Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

The Consequences of Medicare Pricing: An Explanation of Treatment Choice

Elena Falcettoni

2020-063

Please cite this paper as:
Falcettoni, Elena (2020). "The Consequences of Medicare Pricing: An Explanation of Treatment Choice," Finance and Economics Discussion Series 2020-063. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2020.063.

The Consequences of Medicare Pricing: An Explanation of Treatment Choice

Elena Falcettoni *

Abstract

Primary care physicians (PCPs) provide more specialty procedures in less-urban areas, where specialists are fewer. Using a structural random-coefficient model and the demographic and time variation in the data, this paper shows that changes in policy-set reimbursements lead to a reallocation of the suddenly-more-remunerative procedures away from specialists and toward PCPs, and this effect is stronger, the more rural an area is. A reimbursement-unit increase for a given procedure leads to outside-metro PCPs gaining 7-15% market share more than metro PCPs in that procedure, at the expense of specialists. Small metropolitan areas and very rural areas are the most affected.

JEL codes: I18; I13; J2; R12.

Keywords: Primary care physicians; Specialists; Specialty procedures; Rural; Reallocation; Medicare; Fee-for-service.

* Board of Governors of the Federal Reserve System (email: elena.falcettoni@frb.gov). The author would like to thank Francesca Carapella, Mariacristina De Nardi, Nathan Foley-Fisher, Thomas Holmes, Hannah Neprash, Vegard Nygaard, Amil Petrin, James Schmitz, Jr., and Joel Waldfogel for their thoughtful comments and support. The author would also like to thank the participants of the Industrial Organization Workshop at the University of Minnesota, the APPAM International Conference 2020, the American Society of Health Economics Conference 2020 and 2018, the Warwick Ph.D. Economics Conference 2018, the Midwest Economics Association Conference 2018 and 2017, and the Missouri Valley Economic Association Conference 2017 for the great feedback and suggestions received. All remaining errors are the author's only. The views expressed herein are those of the author's and do not represent the views of the Board of Governors of the Federal Reserve System or any person associated with the Federal Reserve System.

1 Introduction

Fee-for-service payments account for about half of physicians' income in the United States. Not only have FFS reimbursements been historically higher for specialty procedures than for typical primary-care procedures, but the gap in reimbursements has also been increasing over the years. Since 1) the reimbursement amount does not depend on the specialty of the physician who carries out the procedure, but only on the procedure itself, and 2) specialty procedures have been compensated increasingly more over the years than typical primary care procedures, this payment system generates financial incentives for primary care physicians (PCPs) to substitute primary care procedures for more specialized, remunerative procedures whenever possible. This paper focuses on a particular margin through which financial incentives arise: the urbanity level of the area where the physician practices, that is, how rural or urban the location is. In particular, I analyze how increases in the reimbursement rates for more-specialized procedures induce PCPs to carry out the procedures themselves rather than referring the patient to a specialist, and shows that PCPs are able to do so more, the more rural the area, because of the lower level of competition coming from specialists in rural areas.

This paper makes two contributions: 1) it provides evidence of the importance of the urbanity margin when financial incentives for PCPs arise, and 2) it provides evidence of the reallocation effect that such incentives have on the specialty that carries out a given procedure, as PCPs respond to such incentives *at the expense of the specialists* within the same geographical market of health care, because rural areas, with fewer specialists, provide less competition. To do so, this paper studies how the financial incentives generated by fee-for-service affect the demand for specialized procedures due to physicians inducing demand for the more remunerative, specialized procedures. In particular, PCPs see patients first and decide whether to refer the patient to a specialist. The reimbursement rate, therefore, generates a financial incentive that affects the decision of PCPs to pass the procedure to a specialist rather than performing it themselves. If this is true, the data should reveal that PCPs are better able to take up a higher share of specialty procedures, the more rural the area, where fewer specialists practice. This paper provides evidence for this novel fact through data analysis, which is the first contribution of this paper.

Second, to understand the reallocation of procedures among specialties discussed in this paper, it is important to realize that the PCP take-up rate in the suddenly-more-remunerative procedures comes at the expense of specialists. Academics and policymakers alike have contended that fee-for-service payment schemes generate financial incentives for physicians. For example, case studies have found supporting evidence that fee-for-service leads to overutilization of procedures without an underlying reason to be found in patient health conditions (see, i.a., Ginsburg 2011; Hoangmai and Ginsburg 2007; Levin & Rao 2004, 2008). Nevertheless, the focus on a particular set of procedures and/or specialties overlooks the reallocation of these procedures between the specialty groups as procedures become more remunerative. If this effect is not included, the entire take-up rate could be incorrectly attributed to overutilization, overstating the effect of the latter. Therefore, this paper provides evidence that it is key to consider two complementary effects caused by fee-for-service payment schemes: 1) a reallocation of the procedures carried out from specialists to PCPs for any given total number of procedures, as shown here, and 2) the increase in the number of total procedures carried out overall without an apparent increase in patients' health issues, that is, overutilization, which is not analyzed in this paper.

I use physician-procedure-level data from Medicare Part B throughout the paper. Medicare, the public health insurance for the elderly, accounts for 20 percent of total health spending and uses a fee-for-service reimbursement scheme for the original Medicare. Note, however, that the results in this paper are likely to generalize to fee-for-service payment schemes outside of Medicare because the Medicare reimbursement system has been documented to influence the payment system of about 80 percent of physicians in the country (Clemens & Gottlieb 2017).

I start by studying how the procedures that PCPs perform vary along the urbanity index. To do so, I create a specialization index of procedures using the data on all procedures carried out by all physicians participating in Medicare Part B from 2012 to 2015. I then classify procedures based on the type of physicians who would usually carry them out and focus on the procedures that can be carried out by all specialty groups. For example, the administration of vaccines is only executed by PCPs, while the removal of bones and bone parts is only done by specialists, and therefore both of these procedure types would be excluded; contrarily, EKGs are carried out by both PCPs and specialists, and are therefore included in the analysis. In particular, this paper focuses on treatments carried out by specialists 60-80 percent of the time and by PCPs the remaining 20-40 percent of the time.[1] The data show that PCPs take on more specialty procedures, the more rural the area. I then run reduced-form tests to provide empirical evidence that the number of specialty procedures carried out by PCPs and their probability to carry them out is affected by changes in the reimbursement amount. I find that an increase in one unit of reimbursement for a given procedure—equal to an increase in reimbursement of about $36—leads to the PCP carrying out that procedure three more times on average and over a 45 percent higher probability of PCPs providing a higher number of the more remunerative procedure.

A recent paper that addresses the incentives generated by fee-for-service is Clemens and Gottlieb (2014). Their analysis uses a 1997 change in Medicare geographical areas and focuses on two main procedures: MRIs and procedures for cardiac patients, such as cataract surgeries. They find that a two-percent increase in payment rates leads to a three-percent increase in care provision and that elective procedures respond more strongly than non-discretionary services. This paper confirms, complements, and generalizes their analysis along several dimensions. Notice that, among the physicians considered in this paper, the median procedure carried out earns the physician a total of 36.45 relative value units, so that a unit increase is equivalent to a 2.7 percent median increase in compensation. Out of those PCPs who perform specialty procedures, performing a specialty procedure three more times is equal to a three-percent increase in the median number of specialty procedures they perform. Therefore, these results suggest that the median PCP displays an elasticity of 1.11 with respect to changes in the reimbursement rates. Therefore, the results suggested by this paper fit well within those found by Clemens and Gottlieb (2014). Notice that removing the procedures discussed in Clemens and Gottlieb (2014) does not dramatically change these values, but leads to a slightly higher elasticity of 1.2, indicating that the results of these papers are not driven only by the procedures analyzed in Clemens and Gottlieb (2014), but that they extend their results to a wider range of procedures.

Finally, I turn my attention to the estimation of the demand curve to be able to model the ability of PCPs to gain market share in specialty procedures at the expense of specialists as areas become more rural. I set up a structural random coefficients model of demand for healthcare where the random coefficient is estimated using the variation in patients' demographic and health measures to model patients as precisely as possible.

[1] Robustness checks, available in the online appendix, show that the results are robust independently of the range chosen, but their effect is stronger, the more specialized the range of procedures is.

The innovation in the model lies in the inclusion of a supply-induced demand mechanism in which PCPs are able to affect the consumers' decision of the specialty-procedure combination. This is motivated by PCPs' ability to offer a treatment to the patient rather than refer the patient to a specialist. In the model, this supply-induced demand mechanism is specified through the inclusion of policy-set reimbursement fees for physicians in the consumers' demand. The identification of the parameter of the reimbursement rate relies on consumers not reacting to the amount of reimbursement the physician receives. This identification strategy would fail if patients' demand for procedures directly responded to changes in the reimbursement amounts. This could be the case if, for instance, patients were responsible for paying for the procedures entirely out of pocket, which amount would clearly vary according to the reimbursement rates. Similarly, this identification strategy could also fail if the reimbursement amounts varied according to the specialty of the physician who provided the procedure. The data allow me to overcome this challenge because all patients considered are insured by Medicare for the procedures analyzed and are therefore responsible for only a co-pay for most procedures, and for at most 20 percent of the cost otherwise as co-insurance. Importantly, all patients would face the same co-insurance amount across all providers considered; the reimbursement units are based on the procedure carried out and not on the specialty of the physician, which therefore limits the patients' ability to respond to their physician's reimbursements. Finally, while patients on Medicare are able to self-refer to specialists, notice that this fact does not impact the results of this paper. As long as self-referral patterns do not change wildly over time, the paper will correctly identify PCPs' ability to gain market share in specialty procedures in more-rural areas at the expense of specialists, based on the set of patients they would see. One would expect the ability to self-refer to be more important in cities, where more specialists are present. While this could explain the higher initial share in specialty procedures carried out by rural PCPs compared with their most urban counterparts, it would not explain the changes in PCPs' market shares in the same procedures over time according their remuneration, as patients would not change their choice to self-refer based on changes in reimbursement fees. The ability to self-refer however indicates that the results in this paper are a lower bound on the reallocation effect for the physician population as a whole, because patients who cannot self-refer according to their insurance contract might be more subject to PCPs' incentive to carry out the procedure in lieu of referring them to specialists.

