EFFECTS OF MEDICARE PART D PLAN SWITCHING ON PRESCRIPTION DRUG EXPENDITURES: A LONGITUDINAL ANALYSIS USING MIXTURE MODELS

by

Victor B. Talisa

B.A., St. Mary's College of Maryland, 2009

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Victor B. Talisa

It was defended on

July 20, 2015

and approved by

Thesis Advisor:
Jeanine M. Buchanich, M.Ed., Ph.D., Research Assistant Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Committee Members:
Ada O. Youk, Ph.D., Associate Professor of Biostatistics, Epidemiology and Clinical & Translational Science, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Yuting Zhang, Ph.D., Associate Professor of Health Economics, Department of Health Policy and Management, Graduate School of Public Health, University of Pittsburgh

Jeanine Buchanich, M.Ed., Ph.D

EFFECTS OF MEDICARE PART D PLAN SWITCHING ON PRESCRIPTION DRUG EXPENDITURES: A LONGITUDINAL ANALYSIS USING MIXTURE MODELS

Victor B. Talisa, M.S.

University of Pittsburgh, 2015

ABSTRACT

Background: Medicare beneficiaries enrolled in a Part D drug plan are faced with choosing from a large number of plans with a complex array of attributes. Unnecessarily high spending may lead to cost-related non-adherence (CRN). The extent to which changes in spending are attributable to sponsor renewal of plans offered (compared to termination or consolidation), or beneficiary plan switching (compared to not switching), is not well known.

Research Aims: This study evaluated whether 1) unanticipated plan consolidations or terminations lead to higher gross total and patient out-of-pocket (OOP) costs compared to plan renewals, and 2) plan switching is associated with gross total and patient OOP costs.

Methods: We used Part D data from N=1,187,469 beneficiaries enrolled between 2006 and 2012 who never received income subsidies and who were enrolled continuously through December to January of two consecutive years. Beneficiaries were classified into one of five plan switching groups at each transition period: plan termination—switch (PT-S), plan consolidation—switch (PC-S) or no switch (PC-NS), and plan renewal—switch (PR-S) or no switch (PR-NS). Longitudinal mixed models using an analytical sample of N=17,812 beneficiaries evaluated plan switching group effect on gross total and patient OOP costs, controlling for covariates. Linear mixed models (LMM) and mixture models were compared.

Results: Model diagnostics suggested that the mixture model provided a better fit than LMM. Neither plan termination nor consolidation led to consistently higher gross total spending or

patient OOP costs. The lowest gross costs were observed in the PC-S group, with spending in PC-NS, PR-NS and PR-S estimated to be 4.6%, 6.9% and 5.2% higher, respectively. Patient OOP spending was significantly higher among PC-NS and PR-NS compared to PC-S and PR-S. Estimated differences were between 4.8% and 7.5%, equivalent to $24 to $60 higher annual patient OOP spending among individuals who did not switch compared to those that switched.

Conclusions: The public health impact of our results is that increased spending may not be a significant concern for individuals exposed to plan consolidation or termination. However, plan switching may not reduce spending (and consequently risk of CRN) as much as has been shown to be possible in previous studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

One of the most valuable things I have learned from my experience in academic research is that great mentors are not easy to come by. I can therefore say that it was through sheer luck that I became a student of Prof. Yuting Zhang. I sincerely thank Prof. Zhang for the honor of working with her and for always inspiring the feeling that I was working on a very important project.

For two years Profs. Jeanine Buchanich and Ada Youk have generously given me as much of their own time and honest input as I have asked for. It is due to their persistent guidance and support that I was able to complete this work.

I would also like to thank Dr. Janet Tooze of Wake Forest School of Medicine for her generosity in sharing her SAS code for the mixture model macro she developed.

Thank you.

1.0 INTRODUCTION

1.1 MEDICARE PART D

Medicare Part D was introduced in 2006 as a result of the Medicare Modernization Act of 2003, and was intended to increase seniors' access to prescription drugs [1]. In many ways the program has performed successfully, as the proportion of Medicare beneficiaries without drug coverage was reduced from 33% to 10%, utilization of essential drugs has risen,[2] and consumer satisfaction is high.[3] Despite its successes, Part D has often been scrutinized for its complexity, which many argue has caused seniors to have difficulty benefitting optimally.

These concerns stem from an essential feature which sets Part D apart from Medicare Parts A and B where beneficiaries receive a one-size-fits-all benefit package with uniform coverage for everyone. In Part D, beneficiaries choose their prescription drug coverage from private insurance providers on a government subsidized marketplace. Instead of a uniform benefit, in Part D seniors are faced with a choice between potentially dozens of plans, the exact number and nature of which may change from year to year, and which may vary with respect to a number of attributes. When the program was initiated in 2006 all Medicare-eligible individuals were required to choose a Part D drug plan or present evidence of creditable coverage or face a monetary penalty. Although beneficiaries were permitted to switch plans as needed at any time during 2006, in subsequent years this activity has only been permitted during the designated

open enrollment period, running from October 15 to December 7. Selected plans take effect January 1. As a result of beneficiary's responsibility to select their own drug coverage, their ability of to efficiently navigate the program has been of concern practically since inception.

There are two ways to receive drug coverage through Part D. The first option, prescription drug (PDP) plans, provide stand-alone drug coverage that allows beneficiaries to remain enrolled in fee-for-service Parts A and/or B for medical outpatient and inpatient medical services, respectively. All Medicare beneficiaries are enrolled in Part A, whereas Part B coverage is optional. Medicare Advantage (MA) plans provide an alternative to the fee-for-service structure of Parts A and B, and through which beneficiaries receive their Part A and Part B benefits through private health plans. These plans, also called Medicare Part C or Medicare Plus Plans, are often structured as managed care organizations (MCOs). Sponsors offering MA plans are generally required to offer at least one plan that is extended to provide prescription drug coverage, referred to as MA-PD plans.[4] Beneficiaries pay a capitated (per person) fee to receive Parts A and B benefits through their MA plans, but through the Part D portion of MA-PD plans essentially subject to the same payment mechanisms as beneficiaries enrolled in PDP plans.

A primary criticism of the Part D program is that plan characteristics are confusingly numerous and the benefit schedule is complex. The Centers for Medicaid and Medicare Services (CMS) requires all plans (PDP and MA-PD) to conform to a non-linear benefit schedule in which year-to-date spending marks beneficiaries' progression through 4 benefit phases, ultimately determining their marginal out of pocket (OOP) costs. Each year, the standard PDP plan benefit includes an initial deductible, a second phase characterized by a 25% coinsurance, a coverage gap (i.e. the "donut hole") in which the beneficiary is subject to 100% of costs, and a

final catastrophic coverage period with a 5% coinsurance.[5] The cycle then restarts on January 1st of each year.

The coverage gap, which places significant financial burden particularly on individuals with high drug costs, has been a source of considerable concern. Once an individual has reached the coverage gap, the likelihood of medication discontinuation increases,[2,6] which in turn may lead to greater health risks and associated medical costs.[7] Called cost-related medication non-adherence (CRN), this phenomenon is well-documented in Medicare Part D and has been observed as recently 2014.[8,9,10] Although increased CRN is associated with lower incomes,[10] the problem has also been reported in older persons with mental illness who exceed qualification requirements for the low-income-subsidy (LIS).[11] In 2012 the initial coverage limit for entry into the gap was $2,930 for the standard benefit, and the catastrophic coverage limit was $6,730, although the specific threshold amounts for the standard benefit increase slightly every year.[5] At the start of the Part D program about 94% of PDP plans and 73% of MA-PD plans offered little to no drug coverage in the gap.[12] While the Patient Protection and Affordable Care Act (ACA) of 2010 has required a 50% price discount for brand-name drugs and 14% coverage of generic drug costs in the gap, 94% of PDPs and 62% of MA-PDs still did not offer any additional gap coverage in 2012 above the mandated minimum.[12] The Patient Protection and Affordable Care Act (ACA) of 2010 has made it a priority to address the problem by gradually phasing out the coverage gap by 2020.[13]

Gap coverage and other attributes are permitted to differ between Part D plans as long as sponsors demonstrate that their plans are “actuarially equivalent” to, or have the same overall value as, the CMS-defined standard benefit. This flexibility extends to a variety of characteristics in addition to gap coverage, including specific premiums costs, cost-sharing, formularies,

catastrophic coverage, and utilization management policies for specific medication classes. In fact, the number of plans offering the standard benefit in 2012 was far outnumbered by plans offering an alternative (but actuarially equivalent) benefit package.[14] Furthermore, the specific selection of PDP plans offered may vary across the 34 PDP and 26 designated MA-PD plan regions, and plan offerings can change on a yearly basis. A typical beneficiary had between 45 to 57 PDP plans and 36 to 51 MA-PD plans to choose from in 2009 depending on the county, while in 2014 the average number had decreased to 35 PDP and 15 MA-PD plans.[5,15,16] In 2012, out of 31.5 million Medicare beneficiaries enrolled in drug plans, 19.6 million were enrolled in PDPs and 11.7 million in MA-PDs.[14]

1.2 PART D PLAN CHOICES

As described above, the Part D program is characterized by a complex benefit design, high variability of attributes between plans, and a potentially large choice set for beneficiaries. In a 2006 survey, about 70% of those enrolled felt that there were too many alternative plans to choose from despite a perceived usefulness of these options, 52% had difficulty understanding how Part D works, and 54% felt that it was difficult to determine whether specific medications were covered by the plans on offer.[3] Another early study revealed that 30% of beneficiaries did not know whether their chosen plan provided gap coverage.[17]

Some attribute this confusion to the more general concept of "choice overload" occurring in markets with too many competing products, causing consumers to experience impaired decision making.[15] A study by Tanius *et al*. has shown this to be the case when seniors were asked to rate a series of fictional prescription drug plans after being given just a subset of the

attributes beneficiaries are faced with when choosing among Medicare plans, including drug premium prices, cost sharing prices, and distance to the closest pharmacy. Their results indicate that that increasing choice set size significantly affects individuals' ability to choose the best plan.[19] To avoid unnecessary redundancy in plan offerings on the Part D marketplace, the CMS became more active in 2008 and 2009 in denying sponsors the ability to sell plans that were too similar to plans already sold by the same provider.[18] Moreover, in 2010 CMS issued regulations to eliminate redundant plans and plans with low enrollment in order to further address the problem of choice overload. Due to this policy the total number of PDP plans dropped from 1,576 in 2010 to 1,109 in 2011.[14]

Plan continuity from year to year is not always assured for non-redundant plans, either. There are essentially 4 possible outcomes for a given Part D plan. Most sponsors renew their plan contracts with CMS each year. However, sponsors are free to change particulars such as premium prices and cost-sharing even when plans are renewed.[4,5,12,14,15] Sponsors may also choose to consolidate several pre-existing plans under one contract, for example in response to an acquisition or merger; other plans are split into several different contracts or terminated by the sponsor.[21] In addition, CMS has given yearly ratings to all plans since 2008, out of 5 stars. Plans with ratings of less than 3 stars for 3 consecutive years are subject to termination.[14]

Given the complexity of the Part D landscape, what kinds of decisions have beneficiaries made with regard to plan choice? In a given year, the non-LIS population was likely to be enrolled in plans that are more than $300 more expensive than the cheapest plan, and often selected plans with features that were overprotective such as generic coverage in the gap.[21,22] It is estimated that about 5.2% of beneficiaries choose the cheapest plan.[21]

1.3 PLAN SWITCHING

Despite being enrolled in plans with seemingly excessive and inappropriate features and costing hundreds of dollars more per year on average than necessary, most Medicare Part D beneficiaries choose not to change plans. Among all non-LIS beneficiaries continuously enrolled in a PDP plan, only about 13% switched plans in a given year between 2006 and 2010, with more than 70% remaining in the same plan throughout that time.[20] Realizing this, the CMS recommends that beneficiaries review their enrollment annually.[23-25] CMS has also implemented a web-based tool to facilitate this process, the Medicare Plan Finder, enabling beneficiaries to quickly compare plans and identify differences in characteristics and pricing. Despite these attempts plan switching rates are relatively low and effective plan optimization is uncommon. The effects of this tendency for Part D beneficiaries to remain enrolled in the same plan year after year are well illustrated by Han *et al*. who point out that in 2006 the insurance company Humana had the lowest premiums and nearly the highest enrollment among all Part D plans.[26] Over the next three years Humana increased premiums by up to 4 times while maintaining an equivalent share of the market. This trend is true in general, as premiums for both PDP and MA-PD plans have continued to rise on average almost 50% from 2006 to 2012.[14] Han *et al.* propose that suboptimal plan choices of most beneficiaries may be due to their reluctance to re-evaluate plans each year, a process which they necessarily experienced in their first year of Part D enrollment and may not want to revisit.[14] This corresponds with results from a study suggesting that seniors may base their decisions disproportionately on premium prices, which were lowest in the early years of Part D.[27]

However, the degree to which this population has learned to navigate the Part D program by exercising their ability to switch plans each year is not as well understood. The only studies to

examine this directly have examined overspending, calculated as the difference between actual spending and hypothetical OOP costs under the plan with the minimum OOP costs for that individual. Ketcham *et al*. observed decreases in overspending 2006 to 2007 among non-LIS beneficiaries and attributed the savings to consumer learning and plan switching, which was more frequent among those who increased their savings.[28] However, a separate study found that overspending increased in 2007-2008 compared to 2006-2007.[22] While these efforts have increased our understanding of the dynamics of plan choices in the Part D program, they largely been limited by cross-sectional approaches or consideration of just a few years of data.[20,22,28] As a result these studies have not been equipped to control for time when measuring the effects of switching on plan outcomes. Moreover, these studies have uniformly excluded individuals' enrolled in Part D plans that are terminated or consolidated with other plans by the sponsor. In such unanticipated situations the beneficiary is involuntarily "forced" to either switch plans (as in the case of plan terminations) or may be unprepared to appropriately evaluate automatic changes in their plans (plan consolidation). Beneficiaries who do not switch plans also may or may not understand that even renewed plans' attributes may not be consistent with previous years. In the LIS population, which is subject to automatic plan enrollment and plan switching based on the level of coverage offered by the subsidy, forced plan switching has been shown to lead to overspending and potentially harmful utilization review policies that limit beneficiaries' use of necessary medications. It therefore stands to reason that the portion of the non-LIS population which is subject to unanticipated or unwanted changes in its prescription drug plans, yet does not react to them, may potentially be negatively affected.

