CCHE CCES
Canadian Centre for Health Economics
Centre canadien en économie de la santé

Working Paper Series
Document de travail de la série

PAYING FOR PRIMARY CARE: THE FACTORS ASSOCIATED WITH PHYSICIAN SELF-SELECTION INTO PAYMENT MODELS

David Rudoler, Raisa Deber, Janet Barnsley, Richard H. Glazier, and Audrey Laporte

Working Paper No: 2014-06

www.canadiancentreforhealtheconomics.ca

March 19, 2014

Canadian Centre for Health Economics
Centre canadien en économie de la santé
155 College Street
Toronto, Ontario

CCHE/CCES Working Paper No. 2014-06
March 19, 2014

Paying For Primary Care: The Factors Associated With Physician Self-Selection Into Payment Models

Abstract

In this paper we use a panel of administrative data to determine the factors associated with primary care physician self-selection into different payment models in Ontario, Canada. We find that primary care physicians will self-select into payment models based on existing practice and individual characteristics. These patterns of self-selection largely follow a utility maximizing model of physician behaviour; physicians with more complex patient populations are less likely to switch into capitation-based payment models where higher levels of effort are not financially rewarded. These findings have implications for future work that considers the impact of payment incentives on provider behaviour, and for governments introducing multiple payment models in a single health-care sector.

JEL Classification: D22; D21; I11

Key words: physician behaviour, financial incentives, administrative data, panel data

Corresponding Authors:

David Rudoler[1]
Doctoral candidate
155 College Street,
Toronto, ON Canada
david.rudoler@utoronto.ca

[1]This study was supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred. This study was funded by CIHR Grants and Studentship Grants (CIHR Grant numbers MOP-114928 and DHE-121276), as well as doctoral award (264515).

Introduction

In this paper we attempt to determine the factors that lead physicians to self-select into different types of remuneration models. The results of this analysis have both practical policy and empirical implications. With respect to the former, the results of this study may enhance our understanding of the impacts of payment reform for healthcare providers, and the implications payment reforms have for physician self-selection when governments introduce multiple payment models into a single market. With respect to the latter, understanding the factors that lead physicians to select into different payment models will also help us better understand how they will react to these payment incentives once they are under their influence. This has implications for the estimation of the impact of payment incentives more generally, as we can more accurately account for self-selection in these estimation procedures.

This research builds off of existing work that has considered the supply decisions of physicians (Brown & Lapan, 1979; Ellis & McGuire, 1986; Thornton, 2013; Thornton & Eakin, 1997). These studies have measured the quantity of services physician choose to provide (Thornton, 2013), or how they may locate their practice in one area as opposed to others (Newhouse, Williams, Bennett, & Schwartz, 1982); and, to some extent, how physicians decide which types of patients they will roster to their practice (Kralj & Kantarevic, 2013; Sibley & Glazier, 2012). Of course, all of these decisions are in some way affected by the payment physicians receive; we know that payment is, at least in part, related to the quantity of services provided, the place in which a physician chooses to locate their practice, and the types of patients they choose to care for. So, what happens when physicians have multiple payment options to select from in a single market? How does this choice bias our observations about supply decisions?

In this study we take a first step toward answering this question by first determining the factors that lead primary care physicians (PCPs) to self-select into different payment models in Ontario, Canada. Payment reform for PCPs in Ontario presents an interesting natural experiment for studying the relationship between payment and physician behaviour. The province of Ontario has undertaken a number of reforms to the way PCPs are compensated, adding mixed payment models to traditional fee-for-service (FFS), which has been the traditional form of remuneration for physicians in Canada.

Policy Context

Canadian healthcare is made up of 13 provincial and territorial healthcare systems, with each province acting as a single payer for publicly insured healthcare services. Physician services are publicly funded, but privately delivered; and traditionally, all physicians in Canada were paid FFS. Many PCPs in Ontario remain in FFS, where fees are set through negotiations between the Ontario Ministry of Health and Long-Term Care (MOHLTC) and the Ontario Medical Association (OMA). The fees and associated service codes are listed in the Ontario Physician Schedule of Benefits, which PCPs use to bill for each service they provide. PCPs practicing under FFS are not required through contract with the province, nor do they have financial incentives to practice in groups or hire multidisciplinary providers.

In the early 2000s, the Ontario government began to introduce a number of alternatives to FFS in an attempt to encourage uptake of primary care by medical students, control costs, and provide incentives to improve access to care. As we will demonstrate in our analysis below, most PCPs in Ontario currently receive an alternative payment to pure FFS, generally enhanced fee-for-service (EFF) or capitation (CAP). Ontario has two EFF models that were implemented in 2003/04: Family Health Groups and the Comprehensive Care Model. The former requires PCPs practice in groups of three or more, while the latter is a solo-physician practice model. PCPs under EFF receive a majority of their revenue through standard FFS payments; however, they also receive a 10% premium for 13 fee codes provided to patients they formally roster to their practice and a small capitation payment. Formal rostering requires patients sign an enrolment form that designates a PCP as their primary source of care. In all cases, rostering is entirely voluntary, though PCPs will receive different payments and financial incentives for rostered patients than they will for non-rostered patients.

Over the past decade, capitation (CAP) models have received considerable uptake from PCPs and patients (Glazier, Zagorski, & Rayner, 2012). In 2001, the Ontario government introduced the Family Health Network model, and in 2005 the Family Health Organization model. PCPs operating under one of these CAP models receive a majority of their payment through an age and sex adjusted payment for each patient they formally roster regardless of service quantity. In 2004, the Ontario government also introduced the Family Health Teams, which are not a payment model, but a practice model that includes interdisciplinary providers (e.g., nurses, dietitians, pharmacists, etc.). The vast majority (∼92%) of PCPs in Family Health Teams also receive capitation payment (Glazier, Zagorski, & Rayner, 2012).

There has been some recent empirical work on the impact of primary care payment reform in and outside of the Canadian context. For instance, Glazier et al. (2009) compared patient and practice characteristics across payment models in Ontario. The authors used cross-sectional data to find that PCPs in CAP models serve patient populations with higher income, and lower morbidity and co-morbidity levels. However, the chosen study design did not allow the authors to control for physician self-selection, making it impossible to attribute these differences to payment incentives. It is possible that differences across models are fully attributable to the self-selection of particular physicians into these models.

Devlin and Sarma (2008) and Sarma et al. (2010) conducted studies in the Canadian context using data from a cross-sectional national survey of physicians to determine the impact of remuneration on physician output. They utilized an instrumental variable approach to control for physician self-selection into different payment models. The authors found PCPs with characteristics associated with productivity and PCPs who expressed desire to engage in non-clinical activities were more likely to choose alternatives to FFS.