Second, similarly to the estimation of demand curves in other industries, the endogeneity of the price paid by patients could lead to a biased estimate of the price coefficient. I observe the amount billed by physicians and the amount reimbursed by Medicare, from which I can impute the price faced by patients. Similarly to a higher-quality product in another industry, a provider could be more popular despite a higher co-pay because of, for example, friendliness, word-of-mouth reputation, or personal attachments. This fact would bias the price coefficient upward, indicating a lower elasticity of patients to prices, when instead the choice of the higher-rate physician could be due to factors that are unobserved by the econometrician. To address this, I use instrumental variables which represent input costs. The instruments I use are malpractice insurance reimbursement fees (which proxy the costs for malpractice insurance) and Medicare-set cost-of-living adjustments (for labor, practice expenses, and malpractice insurance) to proxy for the regional variation in input costs. These costs would increase a physician's pricing menu but not be observed or considered by patients, which is the identifying assumption necessary for the instruments to be valid. In the specifications that utilize the demographic variation in the data for the estimation of the random coefficient, I supplement these instruments with functions of the same instruments, as common in the literature.

The structural results confirm the hypothesis that PCPs increase their share in specialty procedures at the expense of specialists as the procedures become more remunerative, and PCPs are able to do so more, the more rural the area. I find that the same reimbursement-unit increase for a given procedure leads to an increase in the primary care physician share in that specialty procedure by 7-15 percent more in less-urban areas compared with their most-urban counterparts, at the expense of specialists. These results therefore provide evidence that 1) the urbanity of the area is an important factor for PCPs' ability to benefit from the financial incentives generated by fee-for-service payment schemes and 2) PCPs respond to these incentives at the expense of specialists, leading to a reallocation of procedures among the two specialty groups. I find that small metropolitan areas (with a population between 50,000 and 250,000 people) and very rural areas (with a population smaller than 10,000 people) are the most affected.

The rest of the paper is structured in the following way: Section 2 reviews the existing literature. Section 3 introduces the data used in the paper. Section 4 introduces the model. Section 5 provides the Reader with some introductory data analysis. Section 6 discusses the empirical evidence and reduced-form results. Section 7 presents the structural results and discusses the key parameters. Section 8 concludes and discusses future research.

2 Literature

This paper primarily contributes to the strand of health economics literature discussing how physicians respond to financial incentives. Paper and articles have documented an increase in the demand for health goods coming from consumers and an increase in defensive medicine, i.e., physicians' decision to request treatments and procedures out of fear of being sued. Nevertheless, defensive medicine has been shown to only account for 0.46 percent of health spending (Anderson et al. 2005). Consumers, on the other hand, might indeed enjoy a higher consumption of health goods, but it would be puzzling if the entirety of spending was a direct effect of consumer preferences. First, it is difficult to imagine that US consumers, who face higher co-pays than Europeans, are so intrinsically different in their preferences that their utility maximization would lead to higher consumption of more-expensive health goods without any visible health effects. Second and more importantly, this would mean that US consumers switched their preferences quite dramatically, since this unbounded increase in health spending is quite recent. The other main school of thought suggests that the increase in medical spending can be explained by the rise in income because medical spending is a luxury good (Hall and Jones 2007). However, the assumption that medical spending is a luxury good is inconsistent with micro estimates showing that the income elasticity of medical spending is about 0.7 (Acemoglu et al. 2013).

A recent paper that addresses the incentives generated by fee-for-service is Clemens and Gottlieb (2014). Their analysis uses a 1997 change in Medicare geographical areas and focuses on two main procedures: MRIs and procedures for cardiac patients, such as cataract surgeries. They find that a two-percent increase in payment rates leads to a three-percent increase in care provision and that elective procedures respond more strongly than non-discretionary services. This paper confirms, complements, and generalizes their analysis along several dimensions. First, it is able to make more general statements on the reimbursement effect on treatment, without focusing on few case studies, but utilizing both the annual changes in reimbursement rates and the differences in the urbanity of the areas. Second, it documents and analyzes how PCPs specifically are able to increase the number of specialty procedures they provide according not only to the reimbursement level, but to the urbanity level of the area where they practice, leading to a reallocation of specialty procedures from specialists to PCPs. Notice that, among the physicians considered, the median procedure carried out earns the physician a total of 36.45 relative value units, so that a unit increase is equivalent to a 2.7 percent median increase in compensation. The reduced-form results in this paper suggest that a unit increase in the reimbursement rate leads, on average, to the physician carrying out that procedure three more times. Out of those PCPs who perform specialty procedures, performing a specialty procedure three more times is equal to a three-percent increase in the median number of specialty procedures they perform. This leads to an elasticity of about 1.11. Therefore, the results suggested by this paper fit well within those found by Clemens and Gottlieb (2014). Notice that removing the procedures discussed in Clemens and Gottlieb (2014) does not dramatically change these values, but leads to a slightly higher elasticity of 1.2, indicating that the results of these papers are not only driven by the procedures analyzed in Clemens and Gottlieb (2014).

Finally, the use of a structural approach to analyze this issue modeling a supply-distorted structural demand system is on its own a novel approach.

There is an extensive branch of literature analyzing, from a reduced-form perspective, physicians' response to financial incentives in a hospital setting (i.a. Acemoglu and Finkelstein 2008, Finkelstein 2007, American Hospital Association 2008) and in a managed care setting (i.a. Lori 2009). However, the analysis of physicians' responses in their own practices seems to have been overlooked, even though the level and growth of healthcare spending for physicians is not negligible, equaling about 60 percent of hospital care expenditures, or 37.5 percent of total aggregate health care expenditures in the US.

There are also different papers analyzing physicians' response to financial incentives for particular procedures (see, i.a., Gruber & Owings 1994; Grant 2009; Shrank 2005; Jacobson 2006). This paper complements this strand of literature by setting up a structural model of physicians' behavior without confining to a few procedures as specific case studies. This paper generalizes these previous analyses and focuses on how the treatment picked is inherently tied to the profitability of the treatment itself and on the physician's location.

Papers have also been extensively written on the fee-for-service system's impact on health care costs and service overutilization (see, i.a. Ginsburg 2011; Hoangmai and Ginsburg 2007; Levin & Rao 2004, 2008). This paper complements this literature by providing evidence that a change in financial incentives also leads to a reallocation of specialty procedures away from specialists in favor of PCPs. Focusing on a particular set of procedures and/or specialties overlooks the reallocation of these procedures between the specialty groups as procedures become more remunerative. If this effect is not included, the entire take-up rate could be incorrectly attributed to overutilization, overstating the effect of the latter. Therefore, this paper provides evidence that it is key to consider two complementary effects caused by fee-for-service payment schemes: 1) a reallocation of the procedures carried out from specialists to PCPs for any given total number of procedures, as shown here, and 2) the increase in the number of total procedures carried out overall without an apparent increase in patients' health issues, that is, overutilization, which is not analyzed in this paper.

Finally, the evidence of regional variations in Medicare spending has also been mentioned, not only in the health economics literature, but also in the media (Gawande 2007 is probably the predominant example among many) and in the medical literature (see Fisher 2003 for example). This paper makes a contribution in this regard by providing a reason for such variation, i.e. the link between PCPs' location and their ability to respond to financial incentives and to do so more, the more rural a location is.

Methodologically, this paper bases itself mostly on Berry, Levinsohn, Pakes (1995, hereafter: BLP). The Medicare database does not actually provide data on the single patients, but it provides cumulative data on the patients seen by each physician for each procedure carried out, including the number of patients seen for each procedure and the characteristics of these patients, such as their demographics, their race and gender, as well as many health measures on illnesses and diseases, such as the number of patients with cancer. This allows for a simulation of individuals that match the observed characteristic distributions in the data.

Readers who are interested in a deeper discussion on the institutional framework should refer to Appendix Section A.1.

3 Data

The primary source of data for this paper comes from the Centers for Medicare & Medicaid Services (CMS). The Physician and Other Supplier Public dataset provides information on services and procedures provided to Medicare beneficiaries by physicians. It contains information on utilization, actual Medicare reimbursement, and submitted charges. Each line of the dataset is indexed by a National Provider Identifier (NPI), which identifies each physician in the dataset, by a Healthcare Common Procedure Coding System (HCPCS) code, which identifies every procedure carried out by each physician, and by the place of service, indicating whether the procedures were carried out in a facility setting. The data are based on information from CMS administrative claims data for Medicare beneficiaries enrolled in the fee-for-service program. The data cover calendar years 2012 through 2015 and contain the universe of physicians taking part in Medicare Part B for the fee-for-service population. There are a little over 26.3 million observations in the dataset across over a million of physicians.

