annual-report.pdf. [accessed 7-January-2021].

Procuring Medical Devices: The Price Effect of Mergers among Orthopedic Prostheses Producers Web Appendix

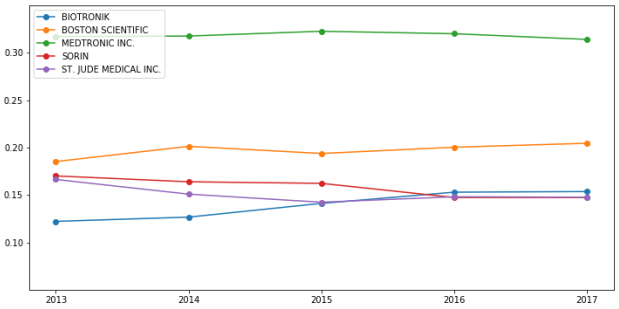
A. Consumption Flows Data and the Pacemaker Market

The data used to calculate the HHI comes from the Italian Ministry of Health - Consumption Flows Data. In Table A.1 we report the detailed expenditures statistics by year and market. These data are also available for the control and we use them in Figure A.1 to illustrate the composition of the main producers in the pacemaker market.

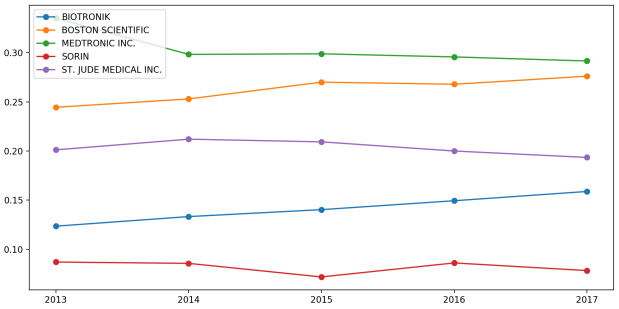
Table A.1: Consumption Flows Data

Market	Year	Yearly expenditures	Average invoiced amount	Amount of invoices	Zimmer	Biomet	HHI
Hip	2013	119,566,711	4,703	25,424	1	5	0.09387
Hip	2014	132,917,502	6,853	19,395	1	4	0.08583
Hip	2015	131,985,751	6,478	20,374	1	1	0.11150
Hip	2016	135,404,193	6,279	21,564	1	1	0.10378
Hip	2017	131,369,136	6,193	21,211	1	1	0.10253
Primary Knee	2013	41,979,232	6,976	6,018	1	3	0.13856
Primary Knee	2014	46,003,917	9,620	4,782	1	3	0.14549
Primary Knee	2015	47,149,102	8,792	5,363	1	1	0.20624
Primary Knee	2016	49,053,333	8,410	5,833	1	1	0.19676
Primary Knee	2017	47,940,109	8,067	5,943	1	1	0.19296
Unicondylar Knee	2013	2,617,740	2,674	979	2	1	0.17141
Unicondylar Knee	2014	2,715,344	3,930	691	1	2	0.16896
Unicondylar Knee	2015	2,582,853	3,068	842	1	1	0.27955
Unicondylar Knee	2016	2,738,824	3,174	863	1	1	0.26322
Unicondylar Knee	2017	2,964,056	3,123	949	1	1	0.28445
Revision Knee	2013	6,971,889	4,865	1,433	1	5	0.39010
Revision Knee	2014	8,099,745	6,929	1,169	1	5	0.40431
Revision Knee	2015	8,707,218	6,632	1,313	1	1	0.42803
Revision Knee	2016	8,877,957	6,429	1,381	1	1	0.36335
Revision Knee	2017	9,319,721	6,783	1,374	1	1	0.37865
Spine	2013	39,706,017	5,379	7,382	3	16	0.15368
Spine	2014	44,046,190	7,250	6,075	3	17	0.14157
Spine	2015	45,696,938	7,296	6,263	3	3	0.14421
Spine	2016	46,198,836	6,956	6,642	5	5	0.13897
Spine	2017	48,787,630	6,845	7,128	6	6	0.12742
Shoulder	2013	8,354,152	1,884	4,434	2	5	0.36762
Shoulder	2014	10,238,431	2,700	3,792	2	5	0.30711
Shoulder	2015	11,099,665	2,680	4,142	2	2	0.32245
Shoulder	2016	11,709,531	2,593	4,516	2	2	0.29070
Shoulder	2017	12,518,530	2,660	4,706	2	2	0.27565
Trauma	2013	108,662,161	1,998	54,390	3	12	0.12542
Trauma	2014	131,014,759	2,947	44,462	3	11	0.12715
Trauma	2015	139,273,822	2,871	48,519	3	3	0.13039
Trauma	2016	147,149,350	2,829	52,008	3	3	0.13146
Trauma	2017	149,909,238	2,700	55,526	3	3	0.12841