Kantarevic and Kralj (2011) considered changes in physician productivity as they switched from traditional FFS to EFF in Ontario, Canada. Within the study, the authors setup a model for the physician's decision to switch to EFF and used a two-part empirical model to control for physician self-selection. However, modeling self-selection was not their primary objective in this paper and, therefore, it is no surprising that they did not report the factors that contributed to this selection process. In a subsequent paper, Kralj and Kantarevic (2013) analyzed differences in quality and quantity outcomes in EFF versus capitation-based practices in Ontario, Canada. Age, sex, expected earnings, and quantities of services delivered were significantly associated with joining a capitation-based model in Ontario.

We expect to build on this existing literature by considering a broader range of factors that may have contributed to PCP self-selection behaviour. We also track this behaviour over a longer period of time, ranging from when all PCPs in Ontario were in FFS practices, to the present day where the vast majority of PCPs are in an alternative payment model.

Theoretical Framework & Empirical Model

We follow the basic theoretical tenets established in the existing literature on physician behaviour (Ellis, 1998; Evans, 1974; Gaynor & Pauly, 1990; Thornton, 2013; Thornton & Eakin, 1997) and largely builds off similar work conducted in the Ontario primary care context (Kantarevic et al., 2011). We assume that PCPs are price taking utility maximizers, whose utility is a function of consumption, leisure and patient benefit.

$$max\ \ U[C, L, H_1, H_2] \tag{1}$$

where C denotes consumption, L denotes leisure, and H_1 and H_2 denote patient benefit in the FFS and capitation contexts respectively. As specified elsewhere in the literature, we define leisure as the PCP's time spent outside of providing healthcare services (Thornton, 2013). It is also assumed that PCPs cannot exit the production process entirely; the physician cannot simply hire other health professionals to provide care and devote all time to leisure. Policies around the delegation of medical services in Ontario prohibit such behaviour. Patient benefit is a function of the quantity of services (Q) provided per patient, where $\bar{Q} = Q_1 + Q_2$, and $H_1 = VQ_1$ and $H_2 = RQ_2$. In order to account for patient case-mix we assume that each physician's pool of patients (V or R) contains both simple and complex patients. Thus V = $V_{Simple} + V_{Complex}$ and R = $R_{Simple} + R_{Complex}$ (Ellis 2008). We assume that the marginal productivity of each individual service is lower for complex patients than for simple patients, which suggests greater quantities of service will be required to generate patient benefit (H) for complex patients. In addition, we assume the quantity of services delivered to each patient (Q) is a function of the physician's own time (M), hired medical labour (N) and capital (K) inputs.

$$Q_i = Q(M, N, K) \tag{2}$$

We assume that these inputs are conditioned on whether the patient is simple or complex, as we assume complex patients require more resource intensive service.

The physician's budget constraint varies under different payment models. Under pure FFS, the PCP received a fixed fee for each individual service provided. As Gaynor and Pauly (1990) suggest, this is a function of the physician's time, hired labour (e.g., nursing time), capital, physician effort, and individual and practice characteristics that affect productivity. Under mixed capitation payment, however, the majority of the physician's revenue is a function of the number of patients enrolled to the practice, not to service quantity. Thus, efficiency and effort are not financially rewarded in capitation models in the same way it is in FFS. The PCP's budget constraint is as follows:

$$P_1VQ_1 + P_2R - w(N) - r(K) = C \tag{3}$$

where P_1Q_1 denotes income received in a FFS context (i.e., fee times the quantity of services provided), and P_2R denotes income received in a capitation context (i.e., price times the number of rostered patients). w denotes the wage paid to hired medical staff (N), and r denotes rents paid for capital (K). The PCP also faced a time constraint:

$$T = L + M_1V + M_2R \tag{4}$$

where M_1V represents the time the physician spends providing care to non-rostered patients and M_2R represents time spent providing care to rostered patients. We also assume the following applies in the FFS context:

$$Q > 0 \tag{5}$$
$$V > 0 \tag{6}$$
$$R \geq R_{min} = 0 \tag{7}$$

In both the FFS and the capitation context, the physician optimizes utility by choosing the number of patients she wishes to care for and the amount of inputs to put into the production process. However, in this context we are interested in the selection of patients. The first order conditions with respect to V and R are as follows:

$$U_{H1}Q_1 + U_C P_1 Q_1 - U_L M_1 = 0 \tag{8}$$
$$U_{H2}Q_2 + U_C P_2 - U_L M_2 + \mu = 0 \tag{9}$$

Changes in the Capitation Context

In the capitation context a number of changes are important for the purposes of this analysis. First, it is important to note that the capitation payment schemes in Ontario are mixed schemes, meaning that physicians receive a majority of their payment from age and sex adjusted capitation payments, and additional FFS payments for 'out-of-basket' services and incentivized services (e.g., diabetes management). Therefore, in the capitation context there is a change in FFS payments, which we represent as $\Delta P_1 = P_1 b < 0$, where b is the proportion of services that are incentivized or 'out-of-basket'. This suggests that average FFS prices will be lower in the capitation context than in the FFS context. Second, capitation income will increase (ΔP_2 >0). And third, the constraint on roster size will come into effect (ΔR_{min} >0). Using the Envelope theorem and evaluating at the optimum value of the endogenous variables we can produce the following value equation:

$$\Delta v = VQ_1 * \Delta P_1 + R * \Delta P_2 \geq \frac{\mu}{U_C} * \Delta R_{min} \tag{10}$$

The left side of equation (10) is the income the physician could expect to earn if they switched to capitation. The right side represents the disutility of switching. Using equations (8) and (9), we can rewrite the right hand side of equation (10) as follows:

$$\frac{\mu}{U_C} * \Delta R_{min} = [\frac{M_1}{M_2} * P_1 Q_1 - P_2] * \Delta R_{min} \tag{11}$$

Equation (11) has been simplified by removing the U_H terms since we assume that physicians will produce equal levels of patient benefit under FFS and capitation, at least in the short-term. The right side of equation (11) suggests that physicians will be less likely to move into capitation models if P_1 is high, and if they are efficient at producing medical services. For instance, Barro and Beaulieu (2003) found that the introduction of performance pay for physicians in a Florida-based managed care company caused the least productive physicians to leave the company, and attracted new physicians who were on average more productive (measured in financial performance). Furthermore, since only FFS compensates physicians for contributing their own time to produce services, physicians with more complex patient pools will less likely to move to capitation if the capitation fee (P_2) is not adequately risk adjusted for case-mix.

Our empirical model attempts to capture the factors that influence individual PCP decisions to switch remuneration schemes, and we must do this without being able to directly observe PCP preferences. Our model variables are listed and defined in Table 1 and described further in next section of this paper.