Despite the wealth of information on payment and utilization for Medicare Part B services, the dataset has a number of limitations. The data may not be representative of a physician's entire practice, because they only include information on Medicare fee-for-service beneficiaries. However, since Medicare influences the payment system of 80 percent of physicians (Clemens & Gottlieb 2017), these data allow for the analysis of physicians' behavior under this payment mechanism, which is then relevant for the greatest majority of physicians in the country. In addition, the data are not intended to indicate the quality of care provided and are not risk-adjusted to account for differences in underlying severity of disease of patient populations. To counter the lack of risk adjustment, demographic data on patients' riskiness and incidence of diseases will be included in the estimation. Despite these limitations, some positive characteristics should be highlighted. First of all, the fact that all beneficiaries are covered by Medicare eliminates the issues related to the status of insurance of the beneficiaries. In particular, it allows me to abstract from other endogenous characteristics related to the insurance status of beneficiaries which would arise if a full dataset (not Medicare only) were used. Moreover, it also allows me to ignore the network effects of different insurance policies as well as their different payment plans. In practice, therefore, this dataset provides a homogeneous universe of insured individuals that only differ by the conditions they have and by the treatments they receive, and not by their insurance situation.

The illness and disease distribution is reflected in the average beneficiary risk scores provided on the "Medicare Physician and Other Supplier Aggregate Table" (i.e., one record per NPI). These data provide information on the health status of the beneficiaries the providers serve for every year of interest together with the rate of incidence of a number of diseases and illnesses among the patients seen by each physician for every year. This information can account for the average health of each physician's patient.

Finally, the Metropolitan Statistical Area definition follows the U.S. Census Bureau definition for Primary Metropolitan Statistical Areas (PMSA). The urban/rural classification also follows the U.S. Census Bureau definition according to the 2010 Census criteria. In particular, an urbanity index equal to 1 indicates a large central metro, i.e. counties in MSAs with a population equal to or greater than 1 million that contain the entire population of the largest principal city of the MSA, or at least 250,000 inhabitants of any principal city of the MSA; an index equal to 2 indicates a large fringe metro, i.e. counties in MSAs with a population equal to or greater than 1 million that do not qualify as large central metros; 3 indicates a medium metro, i.e. counties in MSAs with a population equal to or greater than 250,000 but strictly smaller than 1 million; 4 indicates a small metro, i.e. counties in MSAs with a population equal to or greater than 50,000 but strictly smaller than 250,000; 5 indicates a micropolitan area, i.e. counties in micropolitan statistical areas with a population equal to or greater than 10,000 but strictly smaller than 50,000; and 6 indicates a noncore area, i.e. the most rural classification, with a population that is strictly smaller than 10,000.

4 Model

This section models the ability of PCPs to gain an increasing share in specialty procedures along the urbanity index, at the expense of specialists. In particular, PCPs positively reach to increases in remuneration of a given procedure by increasing the number of times they carry that procedure out instead of referring it to specialists, and PCPs are able to do so more, the more rural the area. This section presents a structural random coefficients model with a supply-induced demand mechanism generated by the physician's utility for reimbursements entering the consumer's utility. The Appendix Section A.2 presents the model setup for the standard multinomial logit model of demand, which is a particular case of the model described here, where stochastic and demographic coefficients are not included. The main difference between the logit demand and the full structural model introduced here is that the marginal utilities of the product characteristics implied by the full model are different across consumers, and determined by the consumer characteristics. This breaks the independence of irrelevant alternatives (IIA) property typical of logit models, from which the simplified model in the Appendix would suffer.

I consider the choice of going to a specialist as the outside option. I assume that at this stage, the physician has already decided what treatments to offer and the patients already know what treatment they need to receive. Therefore, consumers are faced with a product choice given by the type of physician performing a relevant procedure for them.

Consumers' utility is composed of a mean utility and random stochastic coefficients dependent on demographic variables. I assume that the consumer maximizes a weighted sum of utilities, her own and her primary care doctor's, where the doctor's utility only enters the patient's utility in terms of the reimbursement obtained for the procedure carried out, i.e.

$$\begin{cases} u_{ijt} = & \max_j \left\{ \sum_{urban=1}^{6} (1-\gamma_{urban})\, u_{ijt}^{patient} + \mathbb{I}_{\text{PC}} \sum_{urban=1}^{6} \gamma_{urban} u_{jt}^{PC} \right\} \\ u_{jt}^{PC} = & RVU_{jt} \;\; \forall \text{ procedure codes} \\ \mathbb{I}_{\text{PC}} = & 1 \text{ if physician is in primary care (PC)} \end{cases} \tag{1}$$

$u_{ijt}^{patient}$ identifies the consumer/patient's utility while u_{jt}^{PC} identifies the doctor's utility. I define each market t as a Metropolitan Statistical Area (MSA)-year combination. Once the data are cleaned and markets with a single choice are excluded, I am left with 193 MSAs and a total of 764 markets (almost all MSAs are present for all 4 years of interest). Data are available for four consecutive years: 2012, 2013, 2014, 2015. Only PCPs are able to influence the consumers' utility because PCPs see the patient first and can choose to carry out the specialty procedure themselves instead of referring the patient to a specialist.[2] Therefore, $\mathbb{I}_{\text{PC}} = 0$ eliminates the supply-driven demand mechanism. The extent to which physicians respond to the reimbursements depends on the urbanity (this is why γ_{urban} depends on the urbanity). The urbanity index is defined by the Census according to the population of an area and proximity to an MSA. In practice, I will report estimates for γ_{urban} for physicians in less urban areas compared with their most urban counterparts, i.e. physicians in large metropolitan areas with a population that is larger than one million and which fully contain an MSA. A product j is the combination of a procedure, indexed by its respective HCPCS billing code, and the provider group that carries it out, defined as Primary care, Laboratory, Emergency medicine & general surgeons, and Specialists. As usual, the utility of the outside good (specialists) is normalized to zero.

Writing out the variables, the utility is equal to

$$u_{ijt} = \sum_{urban=1}^{6} (1-\gamma_{urban})\left(-\alpha_i p_{jt} + x_{jt}\beta_i + \xi_{jt} + \eta_{ijt}\right) + \mathbb{I}_{\text{PC}} \sum_{urban=1}^{6} \gamma_{urban} RVU_{jt} \tag{2}$$

[2] This assumption assumes away the possibility of patients referring themselves directly to a specialist. Self-referral would be the case in cities more than elsewhere, as specialists are more present in metro areas. This could partially explain the initial higher shares of more-rural PCPs in specialty procedures compared with their urban counterparts. This initial differential would then be included in what I refer to as the "competition" coming from specialists in cities. In other words, PCPs in cities would initially carry out fewer EKGs, for example, both because more patients would go to cardiologists directly and because out of those patients that would see a PCP first, a higher percentage would ask for a referral to a specialist even if offered an EKG by the PCP. However, the use of the time variation where the reimbursement increase is the mechanism that leads to a widening of this initial differential, with rural PCPs gaining a higher share of their urban counterparts. Therefore, even if patients could be part of the reason of why, at the beginning of the period, PCPs are able to do more specialty procedures in more rural areas, they would not choose whether to refer themselves based on changes in the reimbursement fees set by policy across time.

which can be written as:

$$u_{ijt} = \begin{cases} \sum_{urban=1}^{6} (1-\gamma_{urban}) (\alpha p_{jt} + \beta x_{jt} + \xi_{jt}) \quad \text{patients' mean utility} \\ +\mathbb{I}_{\text{PC}} \sum_{urban=1}^{6} \gamma_{urban} RVU_{jt} \quad \text{physician's utility within mean utility} \\ +(-p_{jt},\, x_{jt}) \sum_{urban=1}^{6} (1-\gamma_{urban}) (\Pi D_i + \Sigma \nu_i) \quad \text{stochastic coefficients} \\ +\epsilon_{ijt} \quad \text{iid error term} \end{cases} \tag{3}$$

where $(p_{jt},\, x_{jt})$ are the product characteristics discussed below, ξ_{jt} are product-market unobservables, D_i is a 5x1 vector of consumer i's observable demographic characteristics (age, gender, risk score proxying health issues, income in thousands, and square of the income in thousands), v_i is a 10x1 vector of the effect of consumer i's unobservable characteristics on α_i and β_i parameters; Π is a 10x5 matrix of how α_i and β_i parameters depend on the consumer observables, Σ is a 10x10 matrix of how those parameters depend on the unobservables; and $(\nu_{i\alpha},\, \nu_{i\beta})$, $(\Pi_\alpha,\, \Pi_\beta)$, $(\Sigma_\alpha,\, \Sigma_\beta)$ split the vector into two parts. The values of D_i are picked from the main database as draws of 20 random individuals in each market. The values of ν_i are drawn from a multivariate normal distribution and are independent and identically distributed. The ϵ_{ijt} are drawn from Type 1 extreme value distribution and are independent and identically distributed across individuals, products, and markets.

I estimate four specifications: a random coefficients model without the use of demographic variables, but with a stochastic element on the reimbursement variable (in other words, the D_i vector is set to zero), a random coefficients model with demographic variables (the full version mentioned in this section), and a random coefficients model with and without demographics with the use of Chamberlain optimal instruments. The full specification of the model, a random coefficients model in which the random coefficient is estimated using the demographic variation present in the data, is the one presented here. The other specifications are all particular cases of this general one. Chamberlain optimal instruments are functions of the characteristics and instrumental variables, as commonly done in the literature.