1.4 RESEARCH OBJECTIVES

The effects of plan switching under the potentially unanticipated or involuntary circumstances of plan terminations or consolidations is not well documented. Furthermore, studies of plan switching have focused on measuring the benefits of a hypothetical plan switch by calculating the difference between actual spending and hypothetical spending under a cheaper plan. This approach may not be accurate if assumptions incorporated into the hypothetical spending calculations are not met, for example if economic demand for a particular drug is not stable given changes in prices.[28] Finally, the handful of studies evaluating the effects of switching plans in the non-LIS population have been mostly limited to a couple years of information, leading to conflicting reports as to the degree of benefit associated with switching plans and the strength of choices across a wide range of years.

In order to address these gaps in the literature, we conducted a longitudinal study of non-LIS beneficiaries enrolled in Part D between 2006 and 2012, incorporating actual spending. The current study therefore seeks to evaluate two primary research questions:

1) Is there a negative effect of plan terminations and plan consolidations on gross total costs and patient OOP spending compared to plan renewals?
2) Is plan switching associated with lower gross total costs and OOP spending compared to not switching plans?

2.0 METHODS

2.1 STUDY SAMPLE

2.1.1 Exclusion criteria

The current study utilized Medicare data from years 2006 to 2012, inclusive. Two exclusion criteria were applied when constructing the study sample. First, we excluded beneficiaries that had received the LIS or state buy-in subsidy during at least one month during the study period. Included in this group were individuals who were dually eligible for Medicaid as well as Medicare (dual-eligible), which entitles individuals to copayments as low as $0 throughout the benefit cycle, including the coverage gap. These individuals may therefore not be as sensitive to changes in their plan characteristics that might provide motivation for switching among the non-poor. Furthermore, dual-eligible beneficiaries are automatically randomized to a low-cost "benchmark" PDP plan unless they make another choice. If they chose a plan costing more than the benchmark amount covered by the LIS, the beneficiary is responsible for the difference. This feature not only provides the motivation to remain enrolled in a smaller selection of "benchmark" plans, but it also limits the ability of beneficiaries to remain passively enrolled in the same plan if that plan loses "benchmark" status. The effects of automatic assignment for LIS

beneficiaries have been studied elsewhere, [29,30] and for the purposes of the current study this sub-population has been removed.

Although CMS policy requires enrollment in some type of drug plan, continuous enrollment in Part D is not required if beneficiaries elect to receive benefits through another provider (e.g. receive "creditable coverage"). Some beneficiaries were not continuously enrolled in the Part D program, and re-enrolled after a period of time during which they did not receive the benefit. These beneficiaries were not necessarily excluded from the study. However, after the first discontinuation they were treated as though they did not re-enroll; in other words beneficiaries were treated as not enrolled in Part D for the entire period following first discontinuation.

The second exclusion criterion was applied after determining the first month of discontinuation. This criterion stipulated that individuals were excluded if they had not been enrolled in a Part D plan for at least one continuous two-month period from December of one year to January of the next year (i.e. Dec-Jan transition period). The focus of the study is on end-of-year plan switching in which new plans chosen during the enrollment period are initiated in January. Therefore, beneficiaries not enrolled during the Dec-Jan transition period were not able to contribute to the analysis and were excluded.

The application of the second exclusion criterion is represented in Figure 1. Panel A represents hypothetical enrollment data from an individual who was enrolled in a Part D plan from January to December of 2008, but not enrolled for all other months. Because this period does not include a Dec-Jan transition period, this individual would have been excluded from the final sample. On the other hand, the beneficiary represented in Panel B was continuously enrolled during the period from October 2007 to December 2012, a timeframe including 5

transition periods. The individual in Panel B would have been included in the sample. Finally, panel C1 illustrates a scenario where a beneficiary was enrolled for a continuous period from December 2006 to January 2008, not enrolled for the duration of 2009, and then enrolled again for a period beginning in January 2010. Panel C2 shows enrollment information for this individual as it was analyzed for the current study, where the individual was treated as not enrolled for the entire period following first discontinuation.

2.1.2 Sample size

Table 1 provides sample sizes before and after exclusion criteria were applied. Prior to excluding any beneficiaries, each annual dataset had between 2.5 and 3 million individuals. After merging the annual datasets together and removing individuals who were not enrolled in Part D through at least one enrollment period, the sample size included just over 2 million unique beneficiaries. After fully restricting the sample to those who never received the LIS, final total sample consisted of N=1,187,469 unique beneficiaries.

2.2 PLAN SWITCHING

Although Part D plan enrollees are permitted to switch prescription drug plans at any time, they are disincentivized from switching outside of the designated Part D enrollment period by charging a fee for this activity. In addition, plan termination, consolidation and splitting only take effect on January 1. Therefore this study focused on plan switching events taking effect after the Dec-Jan transition period.

Enrollment period plan switching was classified into five types depending on whether switching occurred and whether the previous year's plan was renewed, terminated, consolidated with other plans or split into multiple plans. Most plans are renewed by the plan sponsor each year. Under these circumstances a beneficiary may choose to keep their enrollment in this plan or switch to a new one. In certain cases, several plans may be consolidated into one plan the following year, or alternatively one plan may be split into several. However, consolidations occur much more frequently than splitting (not shown), and henceforth all consolidations and splits will be simply referred to as consolidations. In cases of plan consolidation the enrollee is notified of the change and is automatically enrolled in the new plan unless the enrollee voluntarily chooses another. A plan that is marked as terminated indicates that the sponsor did not renew the previous year's plan, nor was it consolidated. Terminated plans lead to the enrollee losing Part D coverage unless they voluntarily choose a new plan. To summarize, the five types of switching are as follows:

1. P**lan termination—Switch (PT-S):** The previous year's plan was terminated by the sponsor and the enrollee was forced to actively choose another plan to receive any coverage.
2. **Plan consolidation—No switch (PC-NS)**: The previous year's plan was consolidated or split by the sponsor and the beneficiary chose to remain enrolled in the "matched" plan the following year.
3. **Plan consolidation—Switch (PC-S)**: The previous year's plan was consolidated and the beneficiary chose to switch to a new plan the following year.

4. **Plan renewal—Switch (PR-S)**: The sponsor renewed the plan the beneficiary was previously enrolled in, but the beneficiary actively chose a new plan.
5. **Plan renewal—No switch (PR-NS)**: The sponsor renewed the plan the beneficiary was previously enrolled in, and the beneficiary chose to remain enrolled.

2.3 OUTCOMES

The primary motivation for this study was to gain a better understanding of the effects of plan switching on prescription drug plan spending. Spending was evaluated on two levels: gross total drug cost, and patient OOP spending.

Gross total drug cost was calculated as the sum of the ingredient cost, dispensing fee, sales tax, and the vaccine administration fee if applicable. This value reflects the price paid for the drug at the point of sale. Therefore the gross total cost included a portion covered by the sponsor, the portion that the patient pays out of pocket, and a portion paid by third-party payers if applicable. Exclusion criteria limited the research sample to those without the LIS, excluding the federal government as a possible payer, but other third party payers may have included employers and liability insurers.[31] The amount paid out of pocket by the beneficiary was examined as a separate outcome. This amount represented the cost of the prescription drug event paid by the beneficiary, which includes all copayments, coinsurance and deductibles.

2.4 STATISTICAL ANALYSIS

2.4.1 Analysis of repeated measures data using linear mixed models

In order to understand the effects of switching on drug spending over time, we require a statistical analysis which appropriately evaluates this relationship. Because we are interested in estimating the effects of switching on plan spending, adjusting for covariates, the logical starting point is the basic class of multiple linear regression models. Assuming that our data are normally distributed, let us define this model as:

$$\boldsymbol{y} = \mathbf{X}^{\mathbf{T}}\boldsymbol{\beta} + \boldsymbol{\varepsilon} .$$

Let **y** denote the vector of observed outcome values $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}'$ where n is the total number of observations; let $\boldsymbol{\beta}$ denote the vector of fixed effects parameters $\boldsymbol{\beta} = \{\alpha, \beta_1, \beta_2, \ldots, \beta_p\}'$ where α is the common intercept and p is the number of fixed effects; **X** denotes the design matrix of dimension $n*(p+1)$, containing covariate data for each observation; and **ε** represents the vector of residuals $\boldsymbol{\varepsilon} = \{ \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}'$, which are assumed independent and identically distributed (IID) random variables distributed normally with mean $y_i - \mathbf{X}_i^{\mathbf{T}}\,\boldsymbol{\beta}$ and variance σ_e^2, where y_i and $\mathbf{X}_i$ are the values of y and vector of predictor values corresponding to the i^{th} individual, respectively. Note the independent variance structure (i.e. all covariance terms are equal to 0) for the observations assumed in this model represented by $n*n$ matrix **E:**

$$\mathbf{E}_{nxn} = \begin{pmatrix} \sigma_e^2 & 0 & 0 & & 0 \\ 0 & \sigma_e^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_e^2 & & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_e^2 \end{pmatrix},$$

Furthermore, each observation in **y** is assumed to have been generated from a normal distribution with a mean conditional only on the coefficients $\boldsymbol{\beta}$ (covariate values in X are considered fixed):

$$f\,(y_i \mid \boldsymbol{\beta}, \sigma_e) = (2\pi\sigma_e^2)^{-1/2} * \exp\left\{\frac{-1}{2\sigma_e^2}\left(\boldsymbol{X_i^T\beta}\right)^2\right\}$$

Consistent estimation of the fixed effects parameters would be achieved through maximization of the likelihood function, which in every case is the joint distribution of the observed data:

$$L(\boldsymbol{\beta}, \sigma_e\,; y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} f\,(y_i \mid \boldsymbol{\beta}, \sigma_e)$$

However, as with most longitudinal studies, our data exhibit a "clustered" structure, due to measurements taken repeatedly over time for each individual and therefore the typical assumption of independent observations required by basic linear regression models is violated. While it would be possible to address this problem by introducing n-1 individual-specific intercept terms as fixed effects, this approach is problematic in our case due to the large sample size n. Furthermore, maximum likelihood estimation of the fixed effects parameters of interest would lead to biased variance estimates, due to the fact that the number of nuisance parameters (the n-1 individual specific intercept terms) grows proportionally with the sample size.[32]

Linear mixed models (LMM) provide a much more elegant solution to the problem of within-subject dependency exhibited by longitudinal data. Instead of treating each subject (i.e. cluster-specific) effect as fixed, mixed models assume the cluster-specific effects are a random sample from a population distribution. This cluster-specific term, called the random effect or heterogeneity, is constant over time for each subject, and is often assumed to have a normal distribution with mean 0 and variance σ_u^2. Because our data come from a continuous distribution, it may not be unreasonable to assume the following model structure:

$$y_{it} = \boldsymbol{X_{it}^T\beta} + \boldsymbol{R_i^T} u_i + \boldsymbol{\varepsilon_{it}}\,.$$

$$\boldsymbol{\varepsilon_{it}} \sim N\,(0, \sigma_e^2)$$

$$u_i \sim N\,(0, \sigma_u^2)$$

Where i =1, … n for each subject, and t=1,…k_i for each repeated observation on a specific subject. The design matrix for the random effect, $\boldsymbol{R}$, is of dimension m*n, where m is equal to the total number of observations and n is the sample size. The addition of the random term allows for a separation of within-subject source of variation (σ_e^2) and between-subjects variation (σ_u^2). Secondly this specification allows for a more appropriate variance-covariance matrix in which observations y_{it} within a subject may be correlated. Within-subject residuals $\boldsymbol{\varepsilon_{it}}$ are usually assumed to be independent, and therefore the variance-covariance matrix of the residuals is still represented as diagonal matrix:

$$\mathbf{E} = \mathrm{Var}(\,\boldsymbol{\varepsilon}\,) = \begin{pmatrix} \sigma_e^2 & 0 & 0 & & 0 \\ 0 & \sigma_e^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_e^2 & & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_e^2 \end{pmatrix},$$

with dimension m*m. However, now we also have a between-subjects covariance matrix with dimension n*n:

$$\mathbf{G} = \mathrm{Var}(\,u\,) = \begin{pmatrix} \sigma_u^2 & 0 & 0 & & 0 \\ 0 & \sigma_u^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_u^2 & & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_u^2 \end{pmatrix}.$$

The variance of the vector of outcomes $\boldsymbol{y}$ is therefore:

$$\mathrm{Var}(\,\boldsymbol{y}\,) = \mathrm{Var}(\,\boldsymbol{R^T u}\,) + \mathrm{Var}(\,\boldsymbol{\varepsilon}\,)$$

$$= \boldsymbol{R^T} * \mathrm{Var}(\,u\,) * \mathbf{R} + \mathrm{Var}(\,\boldsymbol{\varepsilon}\,)$$

$$= \begin{pmatrix} V_1 & 0 & 0 & & 0 \\ 0 & V_2 & 0 & \cdots & 0 \\ 0 & 0 & V_3 & & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & V_n \end{pmatrix}.$$

Where V_i is of dimension k_i*k_i:

$$V_i = \begin{pmatrix} \sigma_u^2 + \sigma_e^2 & \sigma_u^2 & \sigma_u^2 & & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 + \sigma_e^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 + \sigma_e^2 & & \sigma_u^2 \\ & \vdots & & \ddots & \vdots \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 + \sigma_e^2 \end{pmatrix}.$$

A third advantage is that mixed models allow available data for each subject to be missing, as in our case where data for each beneficiary are unbalanced. That is, each beneficiary is followed from their specific year of enrollment until first discontinuation of Part D, and the start and end dates are permitted to differ. LMM utilize between-subject information as well as within-subject information to compensate for the presence of missing values, and generally give more precise estimates than fixed effects models assuming that the missing data are missing at random (MAR).