Data and Variable Specification

We use administrative data obtained from the Institute for Clinical Evaluative Sciences (ICES) in Ontario, which has a comprehensive research agreement with the Ontario Ministry of Health and

Long-Term Care. We analyzed a panel of 12-year cross-sections (1999/00 - 2010/11) of all practicing PCPs indexed at March 31 of each year. In each year, we excluded PCPs who were not practicing "comprehensive primary care". While this applies to PCPs in all payment models, the majority of pure FFS physicians listed in our databases are either not practicing primary care regularly, or have practices that focus on other specialties (e.g., sports medicine, tropical medicine, etc.) and, are therefore, not comparable with physicians in other payment models. We excluded these physicians by using an algorithm developed at ICES, which flags physicians who worked less than 50 days annually, whose billed services comprised less than 50% 'core' primary care services (i.e., services that fit within 22 activity areas considered part of 'core' primary care), and provided services within seven or more of these activity areas. We also excluded PCPs who worked in practices that focus on emergency medicine or mental health, all physicians who did not bill at least 8 of the 18 'standard' primary care fee codes two year prior to each panel cross-section, and those physicians who were not listed as general practitioners or family physicians in our databases.

In order to obtain variables for physician practice characteristics we assigned individual patients to physicians using formal and virtual patient rostering. Formally rostered patients included all who signed an enrolment form with a PCP in an EFF or CAP model and remained rostered to that PCP on March 31 of each year. Virtual rostering was determined using an algorithm that counts the number of 'primary care' visits a patient had in the previous two years, links those services to a PCP, and assigns the patient to the PCP with the highest billing costs. It is important to note that pure FFS patients do not sign an enrolment form, and therefore, by definition are all virtually rostered.

To measure patient complexity we used Adjusted Clinical Groups (ACGs), which is a patient case-mix system developed at Johns Hopkins University. The ACG case-mix system groups patients by diagnosis into clinically consistent groups. We use the ACG system to assign patients to 32 Aggregated Diagnostic Groups (ADGs) based on clinical and expected utilization criteria; and, we use the ACG system to categorize patients using Resource Utilization Bands (RUBs), which assigns each ACG category to six levels of morbidity burden and health system utilization ranging from 0 (non-users) to 5 (very-high users). We provide the specifications for all our variables in Table 1.

Analytical Approach

To estimate our equation we used a random effects probit model (using the *xtprobit* command in Stata12) on unbalanced (all PCPs in each year) and balanced (only PCPs that had an observation in each year) panels of PCP practices. A random effects model was preferred since we wanted to observe the coefficients of time invariant effects, which are removed in fixed effects models. However, to allow for our explanatory variables to be correlated with individual effects, we also use the Mundlak specification. To apply this specification we included the within-mean values for explanatory variables. The coefficients on the within-mean variables are generally not interpretable, except that they act as a Wald test (F-test) with the null hypothesis that the coefficients on the within-means equal zero; this tests the existence of unobserved heterogeneity in the individual effect (Wooldridge, 2002). In our results section we compare the coefficients of the balanced and unbalanced panel, and the random effects with the Mundlak specification. Since the random effects probit model is calculated using quadrature, which relies on the number of integration points, we also used the *quadchk* command in Stata12 to test the consistency of our estimates across different integration point values (i.e., 8, 12, and 16). Since our results were robust across all integration values, we used the default number (12) integration points for the purposes of computational efficiency.

Results

Table 2 provides descriptive statistics. Here we compare the results for our unbalanced (n=9,068) and balanced (n=3,775) panels. There is general consistency across these groups, which suggests they should be largely comparable, including within the distribution of PCPs in FFS, EFF and CAP models. We provide more detail with respect to this distribution and its change over time in Figure 1. We can see that in 1999/00 nearly all Ontario PCPs were receiving FFS. There was a sharp decline in the number of pure FFS PCPs starting in 2003/04, and a corresponding increase in the number of PCPs in EFF models. This coincided with the implementation of the Family Health Group and Comprehensive Care Models. The number of PCPs in CAP models increased slowly but steadily until 2004/05 with the introduction of the Family Health Teams and the Family Health Organization model. While, changes have not been as sharp as they were in the FFS-based models, the CAP model is now the most prominent one for Ontario PCPs. Further descriptive statistics on the switching patterns of PCPs over the study period are provided in the Appendix to this paper (see Figures A1-A3). In addition, Table A1 provides descriptive statistics for PCPs who were excluded based on our inclusion/exclusion criteria.

Table 3 provides the results of our estimation of the probability of being in an EFF model, and Table 4 provides the results of our estimation of the probability of being in a CAP model. We report the results of both the random effects probit and with the Mundlak specification as average marginal effects (AMEs). However, we focus on the results of the Mundlak specification, as this is our preferred model. For both the probability of being in EFF and CAP, the lagged dependent variable had a strong and significant effect. With respect to the probability of being in EFF, there was a strong negative effect of being in CAP in the previous year, suggesting PCPs did not switch out of CAP models once they were in them. With respect to the probability of being in CAP, the effect of being in EFF in the previous year was positive, which suggests PCPs switched from EFF to CAP.

In terms of PCP and practice characteristics, PCP group size did not have significant effect on model selection. However, there was a weak but significant relationship between patient roster size and selection; PCPs with larger rosters were less likely to be in EFF and more likely to be in a CAP payment model. Years since graduation had a positive marginal effect on the probability of switching to EFF and CAP. The sex of the PCP did not have a consistent effect across estimations; male PCPs were less likely to be in EFF in the unbalanced panel, and more likely to be in CAP in the balanced panel. Finally, the rurality of the PCP's practice did not have a consistent effect on the probability of being in either model; although, in the unbalanced panel, PCPs in very rural practices were less likely to be in EFF.

With respect to the age distribution of PCPs' roster, having a higher proportion of older patients had a negative effect on the likelihood of being in EFF models, and an inconsistent, but positive effect on being in CAP. It is unclear why this was the case, but may be due to the age-sex risk adjustment in CAP practices, where PCPs receive a higher income for older patients.

However, our results suggest that this age-sex risk adjustment may not have adequately accounted for patient complexity. Patient income distribution had a negative effect on switching to EFF, and a positive effect on switching to CAP in the random effects model, but not in the Mundlak specification. Previous cross-sectional studies have demonstrated that CAP practices enroll wealthier patients (Glazier et al., 2009), however, our results suggest that this relationship does not hold in a panel context. The proportion of patients who were immigrants decreased the likelihood of switching in both models, but the effect is only consistent for switching to EFF. With respect to the morbidity distributions of patient rosters, for the most part, physicians with higher proportions of 'sicker' patients were more likely to switch to EFF, and less likely to switch to CAP models. This relationship holds in the case of utilization (measured in RUBs), and in the number of morbidities (measured in number of ADGs).