I include ten product characteristics, equal to a constant and nine characteristics $(p_{jt},\, x_{jt})$ at a provider group-procedure-market level. These control for the average price to receive the procedure in that market carried out by that provider group, the relative value units earned for that procedure, a proxy for the level of specialization of the physician, the urbanity level, a proxy for the average level of education, and a proxy for competition. Price p_{jt} is derived from the total amount billed by the provider minus the amount reimbursed by Medicare. The relative value units earned for that procedure, which are directly observed, capture the physician's ability to influence demand and are therefore the main variable of interest in this paper. The average level of specialization of all procedures performed by physicians within each group measures the level of specialization of the procedures carried out by the average physician in that market. The average urbanity level of the physicians in that provider group and market, defined for each product, simply captures whether, on average, the procedure-provider group combination is more common in a more urban or a more rural setting. The average Medical Doctor (M.D.) index gives an index of whether the majority of physicians considered in that product-market combination have an M.D. degree. Given that each physician can either be a M.D. (1) or not (0), this variable is between 0 and 1.[3] The proxy for competition is calculated as the

[3] The other type of degree that physicians can hold is a D.O. degree, which categorizes them as osteopathic physicians. Physicians with a D.O. are licensed in all 50 states to practice medicine and surgery, as well to prescribe medications. Anecdotal evidence seems to suggest that patients often display a preference for M.D. physicians, whom they consider better qualified. This

average number of physicians in each physician's own provider group as well as the number of physicians in other provider groups in any given zip code. Therefore, four variables proxy competition: the average number of PCPs, specialists, emergency medicine physicians, and laboratories in close proximity to each physician considered. A higher number indicates higher competition in that product-market combination.

As is usually done in the literature, the utility can be simplified by calling the mean utilities δ's, for the patients and for the PCP respectively, a stochastic coefficient component μ, and the random error:

$$u_{ijt} = \delta_{jt}^{patient} + \mathbb{I}_{\text{PC}}\delta_{jt}^{PC} + \sum_{urban=1}^{6} (1 - \gamma_{urban})\, \mu_{ijt} + \epsilon_{ijt} \tag{4}$$

The market share of product j for consumer of type i in market t is then equal to

$$s_{ijt} = \frac{\exp\left\{\delta_{jt}^{patient} + \mathbb{I}_{\text{PC}}\delta_{jt}^{PC} + \sum_{urban=1}^{6} (1 - \gamma_{urban})\, \mu_{ijt}\right\}}{1 + \sum_{k=1}^{1175} \exp\left\{\delta_{kt}^{patient} + \mathbb{I}_{\text{PC}}\delta_{kt}^{PC} + \sum_{urban=1}^{6} (1 - \gamma_{urban})\, \mu_{ikt}\right\}} \tag{5}$$

and the overall market share of product j in market t can be found by integrating the individual market shares across the individual types, while weighing each type according to its probability in the population. $\hat{P}(D)$ denotes the empirical distribution of the demographic characteristics and $\mathcal{N}(\nu)$ the distribution of the unobserved characteristics:

$$\begin{aligned} s_{jt} &= \int_{\nu}\int_{D} s_{ijt}\mathrm{d}\hat{P}(D)\,\mathrm{d}\mathcal{N}(\nu) \qquad (6) \\ s_{jt} &= \int_{\nu}\int_{D} \frac{\exp\left\{\delta_{jt}^{patient} + \mathbb{I}_{\text{PC}}\delta_{jt}^{PC} + \sum_{urban=1}^{6} (1 - \gamma_{urban})\, \mu_{ijt}\right\}}{1 + \sum_{k=1}^{1175} \exp\left\{\delta_{kt}^{patient} + \mathbb{I}_{\text{PC}}\delta_{kt}^{PC} + \sum_{urban=1}^{6} (1 - \gamma_{urban})\, \mu_{ikt}\right\}} \mathrm{d}\hat{P}(D)\,\mathrm{d}\mathcal{N}(\nu) \end{aligned}$$

The demographic characteristics are picked from the Medicare dataset. In order to proxy for patients, I use the patient cumulative data available for each physician as a representative consumer for each product (specialty-procedure) combination. Since products are defined at a wider level than physicians, thousands of patient proxies are available for each market. I then sort them randomly and pick twenty observations per market of the following characteristics: per capita income (which varies with every zip code), per capita income squared, average risk score (a variable created by Medicare, which measures the general "healthiness" of patients considering many different elements, from diabetes to cancer), gender (as the percentage of patients that were female seen by each physician), and age (as the average age of patients seen by each physician).

5 Data Analysis

To be able to analyze the margins along which PCPs carry out specialty procedures, I first need to define what constitutes a specialty procedure. To do so, for every procedure, I look at the number of times the procedure is carried out by specialists and PCPs, respectively, over the entire dataset (about 26 million

variable controls for this.

Table 1: Percentiles of degree of specialization (the lower the number, the more specialized)

	Percentile
25th percentile	52.06%
50th percentile	63.44%
75th percentile	72.11%
90th percentile	80.99%

Notes: The variable of interest is an indicator variable, from 0 to 1, indicating the frequency of performance of each procedure by PCPs and specialists. This table uses the average of this variable, for each physician, across all procedures carried out by that physician. A value of 0 indicates that a physician performs only specialty procedures, a value of 1 indicates that a physician only performs primary care procedures. This table only shows the degree of specializations for the physicians who perform at least one specialty procedure.

observations across four years). In Figures 1 and 2 and in supplementary results available upon request, I also analyze what percentage of physicians in primary and specialty care, respectively, carry out any given procedure. I then consider the procedures of interest to be those performed by specialists 50-80 percent of the time and by the primary care the remaining 20-50 percent. Robustness checks show that the results are robust independently of the range chosen, but their effect is stronger for tighter ranges. This approach creates a specialization index for each procedure, from 0 to 1. An index value equal to 0 means that the procedure is only carried out by specialists, while an index value of 1 means that it is always carried out by primary care. Therefore, the lower the index value, the more specialized the procedure is.

Having built this index, I can analyze whether any physicians, in general, performs procedures within their specialty. I observe some clustering in the procedures carried out, with PCPs carrying out either mostly primary care procedures (with an average specialization index of at least 0.6) or almost exclusively specialty procedures (with an average specialization index between 0.3 and 0.4). Among those PCPs who perform at least one specialty procedure, it is also observed that they tend to perform a large number of specialty procedures. The median procedure carried out by these PCPs, as shown in Table 1, is one carried out by specialists 60 percent of the time.

It is clear from this evidence, as well as other data analysis reported in the online Appendix, that some PCPs are able to carry out a large number of specialty procedures. The margin through which they are able to do so, which is the main theory analyzed in this paper, is how urban or rural the area where the physician practices is.

To be able to understand this better, I focus on procedures carried out by PCPs 20-50 percent of the time and look at the distribution of physicians and services across the urbanity index, by specialization, as shown in Figure 1.

For robustness, I restrict my attention to highly specialized procedures, i.e. the procedures carried out by specialists 70-80 percent of the time and once again look at the distribution of physicians and services across

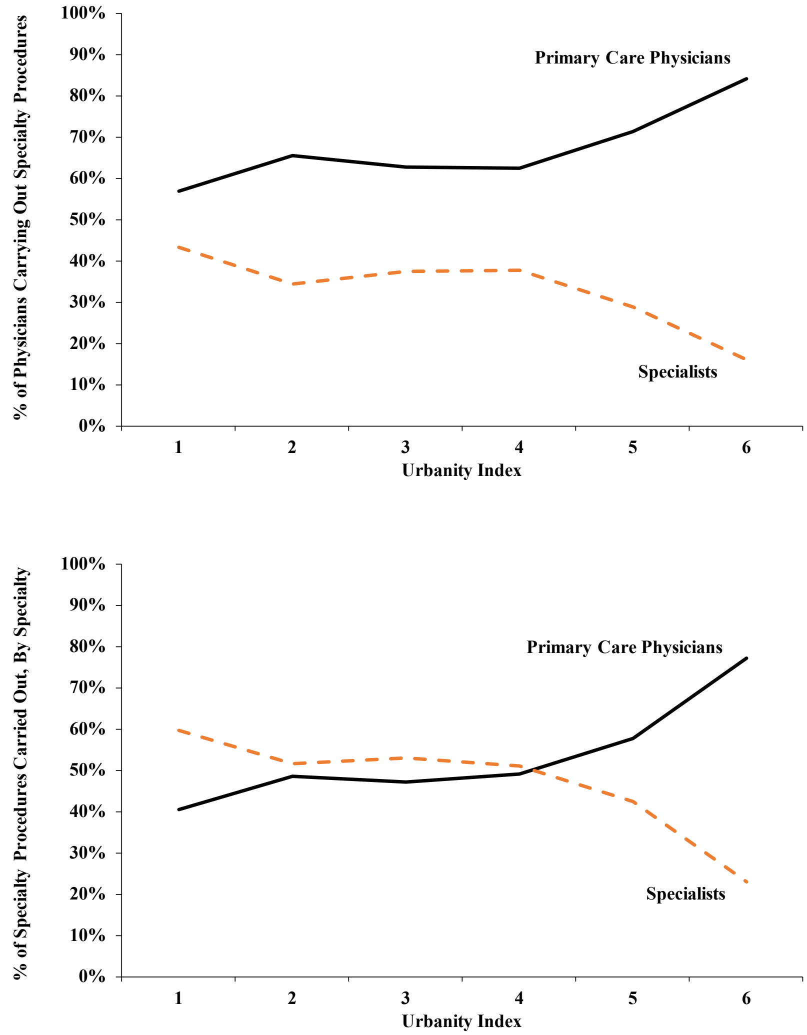

Figure 1: Procedures carried out by specialists 50-80% of the time

Notes: This figure concentrates on procedures carried out by specialists 50-80% of the time. The first figure looks at the percentage of physicians carrying out these procedures (even just once) who are in primary care vs some specialty, across urbanity levels. The second figure looks at the percentage of services provided by primary care vs some specialty, across urbanity levels. The urbanity of the area is an index from 1 to 6, where a higher value denotes a more rural area. Source: Author's calculations based on data from CMS.

the urbanity index, by specialization, as shown in Figure 2.[4]

The preceding data analysis documented the following facts. First, specialty procedures are also executed by PCPs, who seem to make an informed choice about whether to provide specialty procedures. Moreover, those PCPs who provide specialty procedures are likely to perform multiple specialty procedures. Finally, PCPs in less populated, less urban areas are able to perform more specialty procedures than PCPs in more

[4] All the results that are presented in the remainder of this paper are based on the procedures carried out by specialists 60-80 percent of the time. All the results are qualitatively robust to different specialization ranges, and these robustness checks are available upon request.