Figure A.1: Market shares of the first 5 companies.



(a) Pacemakers



(b) Implantable Defibrillators

Note: Source: *Consumption Flows of medical devices* (Italian Ministry of Health, 2017). Pacemakers are identified through the CND code J0105, while implantable defibrillators are identified through the code J010.

B. Reserve Price

We report here some extracts and translation from documents of procurement auctions of MDs that validate the anecdotal evidence provided in Section 3.2.

As we mentioned in Section 3.2, contracting authorities are not obliged to publish any information on the procedure they adopt to set the reserve price of each auction. A preliminary research on the Internet was unsuccessful in providing an answer to this question. However, thanks to the insights provided by MASAN researchers, we were able to identify some formal documents in which these procedures are briefly outlined. They all come from ARIA spa, the CPB of the Lombardy region (referred to also as ARCA in the documents, that is its old name).

Figure B.1: Extract from procurement auctions documents on the reserve price

seguito, per brevità, "**Servizio**"), per un importo massimo di Euro **2.600,00** (duemilaseicento/00), oltre IVA, stimato sulla base delle tariffe della convenzione messa a disposizione da ARIA S.p.A. ed utilizzata in precedenza per l'acquisto di servizi analoghi;

(a) Extract A

Translation: "... estimated on the basis of the contract provided by ARIA S.p.A. and previously used to purchase similar services."

4. con riferimento al suddetto acquisto la Struttura Richiedente ha individuato un importo massimo di **Euro 800,00** (ottocento/00), oltre IVA, stimato sulla base di una preliminare consultazione dei prezzi di mercato;

(b) Extract B

Translation: "... identified a ceiling amount of Euro 800.00, excluding VAT, estimated on the basis of a preliminary inquiry on market prices."

La base d'asta è stata stabilita secondo il seguente criterio: per ciascun lotto si è calcolato il prezzo medio ponderato (PMP) degli attuali prezzi d'acquisto lombardi confrontandola con i prezzi di aggiudicazione della convenzione ARCA in vigore. Definito il PMP di ciascun lotto, si è proceduto a definire il prezzo a base d'asta che, in alcuni casi risulta superiore ai prezzi di aggiudicazione della gara ARCA_2015_51 in ragione della necessità di consentire l'accesso alla gara anche a modelli protesici innovativi e di recente commercializzazione. Tuttavia, al contempo, viene garantito comunque un risparmio complessivo su scala regionale: la spesa totale annuale calcolata con il PMP di ciascun lotto ammonta a 3.155.779,14 €, la spesa totale annuale calcolata con la BdA ammonta a 2.811.000,00€ con una flessione rispetto alla spesa storica del 10,93%.

(c) Extract C

Translation: "The reserve price is set according to the following criterion: for each lot, the weighted average price (PMP) of the current prices at which Lombardy purchases has been computed and compared with the prices awarded in the last ARCA procedure, still valid. Once the PMP for each lot has been established, the reserve price has been computed, resulting occasionally higher than the awarding prices of auction ARCA-2015-51 due to the necessity of allowing innovative prostheses implants and recently-commercialized implants to participate. However, at the same time, savings are guaranteed..."