The gains that follow from using LMM are accompanied by increased complexity in parameter estimation arising from the addition of the random effect. While the random effect is useful for accounting for subject-specific variation, we are not interested in interpreting the estimated random parameter u_i and therefore it is regarded as a nuisance parameter that must be dealt with to estimate the parameters of interest, the fixed effects **β** and the variance parameters. We obtain the marginal distribution of the observed data, $f(y|\boldsymbol{\beta})$, by integrating the joint distribution $f(y_i, u_i|\boldsymbol{\beta}, \sigma_e, \sigma_u)$ with respect to the random effect:

$$f(y_i \mid \boldsymbol{\beta}, u_i) = \int_{\infty}^{+\infty} f(y_i, u_i \mid \boldsymbol{\beta}, \sigma_e, \sigma_u)\, du_i$$

The contribution to the marginal likelihood of the i^{th} subject is therefore

$$L_i(\boldsymbol{\beta}, \sigma_e, \sigma_u; y_{i,t=1}, y_{i,t=2}, \ldots, y_{i,t=k_i}) = \prod_{t=1}^{k_i} \int_{\infty}^{+\infty} f(y_i, u_i \mid \boldsymbol{\beta}, \sigma_e, \sigma_u)\, du_i$$

and the likelihood which can be used to obtain maximum likelihood estimates as follows:

$$L(\boldsymbol{\beta}, \sigma_e, \sigma_u) = \prod_{i=1}^{n} \prod_{t=1}^{k_i} \int_{\infty}^{+\infty} f(y_i, u_i \mid \boldsymbol{\beta}, \sigma_e, \sigma_u)\, du_i$$

The problem with maximizing the marginal likelihood is that often the integral cannot be solved analytically, necessitating alternative methods of computation. Several solutions have been developed, and may be categorized into two types of methods: 1) approximate the integral numerically, such as through adaptive quadrature or the Expectation-Maximization (EM) method, or 2) approximate the integrand such as through Laplace's approximation or quasi-likelihood methods.[33] In order to achieve model convergence, Laplace's approximation was specified for the present study.

As stated previously, the main objective of our analysis is the consistent and efficient estimation of the fixed effects $\boldsymbol{\beta}$ and the variance covariance matrix Var($\boldsymbol{y}$). The key assumptions of the LMM with random effects employed in our analysis are as follows:

1) Linearity: the relationship between each of the predictors and the mean (or transformed mean) outcome is linear.
2) Homoscedasticity: the variance of the within-subjects error term is constant
3) Residuals are normally distributed, with mean $y_i - \mathbf{X}_i^{\mathbf{T}} \boldsymbol{\beta}$ and variance σ_e^2.
4) Compound symmetry: all covariance terms in the covariance matrix are equal, and all variance terms are equal (but covariance may be different from variance terms).
5) Strict exogeneity: the within-subject error terms are not dependent on the values of any independent variables for all time-points:

$$E[\varepsilon_i \mid x_{t1}, x_{t2}, \ldots, x_{tp}] = 0, \text{ given } i, \text{ for all } t=1, 2, \ldots, k_i.$$

6) Random effects u_i are independent across subjects, independent of all covariates, and are normally distributed with mean 0, variance σ_u^2. [34]

2.4.2 Healthcare expenditure data and mixture models

Our analysis involves modelling healthcare expenditure data, specifically gross total and patient OOP costs. Although both of these outcomes are continuous, they exhibit some characteristics which are known to cause problems with models assuming a normal distribution, as the LMM does. First, expenditure data are bounded on the left at 0, often have a long right-hand tail, and may be bi-modal. Second, log or square root transformations are often not normalizing for these data due to the frequently high counts of individuals with expenditures of $0 exhibited by non-users.[35] These qualities have led to the development of novel approaches for data of this type, research which is largely represented in the cost-effectiveness and health econometrics literature. A recent review of these methods has noted the lack of a gold-standard approach that provides unbiased and efficient estimates. [36]

One class of models that has shown promising results for data with excess zeros is the 'two-part' models. These models are mixed-distribution (i.e. mixture) models designed to analyze the data in two parts: a degenerate component (the occurrence of zero) and a continuous component with a support set with an open boundary at 0 to positive infinity. These two components are fit to separate models and estimated separately, an approach which has been shown to perform better than single-equation models. [37]

One such two-part model design which has been tailored to applications with clustered data, and which has been adopted for the present study, was developed and evaluated by Tooze *et al*. [38] In general, we can express a mixture model for a random variable Y as follows:

$$f(y) = \begin{cases} \Pr(Y = 0), & \text{if } y = 0 \\ [1 - \Pr(Y = 0)]h(y) & \text{if } y > 0 \\ 0 & \text{if } y < 0 \end{cases}$$

where $h(y)$ is the probability density for values of $y > 0$. Applying this to our data, we can define a variable Z_{it} to represent whether or not an observed outcome Y_{it} from individual i at time t is equal to 0:

$$Z_{it} = \begin{cases} 0 & \text{if } Y_{it} = 0 \\ 1 & \text{if } Y_{it} > 0 \end{cases}$$

We model the relationship between this outcome and the covariates separately from the values of $Y_{it} > 0$. Because Z_{it} is a binary outcome, we must take advantage of the class of generalized linear mixed models (GLMM) which models the mean of a binary random variable (such as Z_{it}), with range between 0 and 1, using the logit transformation. We therefore model the probability that Z_{it} is equal to 0 (or in our case we model probability of non-zero) using a logistic model with random effects:

$$\text{logit}[\ P(Z_{it} = 1)\ ] = \ln\left(\frac{P(Zit=1)}{1-P(Zit=1)}\right) = \boldsymbol{X^T\ \beta_Z} + \boldsymbol{R^T}\ u_{Zi}$$

$$u_{si} \sim N\ (0,\ \sigma_1^2)$$

Let $\boldsymbol{\beta}_Z$ denote the vector of fixed effects parameters $\boldsymbol{\beta}_Z = \{\alpha_Z, \beta_{Z1}, \beta_{Z2}, \ldots, \beta_{Zp}\}'$ specific to the logistic model; α_Z is the common intercept and subscript p is the number of fixed effects; $\mathbf{X}$ denotes the design matrix of dimension n*(p+1), containing the data for covariates for each observation, and u_{Zi} is the random effect specific to the logistic model of Z_{it}. Using the nomenclature from Tooze et al., we refer to this as the *occurrence variable*. By solving for the probability of observing $Z_{it} = 1$ we can obtain the expression

$$P(Z_{it} = 1|\ \boldsymbol{\beta_Z}) = \boldsymbol{\exp(\ X^T\ \beta_Z} + \boldsymbol{R^T}\ u_{Zi}\boldsymbol{)/(\ 1+\exp(X^T\ \beta_Z} + \boldsymbol{R^T}\ u_{Zi}\boldsymbol{)\ )}\ ,$$

As mentioned above, we also model separately the individuals with values of the outcomes for which $Y_{it} > 0$ is true. Define this random variable $S_{it} = [\ Y_{it} \mid Z_{it} = 1]$ to be the *intensity variable*,

again following the example of Tooze et al. For this we assume a classic LMM structure with normally distributed random error:

$$s_{it} = \boldsymbol{X_i^T \beta_S} + \boldsymbol{R_i^T} u_{Si} + \boldsymbol{\varepsilon_{it}} .$$

$$\boldsymbol{\varepsilon_{it}} \sim \mathrm{N}\ (0, \sigma_e^2)$$

$$u_{si} \sim \mathrm{N}\ (0, \sigma_2^2)$$

Note here that the parameters $\boldsymbol{\beta_S}$ are specific to the model of S_{it}, as is the random error term u_{Si}. Define $\boldsymbol{\theta_Z} = \{ \boldsymbol{\beta}_Z , u_{Zi} \}$ to be the set of parameters specific to the logistic component of the mixture model, and similarly $\boldsymbol{\theta_S} = \{ \boldsymbol{\beta}_S , u_{is} \}$ are the set of parameters specific to the model of the intensity variable. We can now combine these two components in order to define the probability density function of all values of Y_{it} as follows:

$$f\ (y_{it} \mid \boldsymbol{\theta}_Z, \boldsymbol{\theta}_S) = \mathrm{P}(Z_{it} = 0 \mid \boldsymbol{\theta}_Z)*\mathrm{I}[Z_{it} = 0] + \mathrm{P}(Z_{it} = 1 \mid \boldsymbol{\theta}_Z)* f\ (s_{it} \mid \boldsymbol{\theta}_S) ,$$

where $\mathrm{I}[Z_{it} = 0]$ is the indicator function and is equal to 1 when the argument is true and 0 when it is false; let $f\ (s_{it} \mid \boldsymbol{\theta}_S)$ be the probability density function for random variable S_{it} conditional on intensity variable model-specific parameter values $\boldsymbol{\theta}_S$. We can obtain estimates for $\boldsymbol{\theta}_Z$ and $\boldsymbol{\theta}_S$ by maximizing a likelihood function. As before, the random effects have a distribution which must be taken into account in the likelihood. However, because we have two random effects, one from each model component, we assume the random effects are jointly bivariate normally distributed:

$$f\ (u_{Zi}, u_{Si}; \sigma_1 , \sigma_2) = \left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}\right)^{-1} * \exp\left\{\frac{-1}{2\sqrt{1-\rho^2}}\left[\left(\frac{u_{Zi}}{\sigma_1}\right)^2 - 2\rho\left(\frac{u_{Zi}}{\sigma_1}\right)\left(\frac{u_{Si}}{\sigma_2}\right) + \left(\frac{u_{Si}}{\sigma_2}\right)^2\right]\right\}$$

where σ_1^2 is the variance of u_{Zi} and σ_2^2 is the variance of u_{Si}. As with the linear models case, we are interested in maximizing the marginal likelihood function obtained by integrating each observation's contribution to the likelihood with respect to both random variables:

$$L\ (\boldsymbol{\beta_Z}, \boldsymbol{\beta_S}, \sigma_e, \sigma_1, \sigma_2) = \prod_{i=1}^{n} \prod_{t=1}^{k_i} \int_{u_{Zi}} \int_{u_{si}} f\ (y_i, u_{Zi}, u_{Si} \mid \boldsymbol{\beta_Z},\ \boldsymbol{\beta_S}, \sigma_e, \sigma_1, \sigma_2)\ du_{Zi}\ du_{Si}$$

$$= \prod_{i=1}^{n} \prod_{t=1}^{k_i} \int_{u_{Zi}} \int_{u_{si}} f\ (y_{it} \mid \boldsymbol{\theta_Z}, \boldsymbol{\theta_S})\, f\ (u_{Zi}, u_{Si} \mid \boldsymbol{\beta_Z}, \sigma_1, \sigma_2)\ du_{Zi}\ du_{Si}$$

Note that if we assume the random effects are independent, this function could be factored into two parts representing the marginal likelihoods of the logistic and normal components. In other words, maximizing the likelihood functions of each model separately will yield MLE's for the mixture model if we assume ρ = 0. However, the model developed by Tooze allows for estimation of correlated random effects using PROC NLMIXED in SAS. [38]

2.4.3 Study design, model fitting and evaluation

All statistical analysis and data management was performed using SAS version 9.4. Using the total sample obtained after applying exclusion criteria (N=1,187,469), descriptive statistics were calculated, including frequencies and percentages as well as group means and standard deviations. Due to the complexity of the random effect models and the size of the total sample, a 1.5% random sample (N=17,812) of the beneficiaries from the total sample was used for analysis using mixed models, referred to henceforth as the analytical sample. The size of this sample was determined through trial and error, and is approximately the largest sample that would lead to convergence of the models of interest.

In order to understand the effect of plan switching on costs in the year after the switch occurred, two types of models were fit and compared for each outcome. First, a LMM was fit for each outcome, including an individual random effect and adjusting for covariates. The empirical distributions of each outcome were examined using histograms and transformations were applied if necessary in order to obtain approximate normality of the outcomes. Second, a mixture model

consisting of logistic and lognormal components, and allowing for correlated random effects (section 2.5.2), was applied wherever possible using the SAS macro MIXCORR written by and obtained from Tooze *et al*.[38] The MIXCORR macro operates by first defining the occurrence and intensity variables. These variables are then modeled separately in regression models without random effects using PROC GENMOD and specifying the appropriate distribution (e.g. binomial for occurrence and lognormal for intensity). The estimates produced from PROC GENMOD are then used as starting values in the fully-specified correlated random effects mixture model fit using PROC NLMIXED. While MIXCORR also fits an uncorrelated random effects mixture model (assuming random effects from the logistic and lognormal components are independent), we focus on the results from the correlated model. Laplace approximation was specified for all mixtures models using the QPOINTS=1 option in PROC NLMIXED. LMM models were fit using a procedure that mimicked MIXCORR; a fixed effects model was fit using PROC GENMOD and the estimates were used as starting values in a random effects model fit using PROC NLMIXED and Laplace approximation.

The fit of LMM and mixture models was evaluated and compared by assessing the normality of random effects in Q-Q-plots and histograms. In addition, scatter plots of residuals (fixed + random effects) from each model were generated and compared. Point estimates and standard deviations for fixed effects were also compared for each model. All models evaluated the effect of each type of plan switching controlling for covariates. Parameter estimates were considered statistically significantly different from 0 when $p<0.05$. The overall contribution of switching to each model was evaluated, and if statistically significant, pairwise differences between each combination of switching category were evaluated using CONTRAST statements.

Unless otherwise noted, all models adjusted for all covariates. Time independent covariates included gender, race, age at baseline, plan type (whether the beneficiary stayed enrolled in PDP plans, MA-PD plans, or switched between PDP and MA-PD plans) and total number of active plan switches (defined as sum of PT-S, PC-S, PR-S, as well as any mid-year switches). Models also adjusted for region of residence at baseline, where beneficiaries were included into one of 5 state groupings (Table 2), each consisting of two or more PDP or MA-PD regions. These groupings were created in order to reduce the number of dummy variables needed to adjust for geographic location model, as the 34 PDP regions and 26 MA-PD regions do not correspond 1:1 and were too numerous to include without aggregating. Baseline levels of the outcome were also included, defined as the outcome value observed during the first year of enrollment in a Part D plan. Models also adjusted for time-varying calendar year and total number of CCW chronic conditions. Baseline levels of the outcome, total number of plan switches and total number of CCW chronic conditions were square-root transformed and standardized in order to avoid model convergence errors associated with large scaling differences between covariates.