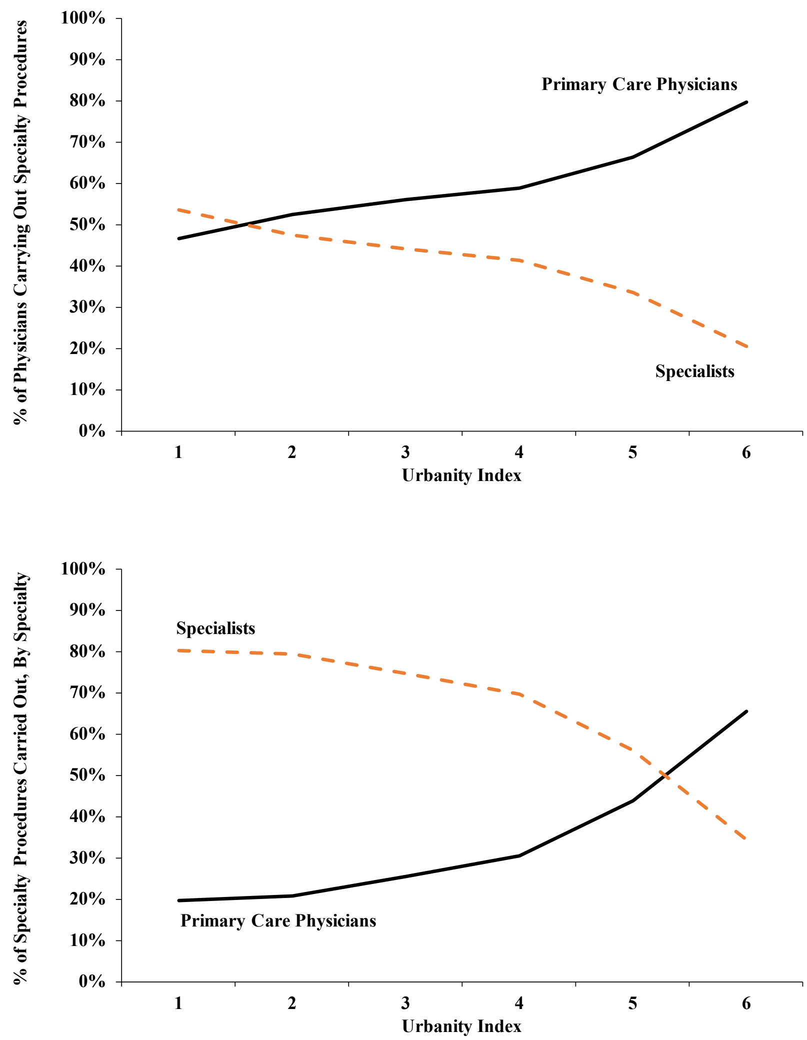

Figure 2: Procedures carried out by specialists 70-80% of the time

Notes: This figure concentrates on procedures carried out by specialists 70-80% of the time. The first figure looks at the percentage of physicians carrying out these procedures (even just once) who are in primary care vs some specialty, across urbanity levels. The second figure looks at the percentage of services provided by primary care vs some specialty, across urbanity levels. The urbanity of the area is an index from 1 to 6, where a higher value denotes a more rural area. Source: Author's calculations based on data from CMS.

urban, densely populated places. This latter finding hints at a tradeoff in their decision-making on what treatments to offer and give to patients in different locations, supporting the theory of this paper. The data therefore support the theory that PCPs take over more specialty procedures in more rural places, where specialists are not as present as they would be in cities. The remainder of this paper will focus on analyzing whether PCPs are able to use this urbanity margin to take on more remunerative procedures.

6 Empirical Evidence and Reduced-Form Results

This Section seeks to provide some empirical evidence as well as discuss some reduced-form results that inform the theory of this paper. To do so, I restrict my attention to the procedures carried out by specialists 60-80 percent of the time and by PCPs the remaining 20-40 percent of the time. As previously discussed, procedures are categorized considering the incidence with which all procedures are carried out across all places and years to be able to classify a procedure as a specialty procedure. All level analyses are performed controlling for area-specific characteristics as well as patients' demographics and health characteristics.

First, I analyze whether there is a relationship between the reimbursement fees and the frequency with which a given specialty procedure is carried out by primary care. Second, I control for fixed effects, exploiting the time variation in the data since reimbursement fees are changed on an annual basis. I then estimate through a logistic regression the effect that changes in the reimbursement fees have on the probability with which a PCP carries out a specialty procedure. I run the estimation both with and without frequency weights. With frequency weights, a higher weight is given to observations that show a higher increase in the frequency with which the more remunerative procedure is carried out.

The typical endogeneity challenge that econometricians face in similar analyses of different industries is less of a concern here because of the focus on Medicare patients. Medicare patients do not have to pay out-of-pocket for the entirety of the bill. For all the procedures considered here, Medicare offers insurance coverage, guaranteeing that patients only have to worry about a co-pay and a co-insurance rate. Therefore, it is highly plausible that patients would not shop around to find a cheaper provider for that particular procedure, but would take the price as given. This assumption would imply that the coefficient on price should be very close to zero.

Nevertheless, there is still the potential of some unobservable effect on the price. To illustrate, consider the fact that choices of physicians are often based on personal matters, such as friendliness of the physician, recommendation from friends or family, or simply him/her being the "family practice of the town." Similarly to a higher-quality product in another industry, this provider could have higher rates that could translate into a higher out-of-pocket cost for patients. This would bias the price coefficient upward, indicating a lower elasticity of patients to prices, when instead the choice of the higher-rate physician would be due to factors that are unobserved by the econometrician. To control for this issue, I use instrumental variables which represent input costs. These costs would increase a physician's pricing menu but not be observed or considered by patients, which is the identifying assumption necessary for the instrument to work. I utilize malpractice insurance reimbursement fees as a proxy for malpractice insurance. Malpractice insurance is a very good proxy for costs: it is higher in cost for higher-cost cities and is higher for higher-risk specialties, which also involve greater costs for machines and equipment. The higher the cost for insurance, the higher the total procedure billing, but malpractice insurance never enters patients' demand functions, as it is as a pure input cost and does not affect other characteristics. I also include the so-called Geographical Adjustment Factors (GAFs), which are cost-of-living adjustments for work, practice expenses, and malpractice insurance, proxying for regional variation in input costs. In the specifications that account for additional heterogeneity, I supplement these instruments with functions of the same instruments, as common in the literature: the interaction term of the work GAF with the practice expense GAF, the square of the work GAF, and the square of malpractice reimbursement units. The coefficient on price once it is instrumented is significant and

Table 2: Number of procedures carried out by PCPs

Variables	$\hat{\beta}$
Procedure Reimbursement	3.33
	(5.44)
Procedure GAF	323.68
	(81.31)
Large Metro w/in MSA	-8.67
	(1.85)
Large Metro	-10.87
	(1.71)
Metro with 250k$\leq$ $pop < 1mil$	-6.29
	(1.49)
Metro with 50k$\leq$ $pop < 250k$	-2.34
	(1.62)
Metro with 10k$\leq$ $pop < 50k$	-0.33
	(1.57)
Constant	-7.63
	(0.29)
Observations	700,855
R^2	0.31

Notes: Linear regression of variables of interest on the number of procedures carried out by PCPs. The urbanity of the area is an index from 1 to 6, where a higher value denotes a more rural area. Geographical, procedure, and year dummies are included, as well as demographic controls. Standard errors in parentheses.

negative, but still close to zero, as theory would suggest.

To provide some empirical reduced-form evidence supporting this paper, let me first report the results of a regression of the number of services provided by primary care for specialty procedures on the reimbursement fees, controlling for costs, and area-specific characteristics as well as patients' demographics and health characteristics. As previously mentioned, I restrict my attention to those procedures carried out by PCPs 20-40 percent of the time (and by specialists 60-80 percent). The results are shown in Table 2.

These results are consistent with the suggested theory. The geographical adjustment factors control for the cost of living in different areas, which explains the high value on the procedure geographical adjustment (Procedure GAF) coefficient, as an increase of 1 in the geographical practice cost indices is a very high increase. To illustrate, a unit increase in the work and practice indices is equal to double the difference between the maximum and minimum value in the same indices, while a unit increase in the malpractice insurance cost index is roughly equal to the difference between the malpractice insurance cost in Wyoming and the malpractice insurance cost in NYC suburbs. The procedure geographical adjustments are higher, where the cost of living is higher, that is, in cities, so the parameters on the geographical cost indices simply reflect the fact that more populated places lead to higher service counts. The Procedure Reimbursement variable is the reimbursement factor. I find that PCPs respond positively to increases in reimbursements, increasing the number of more remunerative specialty procedures they carry out. To be precise, I find that for every unit increase in the reimbursement revenue for a given procedure, equal to about $36, the physician performs that procedure three more times.

Another key component of the analysis is the margin of urbanity. The urbanity index value that is left out for comparison is the value attached to the most rural areas. Comparing the parameters, it is easy to see that the more rural the area, the higher the service count of procedures carried out by PCPs, as expected.