Table 5 provides the negative predictive and positive predictive values for all specifications of our models. We had a very high negative predictive value and a fairly high positive predictive value of between 75 and 80%. We also conducted a RESET test for misspecification and show the results in Table 6. We found some misspecification in some of our estimation, which may be related to omitted variables that we discuss in our conclusions. We also conducted sensitivity analyses by fluctuating the exclusion criteria for the minimum number of rostered patients by 50% in both directions (not reported here), which did not alter our findings, but increasing the threshold did improve our RESET test results. In addition, we also produced a correlation matrix (using corr command) for our independent variables to check for collinearity. Only PCP age (Age) and years since graduation (ysg) had a correlation coefficient greater than 0.7; however, removing PCP Age from the estimation did not alter p-values.

We also estimated pooled pair-wise logit models on the balanced and unbalanced panel to determine if our results were robust when considering payment model as a multinomial outcome. Table 6 provides the results for the EFF compared to FFS, and CAP compared to FFS. Table 7 provides the results for CAP compared to EFF. All results are reported in log odds and all estimations utilized the Mundlak specification. Our results generally hold up under this approach. We show that there is evidence that PCPs are less likely to switch to CAP if they have sicker patients, but that this relationship does not exist as strongly for switching to EFF. This is particularly true when we compare the probability of being in CAP with the probability of being in EFF (excluding FFS PCPs).

Finally, for sensitivity analyses we estimated linear probability models using both random and fixed effects specifications (see Tables A2 and A3 in Appendix), and we calculated the Hausman test to determine if our random effects model is biased by unobserved heterogeneity. We found that the fixed effects approach was preferred in this context. We also found that these results largely mirrored our panel probit estimation. The results of these analyses are reported in the Appendix to this paper.

Conclusions and Policy Discussion

We analyze factors associated with PCP self-selection into different payment models in Ontario, Canada using administrative panel data. Our analysis was conducting using random effects with and without the Mundlak within-means specification to control for unobserved heterogeneity in the individual effect. We also conducted this analysis in an unbalanced and balanced panel. Several coefficient values remained robust across these different specifications, giving us confidence in their interpretation.

However, there are some important limitations. First, we have no variable for expected earnings, nor do we have a variable for service quantity. These variables would help us better proxy expected consumption under different payment models (Kantarevic et al., 2011; Kralj & Kantarevic, 2013); we hope to correct this in future analyses. Second, we do not observe PCP preferences directly. We must proxy these preferences with other variables that may lead to some precision issues. Finally, there may be some limitations with how we have excluded PCPs we do not consider to be delivering 'comprehensive' primary care. We felt these exclusions were necessary to ensure we captured PCPs who would have been eligible to switch payment models; however, our criteria for exclusion may have led to some selection bias issues. Although, as we mentioned earlier, sensitivity analyses on the exclusion of PCPs with fewer than 650 patients did not alter our findings.

Generally, we found that our hypotheses largely hold under empirical testing. Our most important and convincing findings relate to our effort variables; that is, factors that affect the amount of effort required of PCPs to provide care to their patient population. We suspected that PCPs with more effort inducing patient rosters would be less inclined to switch to CAP models if the capitation payment was not adequately risk adjusted. This is because effort is a factor in the physician's production function that is financially rewarded in FFS, but not in capitation payment models (note that Ontario's CAP models are mixed remuneration schemes, so effects may be less pronounced). Our results corroborate this prediction, as we found that PCPs with less affluent and sicker patient rosters were less likely to switch to CAP, whereas we did not see as strong a relationship for EFF models.

This finding has important empirical and policy implications. Previous work by Glazier and colleagues (2009; 2012) demonstrated that CAP models tended to have healthier and higher income patients, but due to the cross-sectional nature of those studies it was unclear whether this was due to physician self-selection or a result of the changes in practice behaviour as a result of financial incentives inherent in the payment models. Our findings suggest that a large component of the differences in patient complexity across models are a result of physician self-selection. We can use this insight in future studies to determine isolate the impact of selection and the ongoing influence of financial incentives. With respect to the policy implications of these findings, we suggest that the implementation of several voluntary physician payment reforms in a single market will encourage the sorting of physicians into models that theory would predict. Without appropriate checks on this behaviour, physicians will sort themselves into the models that will allow them to maximize their individual preferences, which may not necessarily be beneficial for the healthcare system more generally. Further work is needed to determine the full implications of this sorting behaviour.

References

Barro, J. R., Beaulieu, N. D., National Bureau of Economic Research Working Paper 10017. (2003). Selection and Improvement: Physician Response to Financial Incentives

Brown, D. M., & Lapan, H. E. (1979). The suppy of physician services. *Economic Inquiry*, 17(2), 269-279

Devlin, R. A., & Sarma, S. (2008). Do physician remuneration schemes matter? The case of Canadian family physicians. *Journal of health economics*, 27(5), 1168-1181.

Ellis, R. P. (2008). Risk adjustment in health care markets: concepts and applications. *Financing Health Care*, 177-222.

Ellis, R. P. (1998). Creaming, skimping and dumping: provider competition on the intensive and extensive margins. *Journal of health economics*, 17(5), 537-555.

Ellis, R. P., & McGuire, T. G. (1986). Provider behavior under prospective reimbursement. Cost sharing and supply. *Journal of health economics*, 5(2), 129-151.

Evans, R. G. (1974). Supplier-induced demand: some empirical evidence and implications. In M. Perlman (Ed.), (Vol. 6, pp. 162-173). London, Macmillan.

Gaynor, M., & Pauly, M. V. (1990). Compensation and Productive Efficiency in Partnerships: Evidence from Medical Groups Practice. *Journal of Political Economy*, 98(3), 544-573.

Glazier, R. H., Klein-Geltink, J., Kopp, A., & Sibley, L. M. (2009). Capitation and enhanced fee-for-service models for primary care reform: a population-based evaluation. *Canadian Medical Association Journal*, 180(11), E72-E81.

Glazier, R. H., Zagorski, B. M., & Rayner, J. (Eds.). (2012). *Comparison of Primary Care Models* *in Ontario by Demographics, Case Mix and Emergency Department Use*, 2008/09 to 2009/10 (pp. 1-41). Institute for Clinical Evaluative Sciences.

Kantarevic, J., Kralj, B., & Weinkauf, D. (2011). Enhanced fee-for-service model and physician productivity: Evidence from Family Health Groups in Ontario. *Journal of health economics*, 30(1), 99-111.

Kralj, B., & Kantarevic, J. (2013). Quality and quantity in primary care mixed-payment models: evidence from family health organizations in Ontario. *Canadian Journal of Economics/Revue canadienne d'économique*, 46(1), 208-238.

Newhouse, J. P., Williams, A. P., Bennett, B. W., & Schwartz, W. B. (1982). *How Have Location Patterns of Physicians Affected the Availability of Medical Services?* Santa Monica, California: Rand Corporation, 1982.

Sarma, S., Devlin, R. A., & Hogg, W. (2010). Physician's production of primary care in Ontario, Canada. *Health Economics*, 19(1), 14-30.

Sibley, L. M., & Glazier, R. H. (2012). Evaluation of the equity of age-sex adjusted primary care capitation payments in Ontario, Canada. *Health Policy*, 104(2), 186-192.