Panel a) and b) of Figure B.1 shows some extracts from documents called "Delibere a contrarre" and refer to two procurement procedures on services purchased by ARIA.[21]

[21]These documents can be found at the following address `https://www.finlombarda.it/societatrasparente/bandidigaraecontratti/determine`. They represent internal administrative formal acts that contracting authorities often do not release. Furthermore, legislative provisions on these acts do not oblige the contracting authorities to describe the process through which the reserve price has been derived (see art. 11 of D.Leg. n. 163/2006 and art. 192 of T.U.E.L. n. 267/2000). Therefore, we consider the availability of this information to be an exception.

The first clearly states that ARIA based the reserve price on the outcome of previous auctions on the same items. The second says that ARIA set the reserve price after having inquired about the market price for the items it wanted to purchase.

Table B.1

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.079***	0.082***	0.081***	0.083***	0.084***	0.084***
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Treated x Post	-0.076***	-0.079***	-0.086***	-0.061***	-0.060***	-0.060***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.012)
first price	0.086***	0.074***	0.044**	0.087***	0.079***	0.072***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)
negotiation	-0.180***	-0.180***	-0.178***	-0.177***	-0.176***	-0.183***
	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.015)
scoring rule	0.022	0.014	0.012	0.042***	0.036**	0.024
	(0.015)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)
80,000-208,999		0.010	0.011		0.009	0.008
		(0.008)	(0.008)		(0.007)	(0.007)
209,000-417,999		0.007	0.005		0.002	0.001
		(0.009)	(0.009)		(0.008)	(0.008)
418,000-999,999		0.038***	0.035***		0.027***	0.025***
		(0.010)	(0.010)		(0.008)	(0.008)
≥1,000,000		0.033***	0.028***		0.029***	0.024***
		(0.010)	(0.010)		(0.008)	(0.009)
Maturity			0.003			0.005*
			(0.003)			(0.003)
Budg. Plan			0.052***			0.009
			(0.007)			(0.007)
CPB						0.028***
						(0.007)
Observations	5,690	5,690	5,690	7,490	7,490	7,490

Note: Robust standard errors in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

The most interesting document though is that in panel c). It is part of the documentation attached to auction ARCA-2019-024, a procedure by ARIA to purchase spine prostheses that can be easily found at `https://www.sintel.regione.lombardia.it/`. The particular document from which it has been extracted is the "Progetto di Gara", that provides some background information on the auction (the analysis of regional need that led to the creation of the auction, the legislative context, etc.). However, we believe this kind of document is rare and probably not mandatory, as we have not been able to find it for any other auction nor we could find any piece of legislation referring to it. This extract states that the reserve price has been set on the basis of the last procedure issued by ARCA to purchase spine devices and that this price was set not "too low" in order not to prevent

the consideration of more innovative (thus, more expensive) devices that may have been recently commercialized. At the same time, this reserve price was guaranteed to bring about some savings to the regional accounts.

Therefore, we can claim that the anecdotal insights provided by MASAN researchers on how the reserve price is set are confirmed by formal documents.

Finally, we conclude this section with the illustration of a robustness check of our baseline estimates that is directly connected to the measurement of the reserve price. In particular, the exercise involves replacing in the model specification the continuous variable log(Res.Price) with a series of dummy variables capturing four threshold levels of the reserve price at which special procurement regulations apply. These values are: €80,000, €209,000, €418,000 and €1 million. In Table B.1, we report these estimates: the first three columns are for the LHA sample, while the latter three are for the full sample. In both cases, the main qualitative findings on the effect of consolidation discussed in the main text are confirmed.