2.5 DATA MANAGEMENT

2.5.1 Datasets

The CMS Chronic Conditions Data Warehouse (CCW) provided annual data from a random 5% sample of all Medicare beneficiaries. Data were distributed across several different files, including dedicated files for enrollment information, Part D drug events, Part D plan

characteristics, Part D premiums, etc. All datasets were de-identified, but CCW provided unique beneficiary link keys ("bene id") in order to facilitate merging information from the same beneficiary across files within a year as well as between years.

The current study utilized files containing data from years 2006 to 2012. During this timeframe CCW changed the organization and conventions of several datasets, a shift that did not involve changes to the data collected but did impose additional data management issues. Prior to 2010, some beneficiary-level information such as enrollment, eligibility, vital statistics, summarized service utilization and payment, and chronic condition flags were specifically located in the Beneficiary Annual Summary File (BASF), while other enrollment and entitlement variables were located in a Master Beneficiary Summary File (MBSF). From 2010 forward, the MBSF was split into several separate files, each containing specific information such as Medicare Part A/B/D enrollment and entitlement information, chronic conditions, cost and utilization and national death index (NDI), for vital statistics. In addition, starting in 2010 these separate MBSF files incorporated the variables formerly contained in the BASF, which was discontinued.

Data reorganization particularly affected the location of beneficiary demographics, enrollment and entitlement information. The change of file location of key enrollment/demographic information by year is shown in Table 3. The current study utilized enrollment and demographic data from the BASF and MBSF files for years 2006 to 2009, while for years 2010 to 2012 the MBSF files corresponding to Medicare Parts A, B and D were used. These files included many key variables used in determining study eligibility: monthly categorizations of enrollment in Parts A, B and D, as well as eligibility or receipt of LIS; total months of state buy-in; and monthly Part D plan contract ID. They also contained gender, age,

race and state code variables. Comorbidity data was included in the general MBSF file until 2009, but after 2010 was located in a dedicated MBSF Chronic Conditions segment. The presence of common and chronic conditions is captured by 27 CCW Chronic Conditions variables, using claims-based algorithms as a proxy for evidence of the presence of a condition.

Finally, each year Part D plan contracts between plan sponsors and CMS are reevaluated, and may lead to plan terminations, consolidations or plan splitting. These outcomes are recorded for all Part D plans in the plan crosswalk files, which were in turn used to determine plan switching status for each beneficiary from December to January.

2.5.2 Constructing the research sample

Application of the exclusion criteria was performed during assembly of the research sample using SAS. Before combining the annual datasets, some beneficiaries were removed based on logical extensions of the exclusion criteria. For example, beneficiaries not enrolled in a Part D plan in December of 2006 necessarily met at least one exclusion criterion and were removed. Similarly, individuals who were not enrolled in a Part D plan in both January and December of the same given year were excluded from that annual dataset. This process minimized the size of the annual datasets in order to reduce the amount of processing time needed for the merge operation.

Once the initial pruning process was completed, the annual datasets were merged by beneficiary ID. At this stage a string variable was created consisting of a single character for each month of the maximum 84 months of enrollment through the 7 years of the study, similar to the 84-character strings in Figure 1. The character code and its position in the 84-character string

were used to identify whether a beneficiary was “enrolled” or “not enrolled” in a Part D plan for each month of the 7-year study period. If a beneficiary was identified as having discontinued their Part D enrollment during the 84 months, all months following first discontinuation were identified as “not enrolled,” (see section 2.1.1 above). Using this generated string variable, exclusion criteria were applied. LIS status was determined using cost share group information recorded monthly. Beneficiaries were removed if deemed eligible for any amount of LIS subsidy or if their total months of state buy-in exceeded 0 months. Any individual who was not enrolled in a Part D plan for at least one continuous Dec-Jan transition period was also removed.

2.5.3 Plan switching

Plan switching types were assigned using the plan crosswalk files for each year as well as the beneficiary’s plan IDs for the two months of each Dec-Jan transition period. Enrollment datasets included encrypted Part D plan IDs for each month of the year. The plan cross-walk file for a given year contained encrypted plan IDs from that year, as well as the corresponding ID or IDs the plan was mapped to the following year (see Figure 2, Panel A for an example). In addition, the crosswalk file contained a relationship code communicating whether the plan was terminated (in which case there would be no corresponding ID the following year), consolidated (many plans would have the same ID the following year), split (a single plan would have many corresponding IDs the following year), or renewed (same unique ID the following year). Figure 2, Panel B contains a step-by-step schematic example illustrating how plan switching group was determined for example data from transition period 2007 to 2008. Each beneficiary’s plan ID from December of a given year (e.g. 2007 in the example) was merged with the ID for the plan in the crosswalk file corresponding to the same year (e.g. 2007). Next, the beneficiary’s plan ID

from January of the following year (e.g. 2008) was merged with the corresponding ID from the crosswalk file. Switching types were assigned by comparing the actual January plan ID from the second year in the pair (e.g. 2008) to the plan ID in the crosswalk file as well as the relationship code. In the example in Figure 2, beneficiary 1 was assigned PR-NS because the beneficiary's plan from year 2008 matched the plan ID from the crosswalk file for year 2008 and the relationship code indicated that the plan had been renewed. Beneficiary 2 was assigned PC-NS was assigned because actual plan ID for 2008 matched the cross-walked 2008 ID for a consolidated plan. Finally, beneficiary 3 was assigned PC-S because plan ID's did not match, and the plan was consolidated.

A beneficiary continuously enrolled in Part D for Y years would have participated in at most Y-1 Dec-Jan transition periods. Therefore, for modelling purposes, plan switching classifications were associated with spending outcomes from the year following the transition period. Outcomes from the first year of a beneficiary's enrollment were not associated with any plan switching group and were treated baseline levels of spending.

2.5.4 Outcomes

Total gross costs and patient OOP spending were calculated as the average monthly dollars for a given year. Both of these outcomes were obtained from the Part D Event (PDE) data file. The PDE file contains a separate record each time a beneficiary made a prescription drug purchase. Outcome calculations only incorporated costs that were recorded before a beneficiary's first discontinuation from the Part D program (see Section 2.1.1 above). For each year, the sum of these costs was divided by the number of months the beneficiary was enrolled in any plan; therefore, no distinction was made between costs accrued while enrolled in different

plans for those beneficiaries who switched plans mid-year. For example, individuals switching plans in April may have paid more from January to March compared with April to December, but the calculated outcomes would reflect the average spending for any month.

3.0 RESULTS

3.1 TOTAL SAMPLE CHARACTERISTICS

Characteristics of the total sample are displayed in Table 4. In general terms, the sample was majority female and non-Hispanic white, and enrolled exclusively in PDP plans. On average beneficiaries were about 72 years, with approximately 2 CCW conditions. Although a smaller proportion of the sample was enrolled in a Part D plan during the early years, the average number of years followed in the sample was about 5.2 years out of 7. The analytical sample was similar in all characteristics.

Table 5 shows the empirical distribution of plan switching counts across all years. While 94.8% of the sample had not experienced any plan terminations, plan consolidations affected a slight majority of the sample at least one time. Out of the total sample, less than 10% switched at least once following plan consolidation, but more than 20% switched following plan renewal at least once. Overall, 35.9% of the sample had ever switched at least once during the Dec-Jan transition period, 15.2% had ever switched outside the Dec-Jan transition period, and 45% had ever switched plans at least once. About 19.6% of the sample had switched plans two or more times during the study period.

Figure 3 displays the proportion of beneficiaries in each switching category, by Dec-Jan transition period. In each period the largest proportion of the sample was PR-NS. The 2007-2008

period saw the largest proportion of PR-S, while the 2010-2011 period saw the largest proportions of PT-S, PC-NS, and PC-S among the periods observed.

Tables 6 and 7 show descriptive statistics for gross spending and patient OOP, respectively. The proportion of zero values was substantial for each outcome and increased with calendar year. Gross cost means increased from 2006-2009 and then decreased, while patient OOP costs appeared relatively consistent from 2006 to 2010, after which they decreased markedly. Consistent with a right skewed distribution, means exceeded medians for each year, and maxima were several orders of magnitude larger than medians. Furthermore, the 99th percentiles in some cases were at least an order of magnitude lower than the maximum for each year. In general, gross spending was 2-3 times larger than patient OOP costs on average.

3.2 PLAN SWITCHING AND GROSS COST

Figure 4 displays the histograms of untransformed and transformed gross cost. In its original form, the distribution of gross cost is highly skewed, with a large peak at 0 and a very long right tail (panel A). Although the distribution is much closer to normal after log transforming gross cost (after adding 1 unit, panel B), the log transformation of just the non-zero values used in the lognormal component of the mixture model produces a distribution even closer to normal (panel C).

Due to the relatively normalizing effect of the log transformation on gross cost, modelling with LMM proceeded using a log transformation (after adding 1 unit), and included all covariates. Model diagnostics in Figure 5 show that estimated random effects are approximately normal except at the tails, with the lower tail leading to largest violations of the

normality assumption (panels A1 and A2). Residuals were highly heteroscedastic, again corresponding with lower values of gross cost (panel B2). However, a histogram of the residuals suggests approximate normality.

Model convergence was achieved with the mixture model, but not when baseline gross cost was included as a covariate. Diagnostic plots of the model fit after removing baseline gross cost are shown in Figure 6. Compared to random effects from the LMM, the Q-Q plot of the random effects from the lognormal component of the mixture model do not exhibit any regions of egregious departure from the normal line, and the histogram also shows a closer fit to the normal curve (panels B1 and B2). Residual plots show an even distribution around 0 without the structure observed in the LMM model, and normality appears not to be violated (panels C1 and C2). However, the random effects estimated from the logistic component display a bimodal distribution with marked departures from normality (panel A1 and A2).

Table 8 shows unadjusted means of gross cost increasing after the enrollment period for all switching groups. Estimated coefficients, 95% confidence intervals and p-values from both the mixture models and LMM are displayed in Tables 9 and 10, respectively, while and a schematic summary of significant differences is shown in Figure 7. Plan switching variables added significantly to the model fit overall (not shown) when using the LMM structure ($p<0.001$) as well as the mixture model (logistic: $p=0.011$; lognormal: $p=0.002$). In general, the lognormal component estimated the highest gross costs in the PR-NS group, and the lowest costs in the PC-S group. This was substantiated by significant pairwise differences showing gross

costs among PR-NS were 6.6% higher compared to PC-S and 2.2% higher compared to PC-NS [1]. In addition, the lowest cost group, PC-S, had significantly lower costs compared to PC-NS and PR-S. On the other hand, results from the logistic component show that the PT-S group was associated with the highest likelihood that patient OOP costs would be greater than $0 in the year after a given enrollment period. This probability was 102.0% higher than the PC-NS group (p=0.027) and also 102.4% higher than the PC-S group (p=0.022). Finally, the PR-NS group had a 32.9% greater likelihood of positive gross cost compared to the PC-NS group (p=0.002). Estimates from the LMM model echoed those from the mixture model, with PR-NS and PC-S exhibiting the highest and lowest gross costs, respectively. PR-NS was significantly higher compared to PC-S (5.9%, p=0.007) and PC-NS (3.6%, p<0.001).

Greater number of total switching events was associated with significantly greater gross costs in the LMM model (p<.001); however the mixture model suggest that more frequent switchers are at higher risk of gross costs exceeding $0 (logistic p<.001) but not necessarily higher costs among users (lognormal p=0.8891).

3.3 PLAN SWITCHING AND PATIENT OOP COST

Histograms of the untransformed and transformed patient OOP costs are shown in Figure 8. Patient OOP costs displayed a distribution with a long right tail, with a mode somewhat larger

[1] Here and elsewhere in the Results section, reported percent differences for a given pairwise comparison represent the ratio of the two estimated parameters (reported in the corresponding table). The greater of the two estimated parameters is always divided by the smaller.

than $0 (panel A). Applying the log transformation was approximately normalizing (panel B), especially when ignoring values of 0 (panel C).

Figures 9 and 10 display diagnostic plots for the LMM and mixture models of patient OOP cost, respectively. All covariates were included in both models. LMM random effects are approximately normal except for the lower tail (panels A1 and A2). Residuals are also approximately normal, while residuals suggest marked heteroscedasticity (panels B1 and B2).

Random effects from the logistic component of the mixture model also display a departure from normality, again exacerbated at the lower tail (panels A1 and A2). However, the normal model seems to be a good fit for the lognormal random effects (panels B1 and B2) and residuals (panel C2), while residuals appear relatively homoscedastic compared to the LMM (panel C1).

Unadjusted means of patient OOP in Table 11 show slight decreases post-switching for all groups except PR-NS. Estimated coefficients and 95% confidence intervals from models of patient out of pocket spending are shown in Table 12 and 13 for mixture model and LMM, respectively. The schematic summary of significant results in Figure 7 also includes patient OOP models (lower half). Plan switching variables added significantly to the model fit overall (not shown) when using the LMM structure ($p<0.001$) as well as the mixture model (logistic: $p<0.001$; lognormal: $p<0.001$). Results from the lognormal component show the highest patient OOP costs among the PR-NS and PC-NS groups, and the lowest patient OOP costs among the PR-S and PC-S groups. As shown by the mixture model (lognormal component), the PR-NS group had 7.5% higher costs in the year following a given Dec-Jan transition period compared to PR-S ($p<0.001$), and 4.8% higher costs compared to PC-S ($p<0.001$); PC-NS also had 6.2% higher costs compared to PR-S ($p=0.020$). Results from the logistic model component indicated

that the PT-S group was associated with the highest likelihood that patient OOP costs would be greater than 0$ following a given Dec-Jan transition period. This probability was 77.9% higher than the PC-NS group ($p=0.027$) and 62.9% higher than the PC-S group ($p=0.022$). PR-NS also had a 19.7% higher probability of costs being greater than $0 compared to the PC-NS group.