Thornton, J. (2013). The labour supply behaviour of self-employed solo practice physicians. *Applied Economics*, 30(1), 85-94.

Thornton, J., & Eakin, B. K. (1997). The Utility-Maximizing Self-Employed Physician. *The Journal of Human Resources*, 32(1), 98-128.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Tables And Figures

Table 1: Variable Definitions

Variable	Definition
Dependent Variable	
EFF	PCP in EEF = 1, Otherwise = 0
CAP	PCP in CAP = 1, Otherwise = 0
PCP characteristics	
Ysg	Years since graduation
Age	Age of the PCP
Age^2	Polynomial PCP age
CMG	Canadian graduate = 1
Sex	PCP male = 1
Practice Characteristics	
Grpsz	Number of PCPs with a shared group number.
Rurality	Ontario Medical Association Rurality Index of Ontario (RIO): Urban = RIO < 40, Rural = RIO 40 - 74, and Remote = RIO 75+
Npats	Number of formally and virtually rostered patients per PCP
Effort Variables	
Pat. Age distribution	The proportion of each PCP's patient roster that is < 19, 19-44, 45-64, 65-79 and 80+
Prop. Male	The proportion of each PCP's patient roster that is male
Prop. Imm.	The proportion of each PCP's patient roster that is an immigrant
Avg. income quint.	Average income quintile of patient population
Pat. RUB distribution	The proportion of each PCP's patient roster that has an RUB score of 0, 1, 2, 3, 4, & 5
Prop. 6-9 ADGs	The proportion of each PCP's patient roster that has 6-9 ADGs
Prop. 10+ ADGs	The proportion of each PCP's patient roster that has 10 ADGs
Dynamic Variables	
EFF_{t-1}	EFF = 1 in previous year
CAP_{t-1}	CAP = 1 in previous year

Table 2: Descriptive Statistics

Variables	Unbalanced Panel (N= 74,453; n = 9,068; T-bar = 8.21)		Balanced Panel (N = 45,300; n = 3,775; T-bar = 12)	
	Mean	Std. Dev.	Mean	Std. Dev.
FFS	0.47	0.5	0.46	0.5
EFF	0.34	0.47	0.36	0.48
CAP	0.19	0.39	0.18	0.38
Group size	16.51	50.83	16.21	50.31
Npats	1,551.73	645.54	1,706.19	633.55
CMG	0.74	0.44	0.79	0.41
Ysg	24.05	10.51	25.24	9.05
Age	49.71	10.25	50.77	8.9
Male	0.67	0.47	0.7	0.46
Rurality				
RIO < 40	0.75	0.43	0.75	0.43
RIO 40 - 74	0.19	0.4	0.2	0.4
RIO 75+	0.06	0.23	0.05	0.22
Pat. Age Distribution				
Prop. < 19	21.43	8.46	21.5	8.08
Prop. 19 - 44	36.81	9.57	35.68	7.87
Prop. 45 - 64	27.06	6.82	27.69	6.31
Prop. 65 - 79	10.91	5.89	11.2	5.51
Prop. 80+	3.79	3.17	3.94	3.09
Prop. Male	0.47	0.12	0.47	0.11
Avg. Income quint	3.07	0.54	3.08	0.54
Prop. Imm.	0.11	0.13	0.1	0.12
Pat. RUB distribution				
Prop. RUB 0	1.94	2.52	1.97	2.36
Prop. RUB 1	6.99	3.45	6.72	2.94
Prop. RUB 2	21.46	5.09	21.14	4.6
Prop. RUB 3	52.48	5.95	53.01	5.44
Prop. RUB 4	13.28	3.64	13.26	3.43
Prop. RUB 5	3.86	2.44	3.9	2.36
Prop. 0 ADGs	1.93	2.51	1.97	2.36
Prop. 1-5 ADGs	47.73	8.96	47.06	8.05
Prop. 6-9 ADGs	41.44	6.71	41.98	6.03
Prop. 10+ ADGs	8.9	4.08	8.99	3.93

Note: “N” = total number of observations, “n” = unique observations, “T-bar” = average years in panel

Table 3: Average Marginal Effects (Pr. EFF = 1)

Variables	Random Effects Probit Unbalanced	Random Effects Probit Balanced	Random Effects Probit w/ Mundlak Specification Unbalanced	Random Effects Probit w/ Mundlak Specification Balanced
EFF_{t-1}	0.3927***	0.3829***	0.2872***	0.2554***
CAP_{t-1}	-0.2592***	-0.2977***	-0.3306***	-0.3732***
Group size	0.0003***	0.0003***	0.0000***	0.0001***
Npats*10	0.0002***	0.0002***	-0.0002***	0.0003***
CMG	-0.0091**	0.0061	-0.0133***	0.0001
Ysg	0	0.0010*	0.0272***	0.0470***
Age	-0.0008	0.0054**	-0.0002	0.0027
Age^2	0	-0.0000**	0	-0.0002***
Male	-0.0264***	-0.0336***	-0.0088**	-0.01
Rurality				
RIO < 40 (dropped)	-	-	-	-
RIO 40 - 74	-0.0231***	-0.0285***	-0.003	0.0095
RIO 75+	-0.0518***	-0.0606***	-0.0967***	-0.0808
Pat. Age Distribution				
Prop. <19 (dropped)	-	-	-	-
Prop. 19 - 44	-0.0016***	-0.0026***	-0.0027***	-0.0101***
Prop. 45 - 64	0.0012***	0.0010**	-0.0027***	-0.0070***
Prop. 65 - 79	-0.0055***	-0.0059***	-0.0078***	-0.0135***
Prop. 80+	-0.0002	-0.0017*	-0.0089***	-0.0202***
Prop. Male	0.0134	0.0402	-0.2995***	-0.5714***
Avg. Income quint	-0.0036	-0.0199***	-0.0332***	-0.0858***
Prop. Imm.	0.0067	-0.1038***	-0.2945***	-0.4324***
Pat. RUB distribution				
RUB 0-1 (dropped)	-	-	-	-
Prop. RUB 2	0.0062***	0.0063***	0.0134***	0.0185***
Prop. RUB 3	0.0040***	0.0034***	0.0072***	0.0056***
Prop. RUB 4	-0.0007	-0.0042**	0.0053***	0.0037**
Prop. RUB 5	0.0161***	0.0164***	0.0244***	0.0351***
Prop. 0 - 5 ADGs (dropped)				
Prop. 6-9 ADGs	0.0011**	0.0022***	0.0028***	0.0084***
Prop. 10+ ADGs	0.0047***	0.0070***	0.0115***	0.0185***
Pseudo R^2	0.4448	0.4741	0.4751	0.5066

Note: * $p < .10$, ** $p < .05$, *** $p < .001$

Table 4: Average Marginal Effects (Pr. CAP = 1)