C. Robustness of ZIBR model

Table C.1 reports the results from the estimation of equation 9 under different models. These are: ZIBR (the same as Table 2); linear regression estimated through OLS; fractional logit (see Papke and Wooldridge (1996)); beta regression with the outcome variable transformed as Smithson and Verkuilen (2006) indicates (in particular, the corrected response variable is equal to $y^{''} = [(N-1)y^{'} + 0.5]/N$). As we can see, coefficients are similar across all the four models with respect to the sign and the significance levels, reassuring us as to the robustness of our estimates. Some little discrepancies arise when we look at the magnitude of these coefficients. In particular, we can see how the treatment effects under our preferred model, ZIBR, lies somehow in between those estimated under beta regression and the other two models.

A useful tool to assess the validity of the assumptions that our model imposed on the data are quantile residuals.[22] Quantile residuals are robust both under continuous and

[22]These have been introduced by Dunn and Smyth (1996) and are meant to overcome the problems that arise from the use of the other typical GLM residuals, namely Pearson and deviance residual. In many cases, the latter are inspected to assess whether a model can be considered valid for some data, as in standard cases they are supposed to be distributed normally under the true model. However, in other cases, as in discrete outcome models, both Pearson and deviance residuals are far from being normally distributed and they do not provide any meaningful information for model diagnosis (Feng et al., 2017).

Table C.1: Results from the Estimation of Equation (9) under Different Models

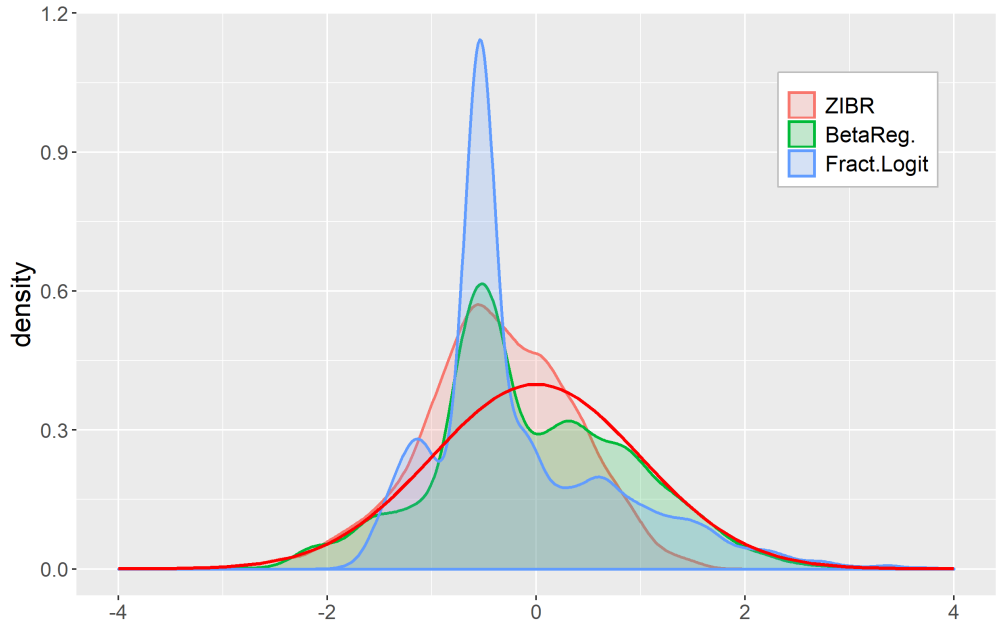
	(1) ZIBR	(2) Lin. Regr.	(3) Frac. Logit	(4) Beta Reg.
Treated	0.076***	0.092***	0.084***	0.032***
	(0.009)	(0.009)	(0.009)	(0.006)
Treated x Post	-0.088***	-0.103***	-0.099***	-0.060***
	(0.012)	(0.012)	(0.013)	(0.007)
First Price	0.035*	0.066***	0.048***	0.077***
	(0.018)	(0.018)	(0.018)	(0.014)
Negotiation	-0.177***	-0.170***	-0.176***	-0.087***
	(0.015)	(0.014)	(0.014)	(0.009)
Scoring Rule	0.013	0.020	0.026*	0.042***
	(0.015)	(0.014)	(0.015)	(0.010)
Maturity	0.003	-0.000	0.001	0.002
	(0.003)	(0.003)	(0.003)	(0.002)
Log Reserve Price	0.010***	0.015***	0.011***	0.010***
	(0.002)	(0.003)	(0.002)	(0.002)
Budg. Plan	0.041***	0.042***	0.038***	0.048***
	(0.007)	(0.008)	(0.007)	(0.005)
Log Personnel	0.010**	0.013***	0.019***	0.009***
	(0.004)	(0.004)	(0.005)	(0.003)
Non-Health Pers. Ratio	-0.362***	-0.298***	-0.289***	-0.157***
	(0.075)	(0.077)	(0.083)	(0.046)
MD Ratio	-0.065**	-0.003	-0.023	0.053***
	(0.029)	(0.030)	(0.032)	(0.017)
Observations	5,690	5,690	5,690	5,690