Next, I create a variable which takes value of 1 whenever a PCP increases the number of times a given specialty procedure is provided from a year to the next, 0 otherwise. I then run a logistic regression to see whether an increase in the reimbursement rate of a given procedure raises the probability of a PCP increasing the number of times she carries out that procedure. The results in Table 3 show that an increase in the reimbursement rate of one value unit for a given procedure, equal to roughly a $36 increase, increases the probability of PCPs increasing the number of times that procedure is carried out by about 16.6 percent. Frequency weights play an important role in this estimation. Accounting for frequency weights, so that a higher weight is placed on physicians who increase the number of procedures carried out more, an increase of one relative value unit in a given procedure, equal again to about $36, leads to a 45.5 percent higher probability of the PCP performing that procedure more.

For robustness, I widen the set to the procedures carried out by PCPs between 20 and 50 percent of the time (and by specialists 50-80 percent of the time). The results are presented in the Appendix Section A.4 and show that the qualitative results are the same.

Table 3: Logit: y=1 if PCPs increase the number of specialty procedures carried out - 20-40%

Variables	(1)	(2)
Change in Procedure Reimbursement	0.17	0.46
	(0.02)	(0.003)
Change in Malpractice Reimbursement	-1.69	-3.39
	(0.06)	(0.01)
Change in Practice Expense Reimbursement	0.25	0.63
	(0.01)	(0.001)
Change in Procedure GAF	9.25	37.02
	(1.78)	(0.19)
Change in Malpractice GAF	-0.04	0.29
	(0.02)	(0.003)
Change in Practice Expense GAF	-3.20	-12.57
	(0.71)	(0.08)
Constant	1.94	2.75
	(0.004)	(0.004)
Frequency Weights	NO	YES
Observations	1,677,036	1,677,036

Notes: Logistic regression of variables of interest on whether or not primary care increases the specialty procedures provided, with and without frequency weights. The procedures included are those that are carried out by PCPs, on average, between 20 and 40% of times as suggested by the specialization index previously discussed. Year dummies are included. Standard errors in parentheses.

7 Structural Model Results

I run four specifications to estimate the model discussed in Section 4: the random coefficients logit model, with and without Chamberlain's optimal instruments, a specification that uses the demographic variation in the data to estimate the random coefficient, and finally a random coefficient logit model using both Chamberlain's optimal instruments and the demographic variation in the data. Table 4 reports all results on the analysis of PCPs increasing their share at the expense of specialists along the urbanity index. All specifications utilize the time variation in the data to identify the parameters. Recall that the relative value units are set by policymakers annually.

A few results should be discussed. First, the price coefficient is very close to zero, but negative, as expected in the previous discussion in Section 6. Second, the urbanity of the area emerges again as a key margin for PCPs. The five variables which look at the interaction of the reimbursement with the urbanity index generate estimates that have to be compared with the reference variable, which value indicates the most rural areas. The results support the empirical evidence shown in Section 6 that PCPs are able to increasingly take on a higher share of specialty procedures, the more rural the area.

Since all the specifications are variation of the logit model, the Reader should recall that the estimated

Table 4: PCPs Turn Specialists

MEAN UTILITY	(1)	(2)	(3)	(4)
Price	-0.001	-0.001	-0.001	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Procedure Reimbursement	-0.067	-0.080	-0.119	-0.046
	(0.20)	(0.01)	(0.05)	(0.01)
Procedure Reimbursement, Large Central Metro	-0.116	-0.115	-0.111	-0.106
	(0.02)	(0.01)	(0.01)	(0.01)
Procedure Reimbursement, Large Fringe Metro	-0.026	-0.030	-0.023	-0.039
	(0.01)	(0.01)	(0.01)	(0.01)
Procedure Reimbursement, Medium Metro	-0.040	-0.040	-0.030	-0.031
	(0.02)	(0.01)	(0.01)	(0.01)
Procedure Reimbursement, Small Metro	0.031	0.029	0.039	0.033
	(0.01)	(0.01)	(0.01)	(0.01)
Procedure Reimbursement, Micropolitan Area	-0.048	-0.048	-0.031	-0.040
	(0.02)	(0.01)	(0.01)	(0.01)
Medical Doctor Degree	-0.432	-0.463	-0.431	-0.538
	(0.03)	(0.02)	(0.02)	(0.02)
Number of PCPs Around	-0.002	-0.002	-0.002	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Number of PCPs Doing Specialty Work	0.575	0.656	0.575	0.915
	(0.21)	(0.07)	(0.07)	(0.07)
Constant	-6.405	-6.447	-6.392	-6.607
	(0.18)	(0.03)	(0.03)	(0.03)
RANDOM COEFFICIENT, Procedure Reimbursement				
SD	0.120	0.120	0.155	0.080
	(0.11)	(0.00)	(0.01)	(0.00)
Specification				
Demographics	NO	NO	YES	YES
Optimal Instruments	NO	YES	NO	YES
Observations	76,133	76,133	76,133	76,133

Notes: The table reports the results from the structural estimations described in Section 7. Standard errors in parentheses. Estimates of control variables are omitted from the table to facilitate the Reader and are available upon request.

Table 5: Change in Physicians' Weight in Consumers' Utility with Respect to Physicians in Large Metro Areas, γ

γ	Base Case	Optimal Instruments	Demographics	Optimal Instruments and Demographics
γ_2-Fringe Metro	9.42%	8.87%	8.80%	6.92%
γ_3-Medium Metro	7.90%	7.79%	8.44%	7.79%
γ_4-Small Metro	15.84%	15.49%	16.18%	14.91%
γ_5-Micropolitan	7.04%	6.93%	8.33%	6.82%
γ_6-Rural	12.30%	12.19%	11.74%	11.18%

Notes: This table shows the increase in the primary care physician's utility weight on the consumer utility, γ , along the urbanity index compared with the most urban classification (urban=1, large metro area), for all specifications considered in this paper. The four specifications are the following: a random coefficients model that uses the BLP (1994) algorithm, the second one uses the same model but includes Chamberlain optimal instruments, the third one uses the same model but lets the random coefficients depend on the empirical distribution of demographic characteristics, and the final one includes both the optimal instruments and the demographics in the estimation.

coefficients on the interaction of the reimbursement with the urbanity index, $\hat{\beta}_{rvu}$, are not exactly equal to the parameters of interest, which are instead equal to: $\gamma_{urban} = \exp\left\{\hat{\beta}_{rvu,urban}\right\} - 1 \;\; \forall$ urban $= 1, ..., 6$. To aid the Reader with the interpretation of the coefficients, a summary of the coefficients of interest γ_{urban} is shown in Table 5, switching the point of view to set the most urban counterparts as the point of reference. Bearing this in mind, I find that an increase in one relative value unit, once again equal to about \$36, PCPs are able to increase their share in specialty procedures between 7 and 15 percent in more rural areas compared with their most urban counterparts, at the expense of specialists. These results provide some evidence that the urbanity is a key margin along which PCPs can use the lower competition coming from specialists to increase their reimbursements by carrying out the specialty procedures themselves rather than referring them to specialists somewhere else. The most affected areas are very rural areas (the dummy of reference, which value is higher than almost every other more urban value) and in small metropolitan areas (which is the only estimate higher than the dummy of reference). Primary care physicians' ability to increase their shares at the expense of specialists along the urbanity index is very robust and significant across all specifications.

8 Conclusion and Future Work

Using data on the insured population of Medicare, this paper has provided evidence for a supply-induced mechanism in the demand for healthcare in the US. First, it has documented the novel fact that PCPs are able to carry out an increasing share of specialty procedure in more rural areas of the United States, where specialists are fewer. Second, this paper has found that PCPs increase their share in specialty procedures more, the more remunerative they become, and are able to do so more, the more rural the area they practice in. From the reduced-form results, this paper has found that an increase in one unit of reimbursement for a given procedure, equal to an increase in reimbursement of about \$36, leads to over a 45 percent higher probability of PCPs providing a higher number of the more remunerative procedure. Structurally, this paper has estimated that the same $36 reimbursement increase for a given procedure leads to an increase in the PCP share in that specialty procedure by 7-15 percent more in less urban areas compared with their most urban counterparts, at the expense of specialists. Small metropolitan areas (with a population between 50,000 and 250,000 people) and very rural areas (with a population smaller than 10,000 people) are the most affected.

This paper has also been the beginning of a more comprehensive research project analyzing physicians' choices bearing in mind how 1) specialty differences are crucial when modeling their behavior, 2) reimbursements are a substantial income component that has been largely ignored in many analyses, and 3) the urbanity of the area comes into play in many different ways, as physicians practice differently according to where they are. Falcettoni (2018) keeps these factors in mind when analyzing the determinants of physicians' location choice, with a focus on the rural shortage of physicians. The question that naturally arises from the results of this paper is whether PCPs should in fact be able to carry out specialty procedures. The rural population in the US, as discussed in Falcettoni (2018), is deeply in need of health care personnel. Since the rural population needs specialty care as well, whether it is specialists or PCPs who carry out the specialty procedures should not matter if the two types of workforce are clear substitutes for those procedures. Nevertheless, the rural population is also sicker, and whether this is related to a lack of access to care or to the quality of the health care personnel in rural areas is an open question on its own, which Falcettoni (2019) tries to address.

Despite the many projects that have sprung from this analysis, many questions arise that still have to be addressed. First, a natural extension to this paper is to analyze the supply side of this issue. In particular, PCPs need to make investment choices to be able to offer specialized treatments. Second, related to Falcettoni (2018), it would be interesting to study how these financial incentives influence medical students' choices. It is plausible that medical students have adjusted their specialty choices since the implementation of the fee-for-service system based on the possibility to get different financial benefits from different locations. This question is important to explain the decrease in students' interest in primary care specialties following the 1992 change in Medicare pricing. This hypothesis is able to rationalize the wage and offering differential across specialties, while also showing that the ability for primary care to perform more specialized procedures in less urban areas is the result of the pricing mechanism described in this paper. Therefore, this theory would be able to explain not only the general decline in the interest for primary care specialties, but also the geographical shifts in the primary care distribution in the country.