Results from the LMM corroborated results from the mixture model, suggesting that in general, the PR-NS group incurred the greatest patient OOP costs, and the PR-S group incurred the lowest: costs among the PR-NS group were 8.2% higher than PC-S ($p<0.001$), 5.1% higher than PR-S ($p=0.016$), and 2.8% higher than PC-NS ($p=0.003$). On the other hand, PR-S costs were and 5.2% lower than PC-NS ($p<0.001$).

Finally, more switching events was associated with a higher likelihood of patient OOP greater than $0 in the logistic portion of the mixture model ($p<0.001$), but the estimate was not significant in the lognormal component ($p=0.059$). The LMM estimated a significantly higher patient OOP cost as total number of switches increased ($p<0.001$).

4.0 DISCUSSION

Every year during the open enrollment period, Medicare Part D beneficiaries are faced with choosing among a large number of prescription drug plans with a complex array of plan attributes. Although remaining in the same plan as the previous year is a possibility if it is renewed by the sponsor, beneficiaries may also be subject to unanticipated changes that might affect the quality of their decisions. It is also not clear to what extent differences in plan expenditures depend on switching plans. The current study presents results from a longitudinal analysis of non-LIS Part D enrollees observed between 2006 and 2012 and who participated in at least one Dec-Jan transition period.

Using random effects models to control for between-subject variability, we sought to identify whether unanticipated plan changes led to higher expenditures, and whether switching led to lower expenditure outcomes in the year directly following the Dec-Jan transition period. Model diagnostics indicated that mixture models in general provided a better fit for the data compared to LMM, but estimates still largely concurred. As shown in the lower half of Figure 7, results indicate that patient OOP costs increased the most among non-switchers (PC-NS and PR-NS groups), but only by a margin of up to 7.5% compared to other groups. Plan terminations did not lead to significantly higher spending among users, but did increase the likelihood of non-zero expenditures by up to 77.9% compared to other switching groups. Finally, plan consolidations did not lead to significantly different patient OOP compared to plan renewals, although the PC-S

group led to lower gross total spending compared to other plan consolidation and renewal groups, including PR-S.

Our results showing significantly higher patient OOP costs among non-switchers is consistent with reports of average increases in costs since 2006—especially premiums—among plans with high enrollment.[14] Individuals who do not re-evaluate the appropriateness of their current plan based on the information provided to them by CMS may be more likely to remain in their current plans and incur higher OOP costs compared to those who switch. This result is also consistent with previous reports suggesting that reductions in overspending may be attributable to switching plans.[18,28] However, our results indicate that patient OOP spending is only between 4.8 and 7.5% higher among non-switchers compared to switchers. Because mean patient OOP spending was about $45 to $59 in a given year, this translates into savings of roughly $2 to $5 per month (or $24 to $60 per year) in savings among switchers compared to non-switchers. This is almost an order of magnitude lower than the potential savings among non-LIS beneficiaries who choose the best plans, which is reported to be between $200 and $368 on average annually.[21,18] Our results indicate that while some individuals may be cutting their over spending by switching more than others, the average beneficiary was saving 6- 8 times less than they could have been. This reinforces the notion that the Medicare population is not choosing the lowest cost plans, even among individuals who find reason to actively re-evaluate their plan enrollment by switching.

Although plan terminations and consolidations may have been unanticipated, and in the case of terminations may have led to a "forced" switch, contrary to our hypothesis these events did not have a detectable negative effect on gross spending or patient OOP costs. While a possible explanation may have been a relatively low frequency of occurrence of this type of

switching, raw means of patient OOP spending pre- and post-PT-S do not suggest any large differences. Results showing higher probability of non-zero spending among the PT-S group may be driven by switching events during the 2010-2011 enrollment period. This period featured the largest proportion of plan terminations among the years considered, at 3.4%, and also corresponds with the 3rd year since CMS began issuing plan ratings and requiring that plans with poor ratings cease to be offered. It is possible that plans with poor ratings did not provide beneficiaries with needed services or coverage, and that following termination, spending may have increased due to new plans facilitating needed prescription drug use.

Our descriptive analysis of the study sample showed that a considerable number of individuals experienced plan consolidations or splitting each year, especially during the 2009-2010 and 2010-2011 enrollment periods. This coincides with the introduction of the CMS policy requiring sponsors to eliminate redundancy and unnecessary plan choices. Interestingly, patient OOP also showed relatively large decreases after 2010, perhaps an indirect effect of the same policy. Other firm mergers and acquisitions may have also contributed to the disproportionately high number of consolidations during these two periods.[20] Our results indicate that plan consolidations, despite being potentially unanticipated, did not lead to detectably poorer plan choices by way of higher OOP spending. Interestingly, however, the largest decreases in gross spending corresponded with the PC-S group, among which gross spending was 5% lower than PR-S for a given year. Future studies may find reason to investigate this finding more closely using models adjusting for baseline levels of gross spending, which we could not do using the chosen modelling method.

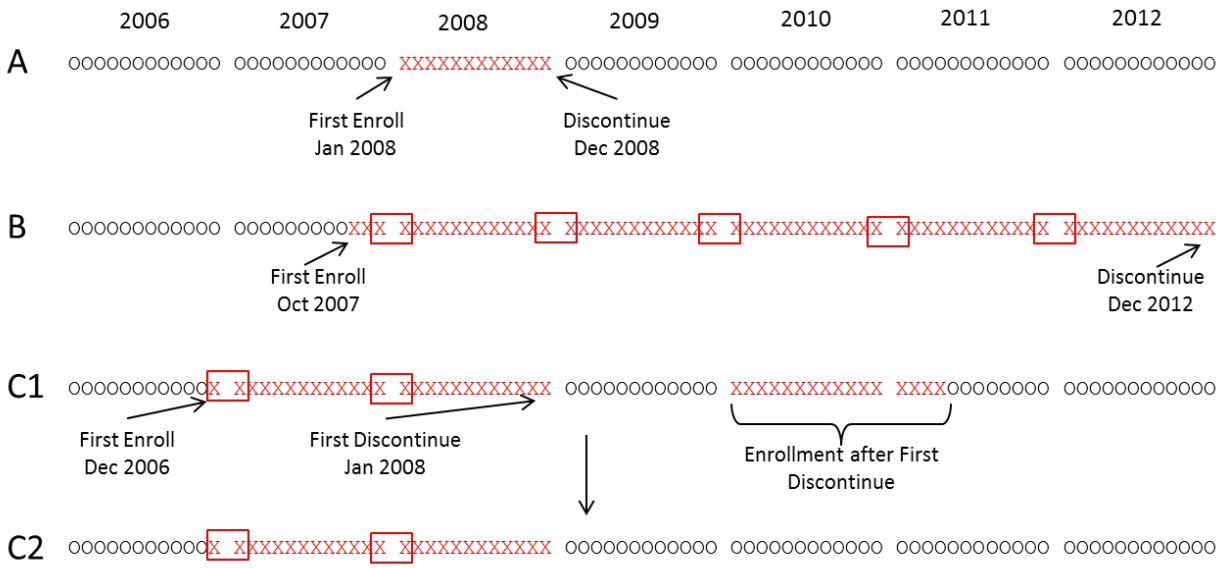
The current study as it was carried out has some notable strengths and weaknesses. The use of mixture models to address the zero-inflated expenditure data seems to have provided a

much better fit compared to LMM, at least in the lognormal components, facilitating confident interpretation of the estimates. Random effects from the logistic components, however, have bimodal distributions, making the normality assumption dubious. Future iterations of this analysis would do well to consider alternative model specifications. Models of gross spending for the current study also failed to converge when baseline spending was included as a covariate, or when geographic location was accounted for using the 34 PDP or 26 MAPD regions. Furthermore, the timeframe studied largely encompasses the 2008 economic depression as well as the introduction of specific CMS policies such as the requirement to limit duplicate plans in 2011. Therefore these results may not be fully generalizable to years beyond 2012.

In summary, the current study reports results from random effects mixture models used to evaluate the effects of plan switching types on Part D gross spending and patient OOP costs, and controlling for time and other covariates. Our results suggest that plan consolidation and plan termination do not appear to be leading to increased beneficiary spending or a resulting negative impact on public health in general. Therefore we have not found evidence to suggest that the policies which enable consolidations and terminations as they are implemented now are in need of revision. Our results also indicate that while switching plans does lead to lower OOP costs, the relative effects of switching are low compared to the potential savings reported by other studies. The public health significance of this study is that individuals incur greater costs if they do not switch plans, and may increase the chances of health-related consequences due to higher costs.[6,7,8,9,10,11] However, even among those individuals who do switch plans, the average OOP cost reduction is only a fraction of what it could be according to previous studies.[21,28] The CMS should perhaps implement more policies or tools for beneficiaries to more easily identify

compatible, minimally costly plans during open enrollment period, for example using an intelligent-decision algorithm such as that suggested by Zhang et al. [30]

APPENDIX: TABLES AND FIGURES



Each line represents a hypothetical beneficiary's status in the Part D program through the 7 years of the study, where each of the 82 months are represented by either an "O" or "X" character. The "O" characters signify that the beneficiary was not enrolled in a Part D plan that month, while the "X" characters indicate enrollment. Red boxes correspond to Dec-Jan transition periods at which plan switching group could be assigned.

Figure 1. Diagram of longitudinal Part D plan enrollment data for 3 hypothetical beneficiaries.

A. Example crosswalk file linking 2007 to 2008 plans

Plan ID Year 2007	Relationship Code	Plan ID Year 2008
Humana_A	Renew	Humana_A
Humana_B	Renew	Humana_B
WellCare_A	Consolidated	Wellcare_Gold
WellCare_B	Consolidated	Wellcare_Gold
WellCare_C	Consolidated	Wellcare_Gold
Cigna_Gold	Split	Cigna_A
Cigna_Gold	Split	Cigna_B
First Health	Terminated	--

B. Determining plan switching status for the 2007-2008 transition period

1. Merge 2007 beneficiary data to plan crosswalk file based on Dec 2007 plan ID

2007 Beneficiary Data				Plan Crosswalk File		
Bene_ID	Plan ID Jan 2007	...	Plan ID Dec 2007	Plan ID Year 2007	Relationship Code	Plan ID Year 2008
Bene_1	Humana_A	...	Humana_A	Humana_A	Renew	Humana_A
Bene_2	WellCare_C	...	WellCare_C	WellCare_C	Consolidated	Wellcare_Gold
Bene_3	WellCare_B	...	WellCare_A	WellCare_A	Consolidated	Wellcare_Gold

2. Merge resulting file to 2008 beneficiary data based on beneficiary ID

2007 Beneficiary Data				Plan Crosswalk File			2008 Beneficiary Data		
Bene_ID	Plan ID Jan 2007	...	Plan ID Dec 2007	Plan ID Year 2007	Relationship Code	Plan ID Year 2008	Plan ID Jan 2008	...	Plan ID Dec 2008
Bene_1	Humana_A	...	Humana_A	Humana_A	Renew	Humana_A	Humana_A	...	Humana_B
Bene_2	WellCare_C	...	WellCare_C	WellCare_C	Consolidated	Wellcare_Gold	Wellcare_Gold	...	Wellcare_Gold
Bene_3	WellCare_B	...	WellCare_A	WellCare_A	Consolidated	Wellcare_Gold	Humana_B	...	Humana_B

3. Compare plan relationship code, plan IDs for 2008 (red box above), to obtain plan switching status

Figure 2. Example crosswalk file and schematic of the process used for determining plan switching group status.

Table 1. Sample size before and after applying exclusion criteria.

Restriction / year	N
No restriction: all individuals	
2006	2,526,596
2007	2,570,615
2008	2,622,588
2009	2,668,398
2010	2,727,742
2011	2,802,032
2012	2,891,669
Enrolled in Part D through at least one enrollment period	2,034,326
Enrolled in Part D through at least one enrollment period and never LIS	1,187,469*

*Final total sample size

Table 2. Definition of each of the 5 state groupings used to control for geographic location.

Group	States/territories included
1	NH, ME, CT, MA, RI, VT, NY, NJ, DE, DC, MD, PA, WV
2	NC, VA, GA, SC, FL, AL, TN, LA, MS
3	MI, OH, IN, KY, IL, WI, IA, MN, MP, NE, ND, SD, WY
4	CO, NM, AZ, NV, CA, ID, OR, UT, WA
5	HI, AK, Other territories

Table 3. File location of key enrollment/demographic data before and after 2010.

Data type	File location 2006-2009	File location 2010-2012
Chronic Condition Indicators	BASF	MBSF (Chronic conditions)
LIS status	Part D Denominator File	MBSF (Part D)
Part D plan ID's (monthly)	Part D Denominator File	MBSF (Part D)
Race	Part D Denominator File	MBSF (Base)
Sex and Age	MBSF	MBSF (Base)
State Buy-In status	MBSF	MBSF (Base)
State code	MBSF	MBSF (Base)

Beneficiary Annual Summary File, BASF; Master Beneficiary Summary File, MBSF; Low-Income Subsidy (LIS)

Table 4. Characteristics of the total study sample as well as analytical sample used for modelling.