Note: Robust standard errors are in parentheses. For model (1), (3) and (4) average marginal treatment effects are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

discrete outcome models and they always follow a standard normal distribution under the true model. Figure C.1 plots the quantile residuals for the ZIBR, beta regression and fractional logit models of Table C.1, the three methods we employed that accommodates a response variable defined over the open interval [0,1). A red line indicating a standard normal distribution is displayed for an easy visual inspection.

Firstly, we have to remark that none of the three densities is normally distributed. Indeed, both the Jarque-Bera and the Shapiro-Wilk tests reject the null hypothesis of normality for all the three curves. However, if we compare the relative performance of the three models we end up supporting the superiority of ZIBR over the other alternatives. Indeed, if the fractional logit model fails completely to match a standard normal, beta regression and ZIBR perform relatively better. Among them, ZIBR may be preferred over beta regression as it is affected by a less severe bi-modality. Although we know this vi-

Figure C.1: Distributions of the Quantile Residuals under Different Models



sual assessment is not enough to declare a model superior to another in a formal way, we believe this choice will not affect the conclusions of the analysis, as the results from the estimation of these two last models proved to be very similar.

Next, we repeat the estimation in Table 3, namely the estimation of ZIBR on the pooled prostheses markets using the full sample (LHA+CPB), dropping all the observations with an outcome variable equal to zero. In this way, we aim at assessing whether the substantial share of responses equal to zero in our data has a role in the results we obtained. Table C.2 presents the results. As we can see from the last row, almost 2,000 observations have been dropped. However, the estimates are similar to those in Table 3, declining by just 1 percentage point and remaining significant at the highest conventional level under all 4 models. The usual pattern of the stability of the coefficient after the inclusion of the auction format (from column (2) onward) is observed. These results are reassuring: the zero-outcome observations constitute almost one third of our sample, but their distribution across time and groups does not have a role in shaping the price effect coefficient. This provides more evidence in favor of the independence of the distribution of these kind of observations with respect to other covariates, making our design more robust and the choice of ZIBR much better grounded.

Finally, we complement the latter set of results with those from a standard hurdle

Table C.2: Beta Regression Estimates for the CPB Sample without Zero Responses

Dependent var: Lot Rebate	(1)	(2)	(3)	(4)
Treated	0.068***	0.082***	0.086***	0.086***
	(0.010)	(0.009)	(0.010)	(0.010)
Treated x Post	-0.015	-0.028**	-0.027**	-0.027**
	(0.014)	(0.013)	(0.013)	(0.014)
First Price		0.019	0.017	0.017
		(0.019)	(0.019)	(0.019)
Negotiation		-0.171***	-0.177***	-0.177***
		(0.015)	(0.016)	(0.016)
Scoring Rule		-0.026*	-0.035**	-0.034**
		(0.015)	(0.016)	(0.016)
Maturity			0.003	0.003
			(0.004)	(0.004)
Log Rsv. Price			0.008***	0.009***
			(0.003)	(0.003)
Budg. Plan			-0.014*	-0.014*
			(0.008)	(0.008)
CPB				-0.003
				(0.010)
Observations	4,850	4,850	4,850	4,850