Finally, as suggested in the introduction, this paper touched upon the relationship between a fee-for-service payment scheme and overutilization. While the issue of overutilization is often discussed, the focus on a particular set of procedures and/or specialties overlooks the reallocation of these procedures between the specialty groups as procedures become more remunerative. If this effect is not included, the entire take-up rate could be incorrectly attributed to overutilization, overstating the effect of the latter. Therefore, this paper provided evidence that it is key to consider two complementary effects caused by fee-for-service payment schemes: 1) a reallocation of the procedures carried out from specialists to PCPs for any given total number of procedures, as shown here, and 2) the increase in the number of total procedures carried out overall without an apparent increase in patients' health issues, that is, overutilization, which is not analyzed in this paper.

References

1. Acemoglu, Daron, Amy Finkelstein, and Matthew J. Notowidigdo. "Income and Health Spending: Evidence from Oil Price Shocks." *Review of Economics and Statistics* 95.4 (2013): 1079-1095.

2. American Hospital Association. "Physician Ownership and Self-Referral in Hospitals: Research on Negative Effects Grows." *Trendwatch*. (April 2008).

3. Anderson, Gerard F., et al. "Health Spending in the United States and the Rest of the Industrialized World." *Health Affairs* 24.4 (2005): 903-914.

4. Armour, Brian S., et al. "The Effect of Explicit Financial Incentives on Physician Behavior." *Archives of Internal Medicine* 161.10 (2001): 1261-1266.

5. Baker, Laurence C. "The Effect of HMOs on Fee-For-Service Health Care Expenditures: Evidence from Medicare." *Journal of Health Economics* 16.4 (1997): 453-481.

6. Barro, Jason, and Nancy Beaulieu. "Selection and Improvement: Physician Responses to Financial Incentives." *National Bureau of Economic Research Working Paper* 10017 (2003).

7. Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile Prices in Market Equilibrium." *Econometrica: Journal of the Econometric Society* (1995): 841-890.

8. Blomqvist, Ake, and Colin Busby. *How to Pay Family Physicians: Why 'Pay Per Patient' Is Better than Fee for Service*. C.D. Howe Institute Commentary 365 (2012).

9. Broomberg, J., and M. R. Price. "The Impact of the Fee-For-Service Reimbursement System on the Utilisation of Health Services. Part I. A Review of the Determinants of Doctors' Practice Patterns." *South African Medical Journal* 78.3 (1990): 130-132.

10. Chaix-Couturier, Carine, et al. "Effects of Financial Incentives on Medical Practice: Results from a Systematic Review of the Literature and Methodological Issues." *International Journal for Quality in Health Care* 12.2 (2000): 133-142.

11. Clemens, Jeffrey, and Joshua D. Gottlieb: "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?." *The American Economic Review* 104.4 (2014): 1320-1349.

12. —. "In the Shadow of a Giant: Medicare's Influence on Private Physician Payments." *Journal of Political Economy* 125.1 (2017): 1-39.

13. Croxson, Bronwyn, Carol Propper, and Andy Perkins. "Do Physicians Respond to Financial Incentives? UK Family Physicians and the GP Fundholder Scheme." *Journal of Public Economics* 79.2 (2001): 375-398.

14. Emanuel, Ezekiel J., and Victor R. Fuchs. "The Perfect Storm of Overutilization." *Journal of the American Medical Association* 299.23 (2008): 2789-2791.

15. Falcettoni, Elena. "Physician Workforce Effect on Health." *Working Paper* (2019).

16. —. "The Determinants of Physicians' Location Choice: Understanding the Rural Shortage." *Working Paper* (2018).

17. Finkelstein, Amy. "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare." *The Quarterly Journal of Economics* 122.1 (2007): 1-37.

18. Fisher, Elliott S., et al. "The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care." *Annals of Internal Medicine* 138.4 (2003): 273-287.

19. Gawande, Atul, and BBC Audiobooks America. *Better: A Surgeon's Notes on Performance*. Picador (Jan. 22, 2008).

20. Ginsburg, Paul B. "Fee-For-Service Will Remain a Feature of Major Payment Reforms, Requiring More Changes in Medicare Physician Payment." *Health Affairs* 31.9 (2012): 1977-1983.

21. Gosden, Toby, et al. *Capitation, Salary, Fee-For-Service and Mixed Systems of Payment: Effects on the Behaviour of PCPs.* The Cochrane Library (2000).

22. Grant, Darren. "Physician Financial Incentives and Cesarean Delivery: New Conclusions from the Healthcare Cost and Utilization Project." *Journal of Health Economics* 28.1 (2009): 244-250.

23. Gruber, Jonathan, and Maria Owings. "Physician Financial Incentives and Cesarean Section Delivery." *National Bureau of Economic Research Working Paper* 4933 (1994).

24. Grumbach, Kevin, et al. "Primary Care Physicians' Experience of Financial Incentives in Managed-Care Systems." *New England Journal of Medicine* 339.21 (1998): 1516-1521.

25. Hall, Robert E., and Charles I. Jones. "The Value of Life and the Rise in Health Spending." *The Quarterly Journal of Economics* 122.1 (2007): 39-72.

26. Hellinger, Fred J. "The Impact of Financial Incentives on Physician Behavior in Managed Care Plans: A Review of the Evidence." *Medical Care Research and Review* 53.3 (1996): 294-314.

27. Hemenway, David, et al. "Physicians' Responses to Financial Incentives: Evidence from a For-Profit Ambulatory Care Center." *New England Journal of Medicine* 322.15 (1990): 1059-1063.

28. Hendee, William R., et al. "Addressing Overutilization in Medical Imaging." *Radiology* 257.1 (2010): 240-245.

29. Hickson, Gerald B., William A. Altemeier, and James M. Perrin. "Physician Reimbursement by Salary or Fee-For-Service: Effect on Physician Practice Behavior in a Randomized Prospective Study." *Pediatrics* 80.3 (1987): 344-350.

30. Hillman, Alan L. "Financial Incentives for Physicians in HMOs." *New England Journal of Medicine* 317.27 (1987): 1743-1748.

31. Hillman, Alan L., Mark V. Pauly, and Joseph J. Kerstein. "How Do Financial Incentives Affect Physicians' Clinical Decisions and the Financial Performance of Health Maintenance Organizations?." *New England Journal of Medicine* 321.2 (1989): 86-92.

32. Jacobson, Mireille, et al. "Does Reimbursement Influence Chemotherapy Treatment for Cancer Patients?." *Health Affairs* 25.2 (2006): 437-443.

33. Levin, David C., and Vijay M. Rao. "Turf Wars in Radiology: The Overutilization of Imaging Resulting from Self-Referral." *Journal of the American College of Radiology* 5.7 (2008): 806-810.

34. McElduff, P., et al. "Will Changes in Primary Care Improve Health Outcomes? Modelling the Impact of Financial Incentives Introduced to Improve Quality of Care in the UK." *Quality and Safety in Health Care* 13.3 (2004): 191-197.

35. Melichar, Lori. "The Effect of Reimbursement on Medical Decision Making: Do Physicians Alter Treatment in Response to a Managed Care Incentive?." *Journal of Health Economics* 28.4 (2009): 902-907.

36. Nevo, Aviv. "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand." *Journal of Economics & Management Strategy* 9.4 (2000): 513-548.

37. Pham, Hoangmai H., and Paul B. Ginsburg. "Unhealthy Trends: The Future of Physician Services." *Health Affairs* 26.6 (2007): 1586-1598.

38. Rasmusen, Eric. *The BLP Method of Demand Curve Estimation in Industrial Organization.* Department of Business Economics and Public Policy, Kelley School of Business, Indiana University. (2007).

39. Rosenblatt, Roger A., and L. Gary Hart. "Physicians and Rural America." *Western Journal of Medicine* 173.5 (2000): 348.

40. Safran, Dana Gelb, Alvin R. Tarlov, and William H. Rogers. "Primary Care Performance in Fee-For-Service and Prepaid Health Care Systems: Results from the Medical Outcomes Study." *Journal of the American Medical Association* 271.20 (1994): 1579-1586.

41. Scott, Anthony, et al. "The Effect of Financial Incentives on the Quality of Health Care Provided by PCPs." *Cochrane Database of Systematic Reviews* 2011.9 (2010).

42. Shrank, William, et al. "Effect of Physician Reimbursement Methodology on the Rate and Cost of Cataract Surgery." *Archives of Ophthalmology* 123.12 (2005): 1733-1738.

43. Stearns, Sally C., Barbara L. Wolfe, and David A. Kindig. "Physician Responses to Fee-For-Service and Capitation Payment." *Inquiry* (1992): 416-425.

44. United States Census Bureau. Census 2010 PMSA Definition:

https://www.census.gov/population/estimates/metro-city/99mfips.txt

45. Vincent, David W. "The Berry–Levinsohn–Pakes Estimator of the Random-Coefficients Logit Demand Model." *Stata Journal* 15.3 (2015): 854-880.