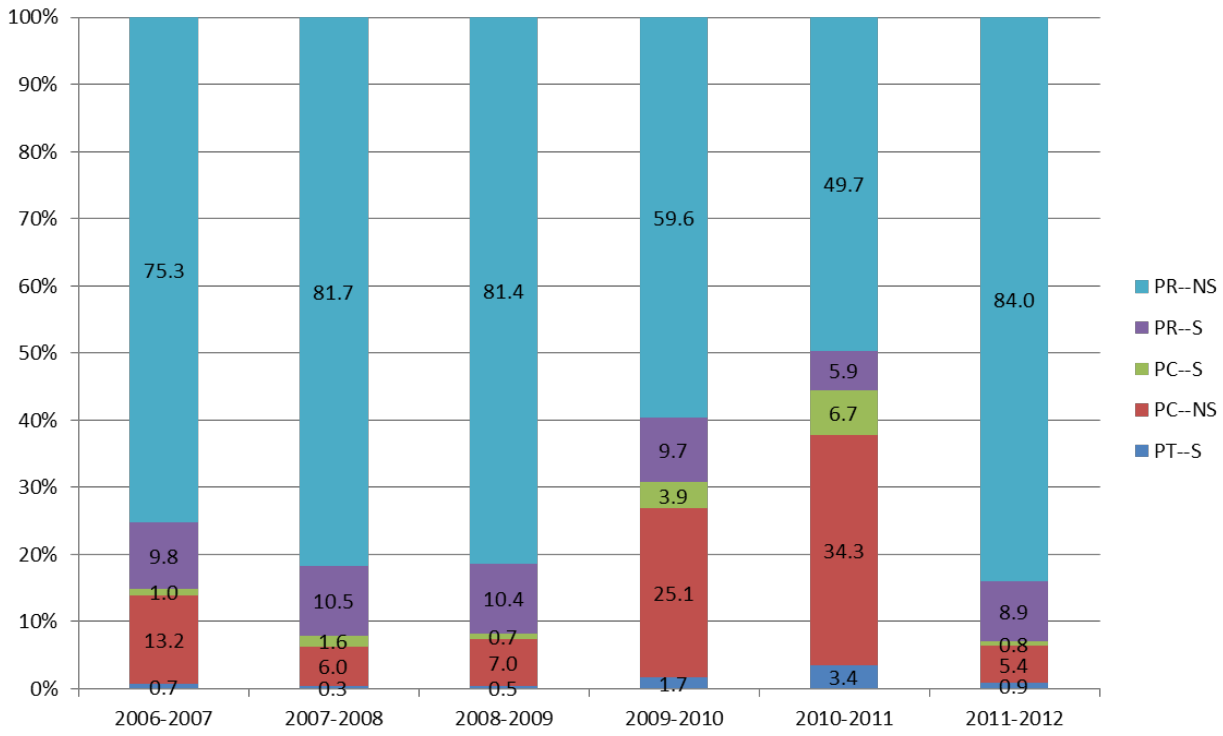
Variable (cell contents)	Total Sample N=1,187,469	Analytical Sample N=17,812
Female (%)	60.2	60.3
Race (%)		
Unknown	3.2	3.5
Non-Hispanic White	82.5	82.0
Black or African American	5.1	5.2
Other	1.2	1.3
Asian/Pacific Islander	1.8	1.7
Hispanic	6.1	6.1
American Indian/Alaska	0.2	0.2
Age at baseline (mean/SD)	72.1 (8.9)	72.2 (8.9)
# of years in sample (mean/SD)	5.2 (1.9)	5.2 (1.9)
% of total sample in each year		
2006	58.3	58.8
2007	67.0	67.3
2008	71.7	72.1
2009	75.5	75.6
2010	78.4	78.7
2011	84.0	84.3
2012	80.3	80.3
Maximum CCW conditions (mean/SD)	2.1 (2.2)	1.9 (2.1)
Plan type (%)		
PDP only	51.5	50.5
MA-PD only	34.9	34.9
Regional PPO	1.0	1.0
Employer-sponsored	0.3	0.3
Non-continuous	12.2	13.3

Standard deviation, SD; Part D Plan, PDP; Medicare Advantage Part D, MA-PD; Preferred Provider Organization; PPO

Table 5. Distribution of plan switching counts per person between 2006 and 2012, as a percent of total sample.

	Frequency per person (% of total sample)				
	0	1	2	3	>3
Plan Termination--Switch	94.8	4.8	0.3	0.0	0.0
Plan consolidation					
No Switch	52.1	33.4	12.7	1.8	0.1
Switch	90.3	8.8	0.8	0.0	0.0
Total	46.0	35.9	15.3	2.6	0.2
Plan Renewal					
No Switch	4.7	20.4	15.9	17.2	42.0
Switch	72.0	20.4	5.5	1.6	0.5
Total end-of-year switch*	64.1	23.3	8.5	2.9	0.3
Total mid-year switch**	81.3	14.8	2.8	0.7	0.4
Total switch***	54.0	26.4	11.6	4.9	3.1

*Sum of plan termination—switch, plan consolidation—switch, plan renewal—switch.
**Plan switches initiated outside of the annual enrollment period
***Sum of end-of-year and mid-year switches



Plan renewal—no switch, PR-NS; plan renewal—no switch, PR-S; plan consolidation--switch, PC-S; plan consolidation—no switch, PC-NS; plan termination--switch, PT-S.

Figure 3. Percent of beneficiaries in each switching category by transition period.

Table 6. Descriptive statistics for gross cost (average dollars per month) by year.

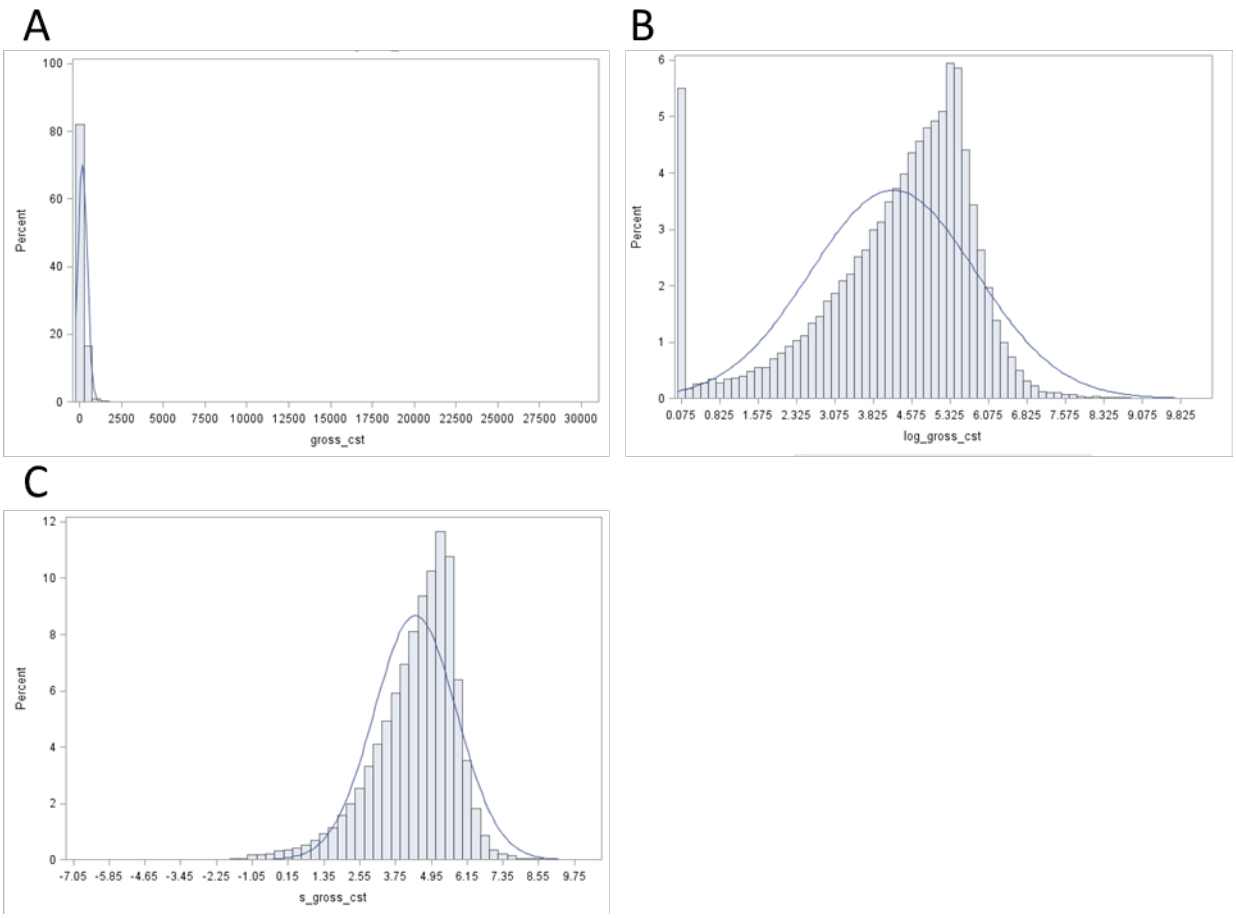
	Year						
Statistic*	2006	2007	2008	2009	2010	2011	2012
% Zero	8.3	7.3	7.9	8.6	10.4	11.3	11.1
Mean*	143.2	154.7	155.6	157.7	154.7	156.9	154.4
SD	184.7	216.8	238.3	263.3	288.3	324.0	370.3
Minimum	0.0	0.0	0.0	0.0	0.0	0.0	0.0
P25	32.5	35.8	32.7	28.7	22.7	19.2	18.2
Median	100.5	107.6	104.2	101.3	92.0	82.2	73.2
P75	197.1	207.8	209.4	215.4	211.8	208.2	194.8
P99	740.4	831.5	883.1	922.6	948.4	1023.3	1077.7
Maximum	9831.8	13476.7	16410.7	20819.2	41071.2	69385.5	53556.1

*All statistics were calculated on the total sample including zero values.
Standard deviation, SD; 25th percentile, P25; 75th percentile, P75; 99th percentile, P99

Table 7. Descriptive statistics for monthly patient OOP (average dollars per month) by year.

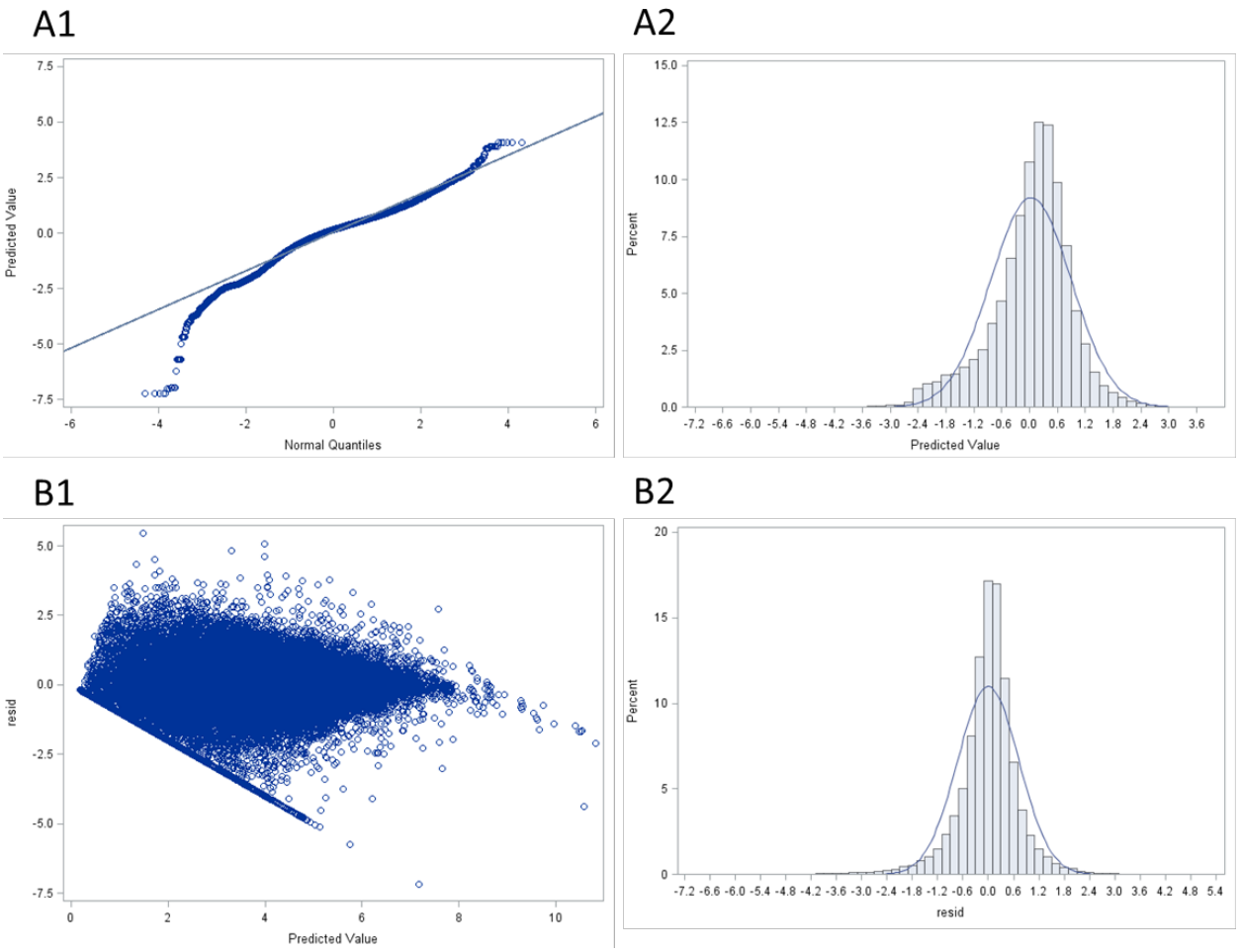
	Year						
Statistic*	2006	2007	2008	2009	2010	2011	2012
% Zero	11.0	12.6	13.6	14.6	15.4	16.7	16.1
Mean*	55.4	58.5	55.8	56.3	55.8	47.2	44.9
SD	68.8	75.6	78.2	80.7	82.5	61.0	60.8
Minimum	0.0	0.0	0.0	0.0	0.0	0.0	0.0
P25	11.3	11.9	10.0	9.3	8.5	7.4	6.8
Median	35.6	35.1	31.7	31.8	30.7	28.2	26.1
P75	68.9	69.5	64.6	65.8	65.2	61.7	57.5
P99	321.9	338.4	355.9	379.4	395.9	270.9	276.4
Maximum	2447.2	3017.3	4730.9	4922.3	4595.7	3671.2	2845.2

*All statistics were calculated on the total sample including zero values.
Standard deviation, SD; 25th percentile, P25; 75th percentile, P75; 99th percentile, P99