Variables	Random Effects Probit		Random Effects Probit w/ Mundlak Specification	
	Unbalanced	Balanced	Unbalanced	Balanced
CAP_{t-1}	0.2810***	0.2995***	0.1901***	0.1824***
EFF_{t-1}	0.0547***	0.0557***	0.0267***	0.0264***
Group size	0.0000*	0	0	0
Npats*10	0.0000***	0.0000**	0.0004***	0.0006***
CMG	-0.003	0.004	-0.0082**	-0.0026
Ysg	-0.0017***	-0.0005	0.0138***	0.0131***
Age	-0.0039***	0.0012	-0.0050***	0.0007
Age^2	0.0000***	0	0	0
Male	-0.0129***	0.0017	-0.0046	0.0159**
Rurality				
RIO <40 (dropped)	-	-	-	-
RIO 40 - 74	0.0076**	0.0023	0.0015*	-0.0034
RIO 75+	0.0131***	0.0171***	0.017	0.0463
Pat. Age Distribution				
Prop. <19 (dropped)	-	-	-	-
Prop. 19 - 44	-0.0007***	-0.0004*	-0.0004	0.0005
Prop. 45 - 64	0.0019***	-0.0001	0.0015***	0.0003
Prop. 65 - 79	-0.0005	0.0004	-0.0010*	0.0029**
Prop. 80+	0.0012**	0.00010*	0.0011	0.0034**
Prop. Male	-0.0700**	-0.0653***	-0.0211	0.1592**
Avg. Income quint	0.0102***	0.0101***	0.0006	0.0001
Prop. Imm.	-0.0963***	-0.1342***	-0.0086	0.0245
Pat. RUB distribution				
RUB 0-1 (dropped)	-	-	-	-
Prop. RUB 2	-0.0156***	-0.0154***	-0.0149***	-0.0154***
Prop. RUB 3	-0.0086***	-0.0083***	-0.0093***	-0.0124***
Prop. RUB 4	-0.0036***	-0.0059***	-0.0045***	-0.099***
Prop. RUB 5	-0.0056***	-0.0085***	-0.0077***	-0.0138***
Prop. 0 - 5 ADGs (dropped)				
Prop. 6-9 ADGs	-0.0040***	-0.0032***	-0.0024***	0.0007
Prop. 10+ ADGs	-0.0077***	-0.0055***	-0.0057***	-0.0033**
Pseudo R^2	0.5626	0.6396	0.5961	0.6538

Note: * p < .10, ** p < .05, *** p < .001

Table 5: Negative & Positive Predictive Values

Model		Random Effects Probit		Random Effects Probit w/ Mundlak Specification	
		Unbalanced	Balanced	Unbalanced	Balanced
EFF					
	NPV	94.73%	94.21%	93.96%	93.07%
	PPV	79.90%	80.02%	79.41%	80.46%
CAP					
	NPV	99.42%	99.86%	98.08%	98.02%
	PPV	77.33%	76.27%	79.45%	79.38%

Figure 1: Enrolment in Payment Models (1999/00 - 2010/11)

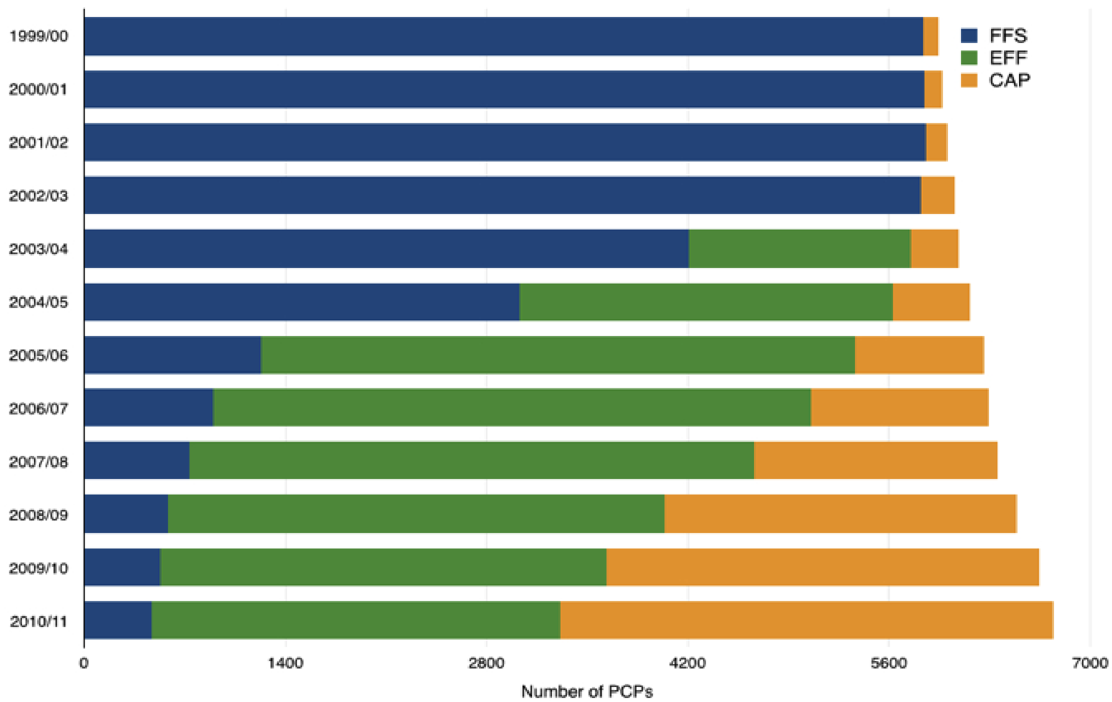


Table 6: Pooled Logit with Mundlak Specification (Base Outcome = FFS)