Note: Robust standard errors are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

model. In Table 6, the first four columns consider the case of dependent variable equal to zero when the winning discount is zero and 1 otherwise; the next four columns consider the case of the winning discount as dependent variable, but restrict the sample to the case when the winning discount is strictly positive. For each of the two cases, Table 6 reports the estimates of the same specifications as in columns (2), (3) and (4) of Table 4. The results for the positive winning discount sample – columns from (5) to (8) – are qualitatively the same as our baseline estimates, indicating a negative effect of the merger on winning discounts ranging between 6 and 7 percentage points. Regarding the estimates in column (2)-(4), they indicate that the merger increased the likelihood of observing auctions with zero rebates.[23] Auctions ending with no discount, although not infrequent, represent degenerate cases where the solicitation of competitive bidding has failed. In this sense, the merger did not exacerbate the problem represented by the zero-discount auctions.

[23]The opposite result in column (1) is likely due to the omission of relevant controls on both the auction formats and the buyer characteristics.

Table C.3: Hurdle Model

VARIABLES	(1) Probit	(2) Probit	(3) Probit	(4) Probit	(5) OLS	(6) OLS	(7) OLS	(8) OLS
Treated	0.050**	-0.031*	-0.033*	-0.022	0.117***	0.122***	0.125***	0.126***
	(0.020)	(0.017)	(0.017)	(0.017)	(0.010)	(0.010)	(0.010)	(0.010)
Treated x Post	-0.068***	0.066***	0.057***	0.045**	-0.062***	-0.072***	-0.070***	-0.071***
	(0.025)	(0.021)	(0.021)	(0.021)	(0.015)	(0.015)	(0.015)	(0.015)
First Price		-0.254***	-0.222***	-0.210***		0.021	0.017	0.017
		(0.027)	(0.028)	(0.027)		(0.019)	(0.020)	(0.020)
Negotiation		0.320***	0.297***	0.298***		-0.192***	-0.196***	-0.197***
		(0.026)	(0.027)	(0.026)		(0.017)	(0.018)	(0.018)
Scoring Rule		-0.232***	-0.219***	-0.193***		-0.036**	-0.047***	-0.045***
		(0.025)	(0.025)	(0.024)		(0.016)	(0.017)	(0.017)
Maturity			0.003	-0.004			0.004	0.003
			(0.005)	(0.005)			(0.004)	(0.004)
Log Rsv. Price			-0.016***	-0.010**			0.011***	0.012***
			(0.004)	(0.004)			(0.003)	(0.003)
Budg. Plan			-0.069***	-0.079***			-0.015*	-0.018**
			(0.012)	(0.012)			(0.008)	(0.009)
CPB				-0.116***				-0.012
				(0.014)				(0.010)
Constant					0.182***	0.235***	0.103***	0.096**
					(0.011)	(0.017)	(0.038)	(0.038)
Observations	7,490	7,490	7,490	7,490	4,850	4,850	4,850	4,850

Note: Robust standard errors in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Columns (1) to (4) report the results from a probit model with dependent variable a dummy which takes value 1 if the rebates is equal to 0 and value 0 otherwise. Column (5) to (8) report the coefficient from the estimation of a linear model through OLS with as dependent variable the rebates for the sub-sample of observations for which this is different from 0.

D. Timing of the Merger Effect

In this section, we present a test for causality in the spirit of Granger, following (Angrist and Pischke, 2008). This is clearly not sufficient to establish a causal link, but it brings one more piece of evidence to our analysis.