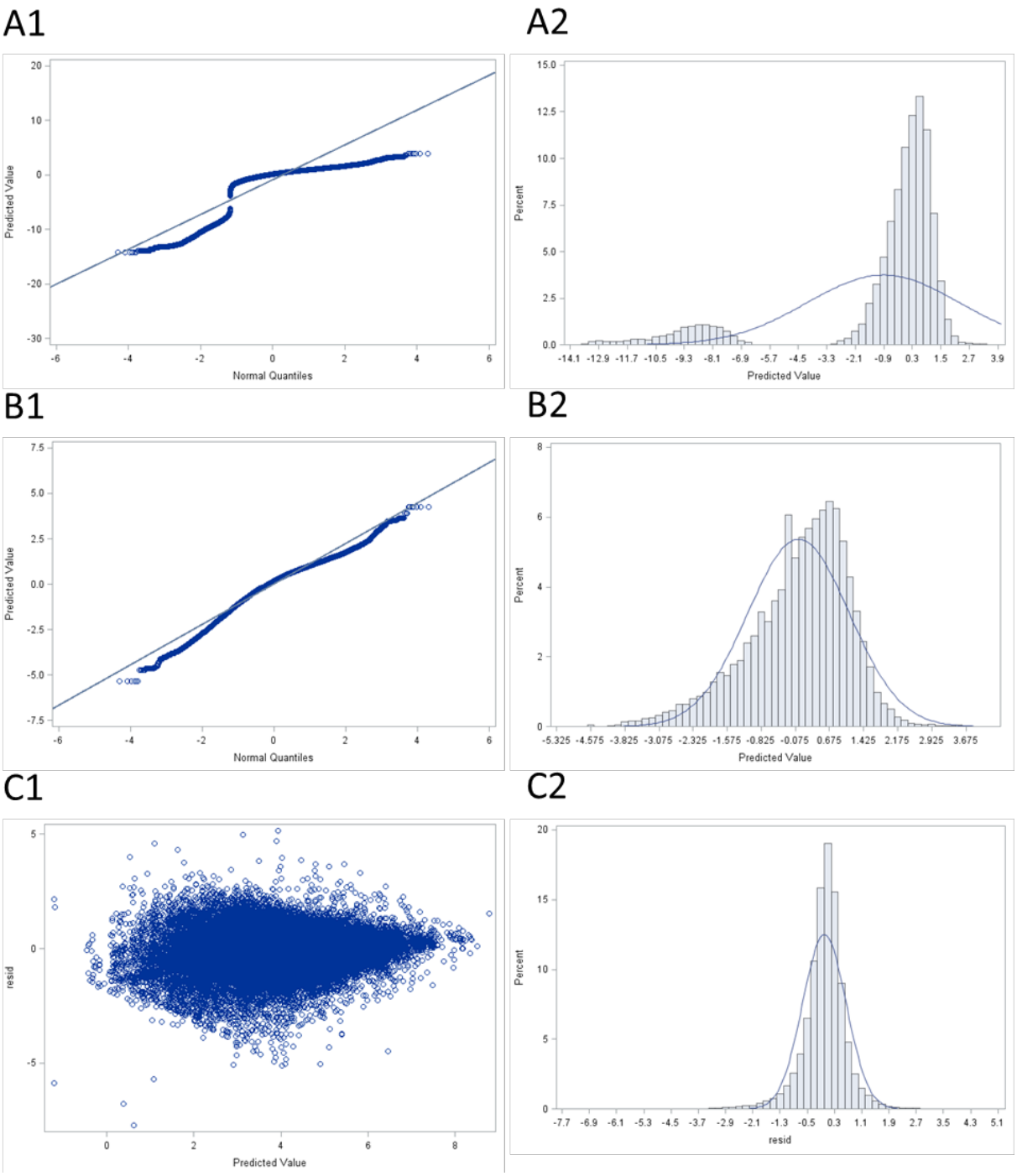
Untransformed gross cost (A); log transformed gross cost + 1 modelled in the LMM (B); and log of non-zero values of gross cost modelled by the mixture model (C).

Figure 4. Histograms with normal curves showing the distribution of untransformed and transformed gross cost.



Q-Q plot of the estimated random effects (A1); histogram of estimated random effects (A2); scatterplot of residuals vs. predicted values (B1); histogram of residuals (B2). Residuals are calculated as ln(gross cost +1) - linear predictor - random effects.

Figure 5. Diagnostic plots for LMM model of gross cost.



Q-Q plots of the estimated random effects from logistic component (A1) and lognormal component (B1); histograms of estimated random effects from logistic component (A2) and lognormal component (B2); scatterplot of residuals vs. predicted values from lognormal component (C1); histogram of residuals from lognormal component (C2). Residuals are calculated as ln(gross cost intensity variable) – linear predictor – random effects.

Figure 6. Diagnostic plots for mixture model of gross cost.

Table 8. Unadjusted means and standard deviations of gross cost (in average dollars per month) before and after switching, by switch type.

	Before		After	
Switch Type	**Mean**	**SD**	**Mean**	**SD**
PT-Switch	175.5	310.0	177.0	342.3
PC-No Switch	153.3	259.2	158.7	291.5
PC-Switch	149.8	240.5	152.8	300.9
PR-Switch	157.8	229.3	162.7	274.9
PR-No Switch	158.9	257.8	163.8	291.9

Standard deviation, SD; Plan termination, PT; plan consolidation, PC; plan renewal, PR.

Table 9. Estimated coefficients, confidence limits and p-values from mixture model of gross total cost.

Covariate (reference level)	Logistic Component			Lognormal Component		
	Estimate	Confidence Interval	P-value	Estimate	Confidence Interval	P-value
Plan Switch Type (PT-S)						
PC-NS	0.50	0.28-0.89	0.02	1.01	0.95-1.07	0.81
PC-S	0.49	0.25-0.98	0.04	0.96	0.90-1.03	0.27
PR-S	0.63	0.34-1.14	0.13	1.01	0.96-1.07	0.66
PR-NS	0.66	0.37-1.17	0.15	1.03	0.97-1.09	0.31
Total Number of Switches	1.58	1.33-1.87	<.01	1.00	0.98-1.02	0.89
Residual Variance	3.29**	--	--	0.52	0.52-0.53	<.01
Random Effect Variance	85.62	74.12-97.12	<.01	1.44	1.40-1.47	<.01
Random effect covariance ***	1.34	1.14-1.54	<.01	1.34	1.14-1.54	<.01

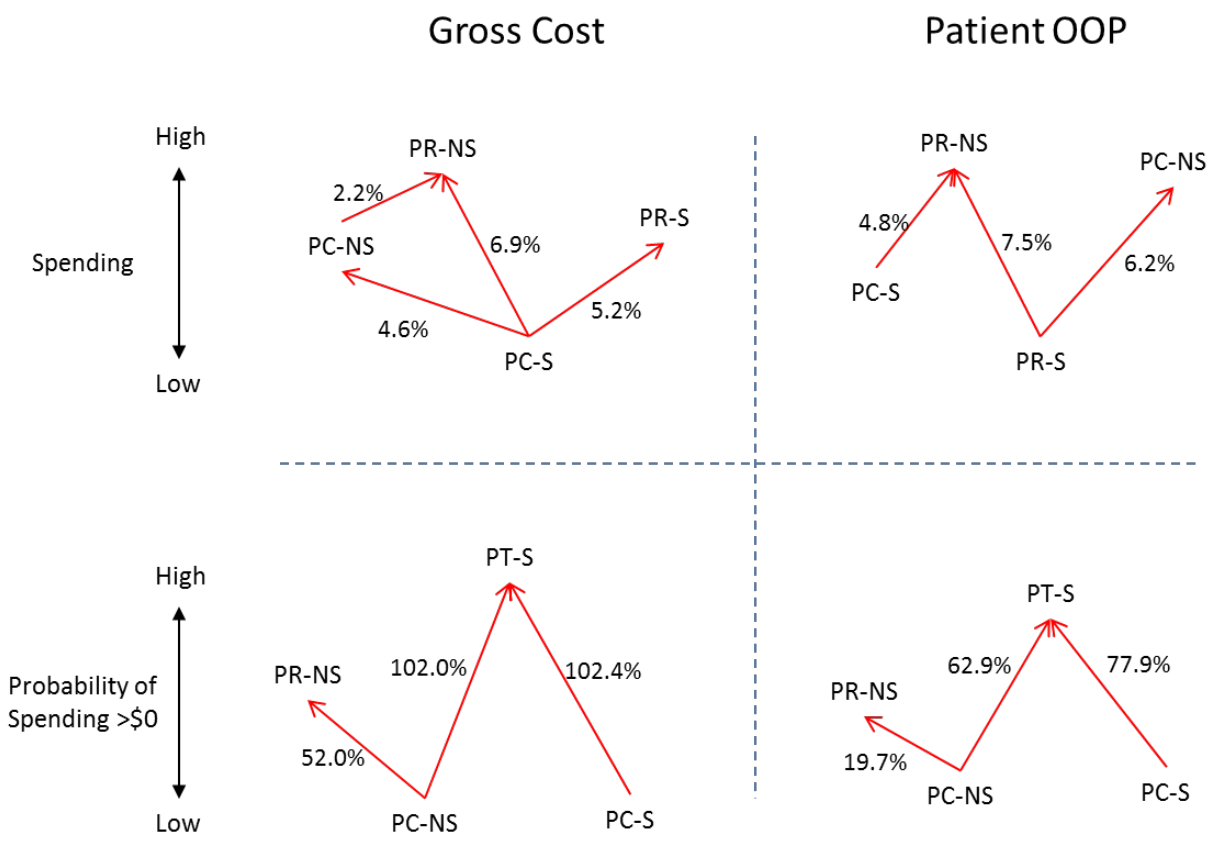
*Occurrence or intensity model

**Not estimated--residual variance from logistic models is always $\pi^2/3 = 3.290$.

***Random effect covariance is not component-specific

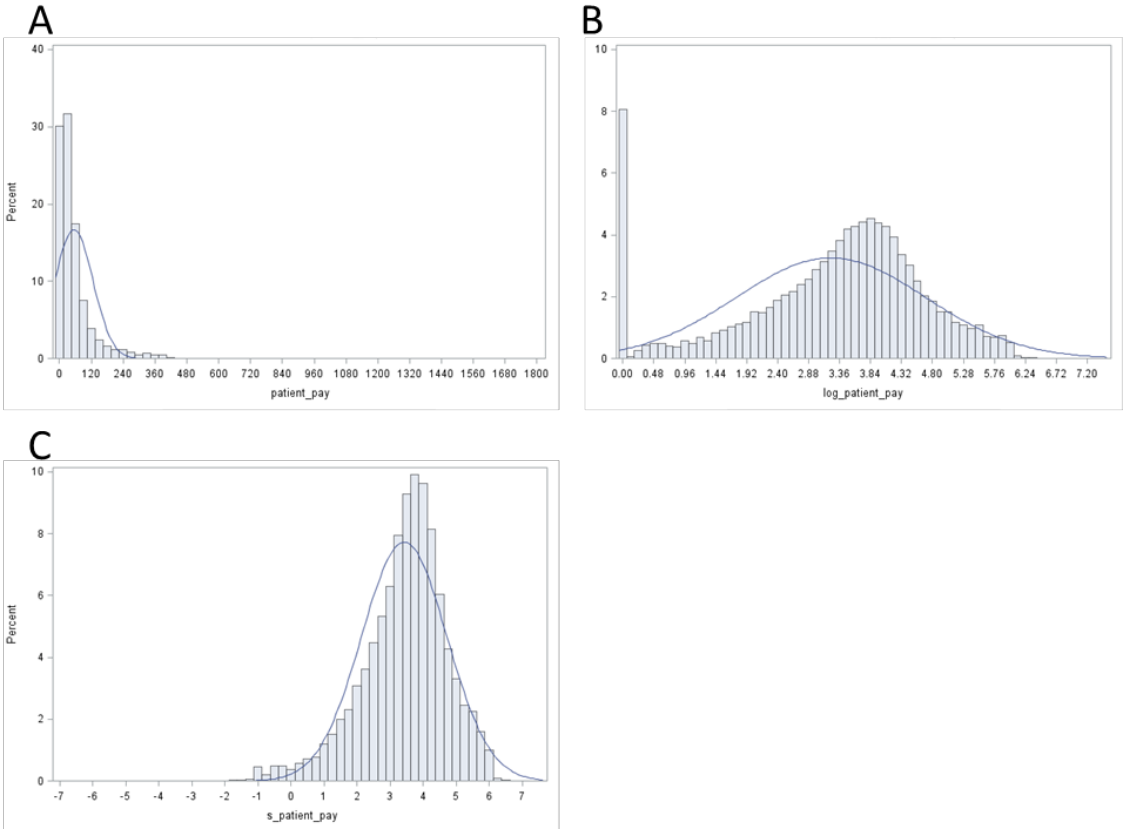
Table 10. Estimated coefficients, confidence limits and p-values from LMM of gross total cost.