	EFF = 1 (Exclude: CAP)		CAP = 1 (Exclude: EFF)	
Variables	Unbalanced	Balanced	Unbalanced	Balanced
EFF_{t-1}	6.2330***	6.4200***	6.6700***	8.8572***
CAP_{t-1}	1.5441***	1.8293***	8.9872***	10.6054***
Group size	0.0013**	0.0009	0	0.0018
Npats	0.0003***	0.0021***	0.0015***	0.0027***
CMG	-0.1419**	-0.0107	-0.3597**	-0.0918
Ysg	-0.9175***	1.2300***	0.4849***	0.2809
Age	-0.0896***	0.0166	-0.1674***	0.0727
Age^2	0.0041***	-0.0069***	-0.0009	-0.0012
Male	-0.0941	0.1422	-0.3733**	0.0344
Rurality				
RIO <40 (dropped)	-	-	-	-
RIO 40 - 74	0.2545	0.4582	0.4746	-0.3663
RIO 75+	-0.7925**	0.613	-1.0856	-0.1422
Pat. Age Distribution				
Prop. <19 (dropped)	-	-	-	-
Prop. 19 - 44	-0.0483***	-0.1710***	-0.0818***	-0.1913***
Prop. 45 - 64	-0.0339***	-0.1086***	0.0008	-0.1619**
Prop. 65 - 79	-0.1089***	-0.1754***	-0.1812***	-0.1643**
Prop. 80+	-0.0873***	-0.2801***	-0.0291	-0.1851**
Prop. Male	-4.3773***	-1.6282	-1.7708	8.2489
Avg. Income quint	-0.2416*	-1.1451***	0.138	0.9112*
Prop. Imm.	-2.5212***	-3.0624**	0.2861	2.9921
Pat. RUB distribution				
RUB 0-1 (dropped)	-	-	-	-
Prop. RUB 2	-0.0906***	-0.0708**	-0.5193***	-0.7892***
Prop. RUB 3	-0.0526***	-0.1242***	-0.2267***	-0.4274***
Prop. RUB 4	0.0670***	-0.1726***	-0.0939**	-0.4110***
Prop. RUB 5	0.1371***	0.2129***	-0.1209**	-0.2942**
Prop. 0 - 5 ADGs (dropped)				
Prop. 6-9 ADGs	-0.0280**	0.0642***	-0.0949***	0.04
Prop. 10+ ADGs	0.1616***	0.3183***	-0.0174	0.3073***
Pseudo R^2	0.6561	0.7211	0.9042	0.9379

Note: * $p < .10$, ** $p < .05$, *** $p < .001$

Table 7: Pooled Logit with Mundlak Specification (Base Outcome = EFF)

	CAP = 1 (Exclude: FFS)	
Variables	Unbalanced	Balanced
EFF_{t-1}	-0.3881***	-0.033
CAP_{t-1}	6.5798***	7.2252***
Group size	-0.0003	-0.0009*
Npats	0.0009***	0.0014***
CMG	-0.0663	-0.0047
Ysg	-0.1114*	-0.2083**
Age	-0.0554**	0.0098
Age^2	0.0017**	0.0020**
Male	0.1335	0.5259***
Rurality		
RIO <40 (dropped)	-	-
RIO 40 - 74	0.1876	0.3349
RIO 75+	0.7993	1.9556*
Pat. Age Distribution		
Prop. <19 (dropped)	-	-
Prop. 19 - 44	0.001	0.04032**
Prop. 45 - 64	0.0281**	0.0077
Prop. 65 - 79	0.0400**	0.085**
Prop. 80+	0.0615**	0.1335**
Prop. Male	-0.1542	7.5641**
Avg. Income quint	0.0438	0.3713
Prop. Imm.	-0.8239	-1.6864
Pat. RUB distribution		
RUB 0-1 (dropped)	-	-
Prop. RUB 2	-0.3344***	-0.3066***
Prop. RUB 3	-0.2009***	-0.2142***
Prop. RUB 4	-0.1155***	-0.1842***
Prop. RUB 5	-0.2725***	-0.3410***
Prop. 0 - 5 ADGs (dropped)		
Prop. 6-9 ADGs	-0.0571***	0.0167
Prop. 10+ ADGs	-0.1494***	-0.1360***
Pseudo R^2	0.6509	0.6546

Note: * $p < .10$, ** $p < .05$, *** $p < .001$

Appendix A: Additional Tables & Figures

Table A1 provides descriptive statistics for physicians who were excluded from the sample for having a patient roster of 650 patients or less, or for not being a comprehensive primary care physician.

Table A1: Descriptive Statistics for Excluded PCPs

Variables	Excluded Observations (N=31,757; n=7,014; T-bar = 4.53)	
	Mean	Std. Dev.
FFS	0.99	0.09
EFF	<0.01	0.06
CAP	<0.01	0.06
Group size	16.12	50.6
Npats	102.51	148.13
CMG	0.68	0.47
Ysg	23.4	14.01
Age	49.32	13.71
Male	0.64	0.48
Rurality		
RIO < 40	0.8	0.4
RIO 40 - 74	0.15	0.36
RIO 75+	0.05	0.23
Pat. Age Distribution		
Prop. <19	32.57	36.17
Prop. 19 - 44	35.32	28.95
Prop. 45 - 64	19.47	20.95
Prop. 65 - 79	7.43	13.71
Prop. 80+	5.21	15.04
Prop. Male	0.54	0.25
Avg. Income quint	3.01	0.81
Prop. Imm.	0.13	0.19
Pat. RUB distribution		
RUB 0	0.73	4.27
RUB 1	8.65	13.19
RUB 2	25.76	21.37
RUB 3	43.01	23.12
RUB 4	13.97	17.61
RUB 5	7.87	16.83
Prop. 0 ADGs	0.73	4.26
Prop. < 6 ADGs	57.09	27.87
Prop. 6-9 ADGs	32.63	22.5
Prop. 10+ ADGs	9.54	17.03

Note: "N" = total number of observations, "n" = unique observations, "T-bar" = average years in panel

Table A2 provides the results for the linear probability model where EFF = 1. We used random and fixed and effects models. The Hausman test suggests that the fixed effects model is preferred.

Table A2: Linear Probability Model (Pr. EFF = 1)

	Random Effects		Fixed Effects	
Variables	Unbalanced	Balanced	Unbalanced	Balanced
Pat. Age Distribution				
Prop. <19 (dropped)	-	-	-	-
Prop. 19 - 44	-0.0018***	-0.0027***	-0.0075***	-0.0125***
Prop. 45 - 64	0.0012***	0.0011**	-0.0059***	-0.0095***
Prop. 65 - 79	-0.0058***	-0.0062***	-0.0132***	-0.0174***
Prop. 80+	0.0002	-0.0018**	-0.0152***	-0.0228***
Prop. Male	0.0001	0.0313	-0.3164**	-0.3073**
Avg. Income quint	-0.003	-0.0196***	-0.0392***	-0.0624***
Prop. Imm.	0.0096	-0.1074***	-0.2844***	-0.2990***
Pat. RUB distribution				
RUB 3 (dropped)	-	-	-	-
Prop. RUB 2	0.0043***	0.0045**	0.0204***	0.0250***
Prop. RUB 3	0.0030***	0.0024***	0.0163***	0.0197***
Prop. RUB 4	-0.0021**	-0.0055***	0.0115***	0.0142***
Prop. RUB 5	0.0156***	0.0162***	0.0376***	0.0468***
Prop. 0-5 ADGs (dropped)				
Prop. 6-9 ADGs	0.0007*	0.0019***	-0.0012	0.0012
Prop. 10+ ADGs	0.0052***	0.0075***	0.0153***	0.0180***
R^2	0.5951	0.5975	0.2009	0.2964
Hausman	0	0	0	0

Note: * p < .10, ** p < .05, *** p < .001

Table A3 provides the results for the linear probability model where CAP = 1. We used random and fixed and effects models. The Hausman test suggests that the fixed effects model is preferred.