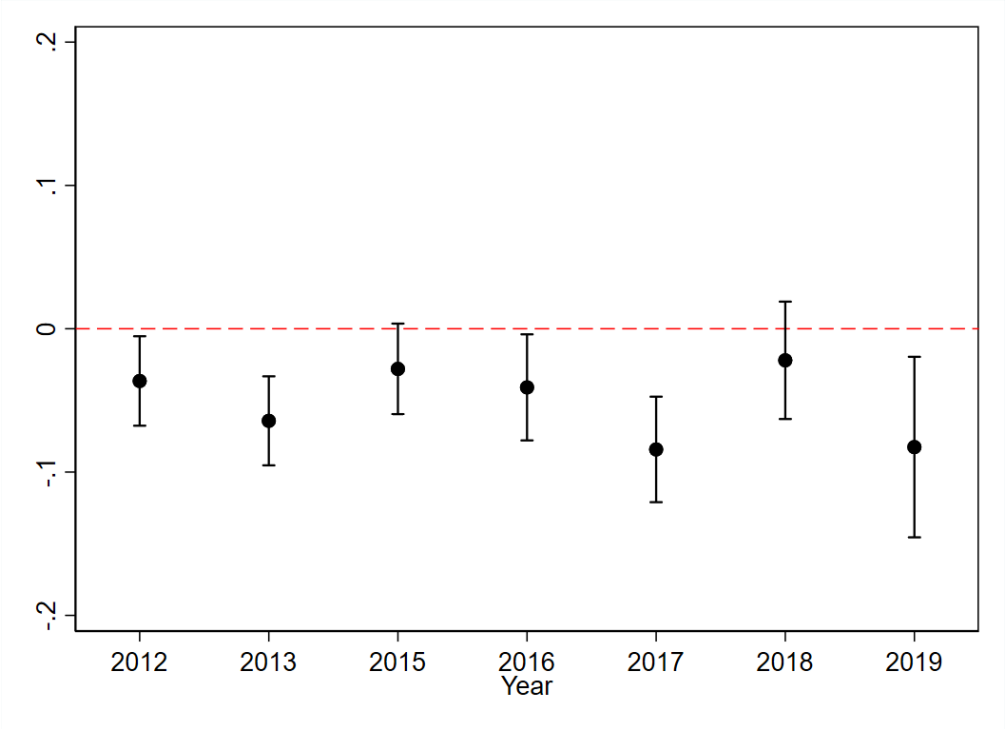
Following Autor (2003) and adapting his strategy to our ZIBR model, we estimate

$$E(y_i|\mathbf{x}_i) = g_1^{-1}(\mathbf{x_i}^\top\beta) = g_1^{-1}\left[\beta_0 + \alpha_t + \beta_1 Treat_i + \sum_{\tau=0}^{4}\beta_{-\tau}D_{i,t-\tau} + \sum_{\tau=1}^{3}\beta_{+\tau}D_{i,t+\tau} + \beta_3 Z_{i,t}\right]$$

where the sums on the right-hand side allow for 4 lags ($\beta_{-1}, \beta_{-2}, \beta_{-3}, \beta_{-4}$) or post-treatment effects and 3 leads ($\beta_{+1}, \beta_{+2}, \beta_{+3}$) or anticipatory effects. To ease notation, we substituted the interaction $Treat_i$ x $Post_t$ by the dummy $D_{i,t}$. If the treatment causes the outcome, then leads should not be significantly different from zero in this model.

Figure D.1 reports the coefficients of leads and lags resulting from the maximum like-

Figure D.1: Timing of the Merger Effect



Note: Estimated impact of the merger for years before, during, and after its occurrence. The black vertical bars are the confidence intervals at 5% significance level. The outcome variable is the percentage rebates. To avoid collinearity we remove the lead for the year preceding the treatment (2014).

lihood estimation of the above model. The pattern of the year-treatment interactions reveals that the treatment effect is pretty constant in the years before the merger and close to zero, while it more clearly moves along a negative trend starting in 2016 with the only exception of 2018 when the estimate bounces back close to zero. Thus, except for 2018, the stability of the level of the coefficients after 2015 and the relatively constant degree of significance suggests that something is negatively affecting the response variable in the years after 2015, and to our knowledge the merger offers the best explanation.