Appendix

A.1 Institutional Framework

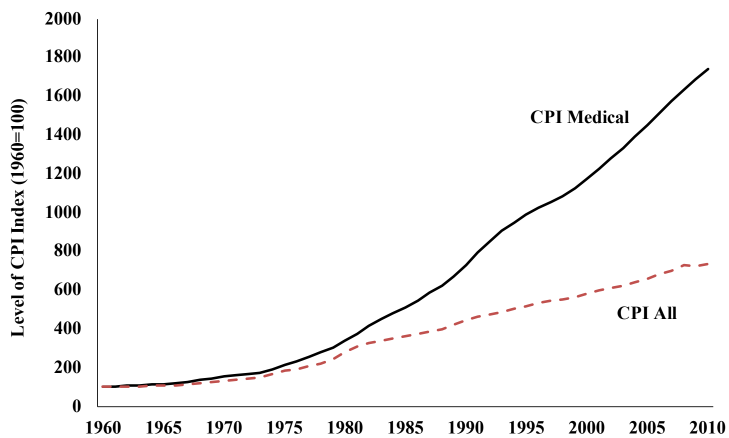

Figure A.1: Medical CPI vs. All CPI

Notes: This figure shows the Consumer Pricing Index for all goods vs. only medical goods. The growth of CPI for medical goods changed dramatically both following the birth of Medicare in 1967 and the birth of the Resource-Based Relative Value Unit System in 1992. Source: CMS.

As shown in Figure A.1, the increase in health expenditure has been long-lasting, with a clear increase in its growth since the 80s. The first full year under Medicare coverage (1967) clearly led to an increase in medical costs.

The 1992 fee-for-service system is called the Resource-Based Relative Value Scale (RBRVS). The system was based on some initial rates and geographical adjustment factors, which would be reviewed on an annual basis by the RVS Update Committee (RUC). The RUC was meant to only have an advisory role, but its recommendations are accepted 97 percent of the time, making it *de facto* the fee-setting organization.

The Reader should bear in mind that the fee-for-service system is not new to 1992. The system before, the Usual, Customary, and Reasonable (UCR) system, was still based on a fee-for-service reimbursement; however, these reimbursements were not standardized across physicians, and tractability was not possible due to lack of information on individual pricing. This pricing system matched with lack of transparency is what prompted discussions at the beginning of 1990 to reform it. This paper shows that the new pricing system exacerbated the issue, leading to a change in physicians' decision making. Moreover, it allows the researcher to be able to estimate the impact of this pricing on physicians' choices due to the fee standardization (and its availability publicly).

For each procedure j in market t, the reimbursement is equal to:

$$Reimbursement_{jt} = Constant_t * RVU_{jt} * GAF_t \tag{7}$$

The constant only depends on the year and is equal across specialties and procedures. The relative value units change according to the procedure as well as the year, and the geographic adjustment factors (GAFs) depend on the market (area and year).

The constant, called the Conversion Factor (CF), is a national adjustment factor, which is identical across specialties, areas, and procedures. The 2017 CF is equal to $35.8887. The GAFs are a proxy for cost of living, adjusting for differences in input costs across payment regions.

The RUC's recommendations across the years have been constantly widening the gap between the procedure reimbursements usually carried out by specialists and those regularly carried out by primary care.

Since the reimbursement does not depend on who carries out the procedure, but only on the procedure itself, and specialty procedures are more highly priced than typical primary care procedures, this payment system generates financial incentives for PCPs to substitute primary care procedures for more specialized, remunerative procedures when possible.

A.2 Standard Multinomial Logit Demand

This section presents what the model would look like if it was simplified to a standard multinomial logit model of demand.

I assume the household to face the following utility function when buying product j in market t:

$$\begin{aligned} u_{ijt} &= x_{jt}\beta + \lambda RVU_{jt} - \alpha p_{jt} + \epsilon_{ijt} \qquad \epsilon_{ijt} \text{ iid} \sim \text{T1EV} \\ \delta_{jt} &= x_{jt}\beta + \lambda RVU_{jt} - \alpha p_{jt} \end{aligned}$$

Markets, products, and characteristics are the same as in Section 5.

The simple logit demand assumes all individuals to be identical. Then, the market share of product j is equal to the probability that product j yields the highest utility, which happens if the ϵ disturbance of product j is high enough with respect to the alternatives:

$$s_{jt} = \Pr\left(x_{jt}\beta + \lambda RVU_{jt} - \alpha p_{jt} + \epsilon_{ijt}\right) > \Pr\left(x_{kt}\beta + \lambda RVU_{kt} - \alpha p_{kt} + \epsilon_{ikt}\right) \quad \forall k \neq j \tag{8}$$

Thanks to the properties of the type-1 extreme value distribution, this reduces to the share of product j being equal to:

$$s_{jt} = \frac{\exp\{x_{jt}\beta + \lambda RVU_{jt} - \alpha p_{jt}\}}{1 + \sum_{k=1}^{1175} \exp\{x_{kt}\beta + \lambda RVU_{kt} - \alpha p_{kt}\}} \tag{9}$$

Table A.1: OLS Regression of Price on Chosen Instruments

Variables	y = Price
Procedure GAF	840.80
	(52.15)
Practice Expense GAF	-82.06
	(23.40)
Malpractice GAF	13.18
	(5.59)
Malpractice Reimbursement	748.50
	(1.83)
Constant	-677.70
	(49.00)
Observations	142,112
R^2	0.54

Notes: This table shows the coefficients obtained by regressing price on the chosen instrumental variables. Standard errors in parentheses.

To estimate this demand, notice that:

$$
\begin{aligned}
\log\left(\frac{s_{jt}}{s_{0t}}\right) &= \\
\log(s_{jt}) - \log(s_{0t}) &= \log\left(\frac{\exp\{x_{jt}\beta + \lambda RVU_{jt} - \alpha p_{jt}\}}{1+\sum_{k=1}^{1175}\exp\{x_{kt}\beta + \lambda RVU_{kt} - \alpha p_{kt}\}}\right) - \log\left(\frac{\exp\{0\}}{1+\sum_{k=1}^{1175}\exp\{x_{kt}\beta + \lambda RVU_{kt} - \alpha p_{kt}\}}\right) \\
\log(s_{jt}) - \log(s_{0t}) &= \log\left(\exp\{x_{jt}\beta + \lambda RVU_{jt} - \alpha p_{jt}\}\right) - \log\left(1+\sum_{k=1}^{1175}\exp\{x_{kt}\beta + \lambda RVU_{kt} - \alpha p_{kt}\}\right) + \\
&\quad - \log\left(\exp\{0\}\right) + \log\left(1+\sum_{k=1}^{1175}\exp\{x_{kt}\beta + \lambda RVU_{kt} - \alpha p_{kt}\}\right) \\
\log(s_{jt}) - \log(s_{0t}) &= \log\left(\exp\{x_{jt}\beta + \lambda RVU_{jt} - \alpha p_{jt}\}\right) - \log(1) \\
\log(s_{jt}) - \log(s_{0t}) &= x_{jt}\beta + \lambda RVU_{jt} - \alpha p_{jt}
\end{aligned}
$$

The estimation would then be equivalent to an IV regression with the same instruments described in the body of the text.

A.3 Model-Implied Marginal Cost Regression on Input Costs

Table A.1 shows the results from a regression of the instruments discussed in Section 6 on price. The instruments explain about 54 percent of the price variance with highly significant coefficients.

The marginal costs can be backed out from the estimates of the model. Table A.2 shows the results from a regression of the marginal costs implied by the model on the instrumental variables chosen for a random market, i.e., rural Kansas. Marginal costs are well fitted by the instrumental variables representing input

Table A.2: Regression of Marginal Costs on Cost Shifters, Rural Kansas

Variables	(1)	(2)
Procedure GAF	774.20	766.19
	(1.68)	(1.66)
Practice Expense GAF	-72.10	-71.67
	(1.74)	(1.73)
Malpractice GAF	0.07	0.07
	(1.40)	(1.39)
Procedure GAF^2	-417.51	-413.38
	(1.69)	(1.66)
Malpractice Reimbursement	0.42	0.42
	(13.71)	(13.67)
Malpractice Reimbursement^2	-0.04	-0.04
	(6.81)	(6.80)
Procedure GAF*Practice Expense GAF	71.53	71.12
	(1.74)	(1.73)
Random Error		-0.004
		(0.44)
Constant	-353.80	-349.93
	(1.65)	(1.63)
Observations	272	272
R^2	0.74	0.75

Notes: Exemplary results for one market. Marginal costs estimated are regressed on the instrumental variables used in the paper, showing a good fit of cost variance. T-Statistics in parentheses.

costs, with an R^2 equal to 0.74.

A.4 Robustness of Reimbursement Effect on Probability

Table A.3 shows the results from the logistic regression as explained in Section 6 on a widened range of procedures, carried out by specialists 50 to 80 percent of the time, and by PCPs the remainder of the time. The positive effect of the reimbursement rate is consistent. The increase in the probability of providing specialty procedures, independently of the number of services provided, is about the same as before. Once I include frequency weights, the probability is about half. This is due to the fact that I am now including many procedures that are generally provided by PCPs to begin with, so that these procedures are more homogeneously carried out across the physician population.

Table A.3: Logit: y=1 if PCPs increase the number of specialty procedures carried out - 20-50%

Variables	(1)	(2)
Change in Procedure Reimbursement	0.14	0.21
	(0.02)	(0.003)
Change in Malpractice Reimbursement	-0.62	-2.00
	(0.03)	(0.01)
Change in Practice Expense Reimbursement	0.02	0.50
	(0.01)	(0.001)
Change in Procedure GAF	10.30	32.83
	(1.40)	(0.15)
Change in Malpractice GAF	-0.003	0.17
	(0.02)	(0.002)
Change in Practice Expense GAF	-2.64	-7.34
	(0.56)	(0.06)
Constant	1.77	2.58
	(0.003)	(0.003)
Frequency Weights	NO	YES
Observations	2,304,888	2,304,888

Notes: Logistic regression of variables of interest on whether or not primary care increases the specialty procedures provided, with and without frequency weights. The procedures included are those that are carried out by PCPs, on average, between 20 and 50% of times as suggested by the specialization index previously discussed. Year dummies are included. Standard errors in parentheses.