Table A3: Linear Probability Model (Pr. CAP = 1)

	Random Effects		Fixed Effects	
Variables	Unbalanced	Balanced	Unbalanced	Balanced
Pat. Age Distribution				
Prop. <19 (dropped)	-	-	-	-
Prop. 19 - 44	-0.0014***	-0.0011***	0.0001	0.0004
Prop. 45 - 64	0.0024***	-0.0003	0.0026***	0.0010*
Prop. 65 - 79	-0.0007**	0	0.0007	0.0027**
Prop. 80+	0.0013**	0.0010*	0.0047***	0.0074***
Prop. Male	-0.0738***	-0.0672***	0.0079**	0.2106***
Avg. Income quint	0.0096***	0.0012***	0.0106*	-0.001
Prop. Imm.	-0.0277**	-0.0373***	0.2558***	0.2985***
Pat. RUB distribution				
RUB 3 (dropped)	-	-	-	-
Prop. RUB 2	-0.0267***	-0.0193***	-0.0280***	-0.0279***
Prop. RUB 3	-0.0150***	-0.0105***	-0.0203***	-0.0201***
Prop. RUB 4	-0.0096***	-0.0082***	-0.0155***	-0.0174***
Prop. RUB 5	-0.0136***	-0.0118***	-0.0214***	-0.0241***
Prop. 0-5 ADGs (dropped)				
Prop. 6-9 ADGs	-0.0057***	-0.0038***	-0.0034***	-0.0023***
Prop. 10+ ADGs	-0.0080***	-0.0051***	-0.0047***	-0.0043***
R^2	0.7293	0.7422	0.5492	0.6249
Hausman	0	0	0	0

Note: * p < .10, ** p < .05, *** p < .001

Fig. A1 -- Number of Times Switched Payment

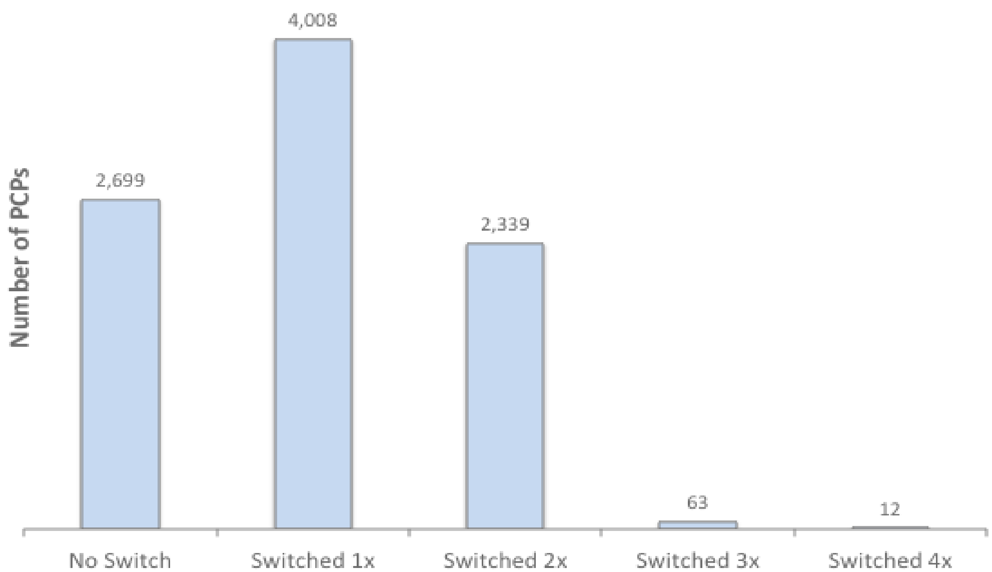


Fig. A2 -- No. of Switches (2000/01 - 2010/11)

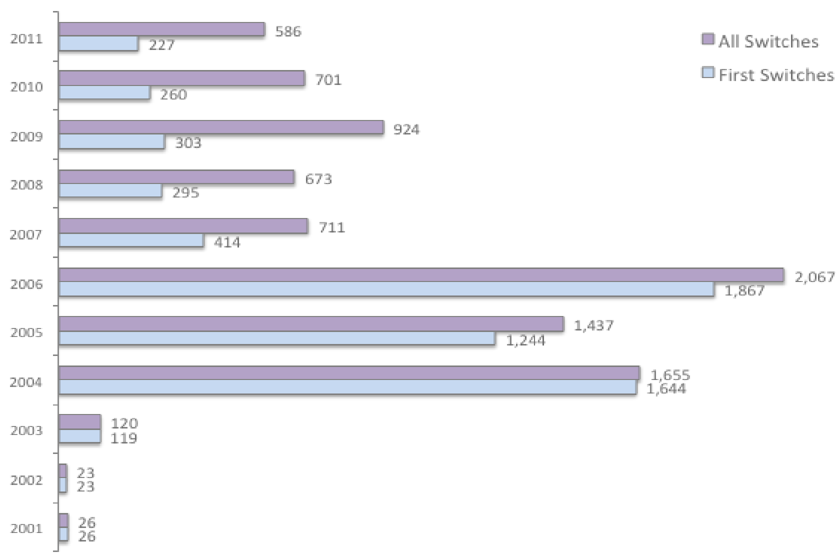
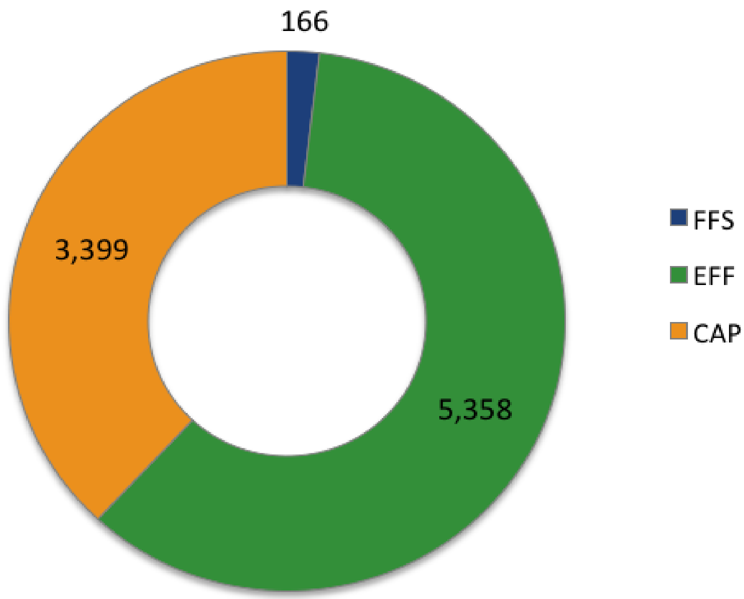


Fig. A3 - # of Switches to Each Model



Note: “FFS” = Fee-for-service, “EFF” = Enhanced fee-for-service, “CAP” = Blended capitation