Covariate (reference level)	Estimate	Confidence Interval	P-value
Plan Switch Type (PT-S)			
PC-NS	0.99	0.93-1.05	0.76
PC-S	0.97	0.90-1.04	0.37
PR-S	1.01	0.95-1.07	0.87
PR-NS	1.03	0.97-1.09	0.37
Total Number of Switches	1.06	1.04-1.08	<.01
Residual Variance	0.66	0.65-0.67	<.01
Random Effect Variance	0.94	2.49-2.61	<.01



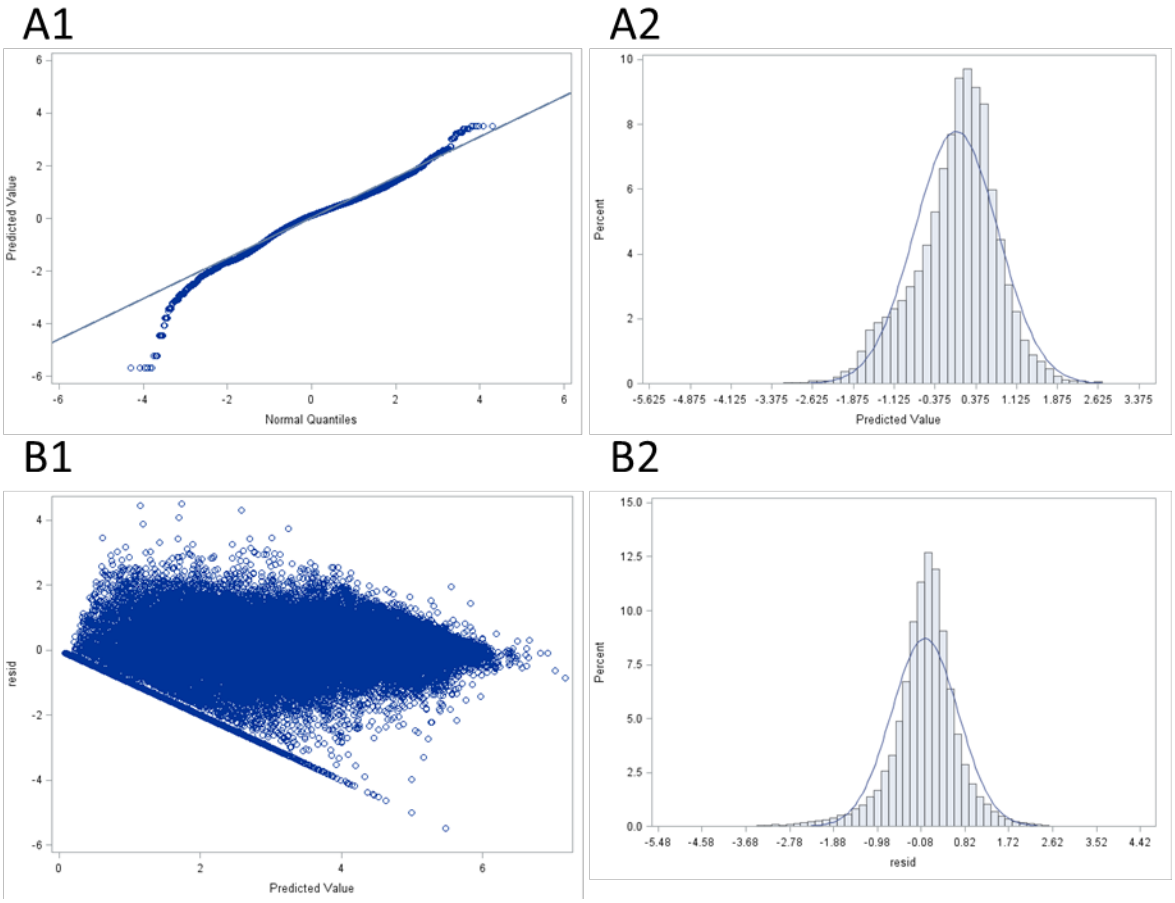
Schematic summary of significant differences from lognormal (top half) and logistic (bottom half) components of the mixture model controlling for covariates. Omitted groups are not significantly different from other groups. A red arrow signifies that the group at the head of the arrow had significantly higher spending (or probability of spending >$0) compared to the group at the arrow's tail. Arrow labels represent estimated percent differences. Out of pocket, OOP; plan termination—switch, PT-S; plan consolidation—no switch, PC-NS; plan consolidation—switch, PC-S; plan renewal—no switch, PR-NS; plan renewal—switch, PR-S.

Figure 7. Schematic summary of significant differences in spending among plan switching groups from mixture models.



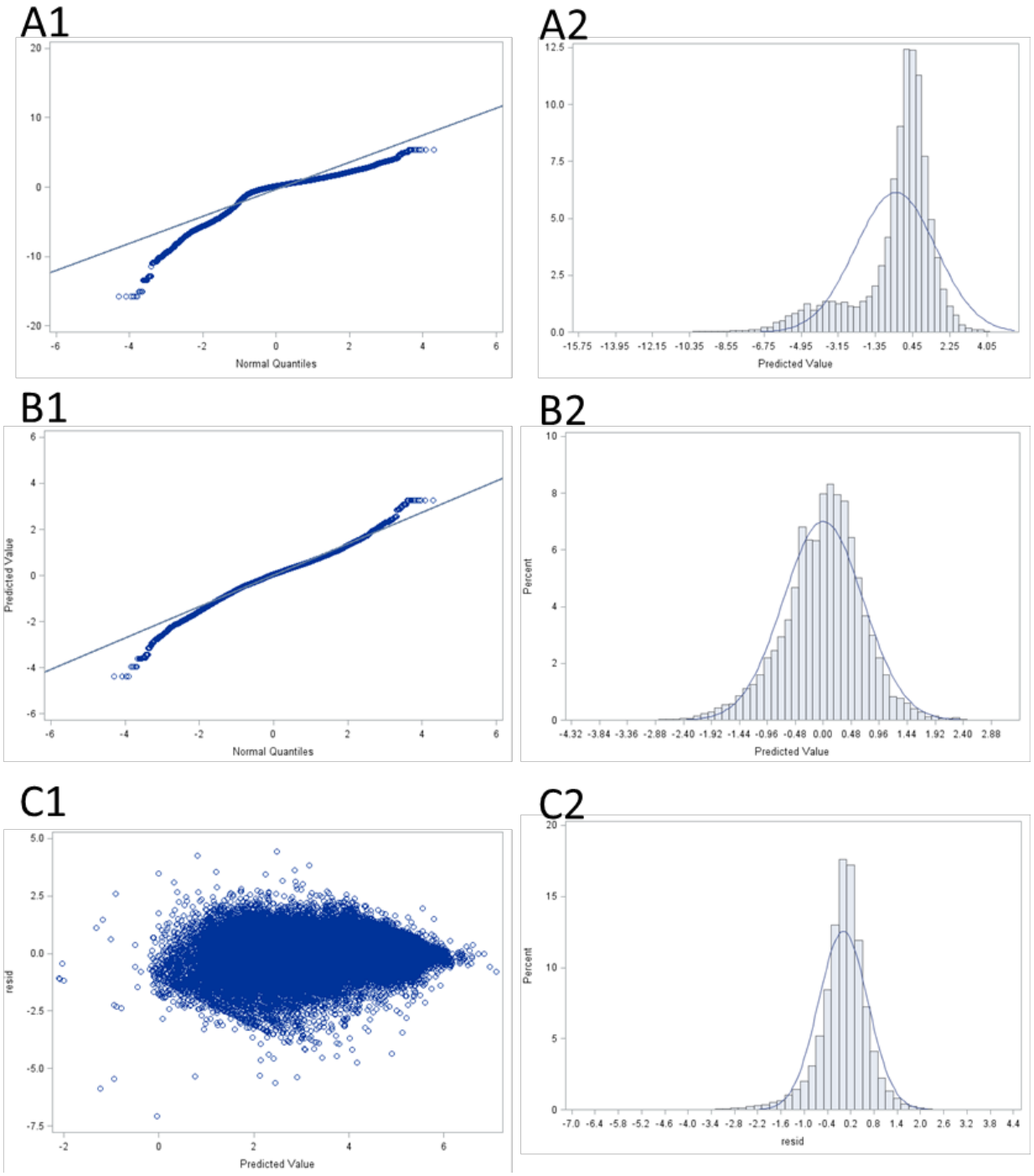
Untransformed patient OOP costs (A); log transformed patient OOP cost + 1 modelled in the LMM (B); and log transformed intensity variable derived from patient OOP costs and modelled by the mixture model (C).

Figure 8. Histograms with normal curves showing the distribution of untransformed and transformed patient OOP costs.



Q-Q plot of the estimated random effects (A1); histogram of estimated random effects (A2); scatterplot of residuals vs. predicted values (B1); histogram of residuals (B2). Residuals are calculated as ln(patient OOP cost +1) - linear predictor - random effects.

Figure 9. Diagnostic plots for LMM model of patient OOP costs.



Q-Q plots of the estimated random effects from logistic component (A1) and lognormal component (B1); histograms of estimated random effects from logistic component (A2) and lognormal component (B2); scatterplot of residuals vs. predicted values from lognormal component (C1); histogram of residuals from lognormal component (C2). Residuals are calculated as ln(patient OOP intensity variable) – linear predictor – random effects.

Figure 10. Diagnostic plots for mixture model of patient OOP costs.

Table 11. Unadjusted means and standard deviations of patient OOP cost (in average dollars per month) before and after switching, by switch type.

	Before		After	
Switch Type	**Mean**	**SD**	**Mean**	**SD**
PT-Switch	53.4	68.6	52.1	66.3
PC-No Switch	58.4	79.1	56.7	73.0
PC-Switch	58.4	81.9	54.9	76.4
PR-Switch	57.4	73.3	55.4	75.0
PR-No Switch	54.0	66.8	54.3	68.2

Standard deviation, SD; Plan termination, PT; plan consolidation, PC; plan renewal, PR.

Table 12. Estimated coefficients, confidence limits and p-values from mixture model of patient OOP cost.

Covariate (reference level)	Logistic Component			Lognormal Component		
	Estimate	Confidence Interval	P-value	Estimate	Confidence Interval	P-value
Plan Switch Type (PT-S)						
PC-NS	0.61	0.40-0.95	0.03	1.02	0.97-1.08	0.39
PC-S	0.56	0.34-0.92	0.02	0.99	0.93-1.06	0.74
PR-S	0.55	0.37-0.88	0.01	0.96	0.91-1.02	0.20
PR-NS	0.74	0.48-1.12	0.15	1.04	0.98-1.09	0.19
Total Number of Switches	1.18	1.08-1.29	<.01	1.01	1.00-1.03	0.06
Residual Variance	3.29**	--	--	0.50	0.50-0.51	<.01
Random Effect Variance	13.11	11.35-14.86	<.01	0.57	0.56-0.59	0.57
Random effect covariance ***	0.93	0.86-1.01	<.01	0.93	0.87-1.01	<.01

*Occurrence or intensity model

**Not estimated--residual variance from logistic models is always $\pi^2/3 = 3.290$.

***Random effect covariance is not component-specific

Table 13. Estimated coefficients, confidence limits and p-values from LMM of patient OOP cost.

Covariate (reference level)	Estimate	Confidence Interval	P-value
Plan Switch Type (PT-S)			
PC-NS	0.99	0.94-1.05	0.81
PC-S	0.97	0.91-1.04	0.39
PR-S	0.94	0.89-1.00	0.05
PR-NS	1.02	0.97-1.08	0.46
Total Number of Switches	1.01	1.01-1.02	<.01
Residual Variance	0.08	0.08-0.08	<.01
Random Effect Variance	0.10	0.09-0.10	<.01

BIBLIOGRAPHY

1. Medicare Prescription Drug, Improvement, and Modernization Act. Pub. L. no. 108-173 (2003).

2. Schneeweiss S, Patrick AR, Pedan A, Varasteh L, Levin R, Liu N and Shrank WH (2009) The effect of Medicare Part D coverage on drug use and cost sharing among seniors without prior drug benefits. *Health Affairs* 29:w305-w316.

3. Heiss F, McFadden D and Winter J (2006) Who failed to enroll in Medicare Part D, and why? Early Results. *Health Affairs* 29(5):w344-354.\

4. Henry J. Kaiser Family Foundation (2015) Medicare Advantage: Fact Sheet. Accessed June 2015: http://kff.org/medicare/fact-sheet/medicare-advantage/

5. Henry J. Kaiser Family Foundation (2015) The Medicare Part D Prescription Drug Benefit: Fact Sheet. Accessed January 2015: http://files.kff.org/attachment/medicare-prescription-drug-benefit-fact-sheet

6. Kaplan CM and Zhang Y (2014) The January Effect: Medication reinitiation among Medicare Part D Beneficiaries. *Health Econ.* 23(11):1287-1300.

7. West JC, Wilk JE, Rae DS, Muszynski IL, Rubio-Stipec M, Alter CL, Sanders KE, Crystal S, Regier DA (2010) First-year Medicare Part D prescription drug benefits: Medication access and continuity among dual eligible psychiatric patients. *J Clin Psychiatry* 71(4):400-410.

8. Kurlander JE, Kerr EA, Krein S, Heisler M, Piette JD (2009) Cost-related nonadherence to Medications among patients with diabetes and chronic pain. *Diabetes Care* 32: 2143-2148.

9. Zhang Y, Baik SH, Lave JR (2013) Effects of Medicare Part D coverage gap on medication adherence. *Am J Manag Care* 19(6):e214-e224.

10. Zivin K, Madden JM, Graves AJ, Zhang F, Soumerai SB (2009) Cost-related medication nonadherence among beneficiaries with depression following Medicare Part D. *Am J Geriatr Psychiatry* 17(12): 1068-1076.

11. Bakk L, Woodward TA, Dunkle RE (2014) The Medicare Part D coverage gap: Implications for non-dually eligible older adults with a mental illness. *Journal of Gerontological Social Work* 57:37-51.

12. Hoadley J and Summer L (2014) Medicare Part D in its ninth year: The 2014 marketplace and key trends 2006-2014. The Henry J. Kaiser Family Foundation. Accessed July 2015.

13. Patient Protection and Affordable Care Act. Pub. L. no. 111-148 (2010).

14. Hoadley H, Summer L, Hargrave E, Cubanski J, Neuman T (2012) Analysis of Medicare prescription drug plans in 2012 and key trends since 2006. The Henry J Kaiser Family Foundation. Accessed January 2015.

15. Hargrave E, Hoadley J, Cubanski J, Neuman T (2009) Medicare prescription drug plans in 2009 and key changes since 2006: Summary of findings. The Henry J Kaiser Family Foundation. Accessed January 2015: https://kaiserfamilyfoundation.files.wordpress.com/2013/01/7917.pdf

16. Hoadley J, Hargrave E, Summer L, Cubanski J and Neuman T. 2013. To switch or not to switch: Are Medicare beneficiaries switching drug plans to save money? Henry J. Kaiser Family Foundation. Accessed January, 2014: https://kaiserfamilyfoundation.files.wordpress.com/2013/10/8501-to-switch-or-not-to-switch1.pdf

17. Henry J. Kaiser Family Foundation (2006) Voices of beneficiaries: Early experiences with the Medicare drug benefit. Henry J. Kaiser Family Foundation. Accessed July, 2015: < https://kaiserfamilyfoundation.files.wordpress.com/2013/01/7504.pdf>

18. Ketcham JD, Lucarelli C (2015) Paying attention or paying too much in Medicare Part D. *American Economic Review* 105(1):204-233.

19. Tanius BE, Wood S, Hanoch Y (2009) Aging and choice: Applications to Medicare Part D. *Judgement and Decision Making* 4(1):92-101.

20. Hoadley J, Hargrave E, Summer L, Cubanski J, Neuman T (2013) To switch or not to switch: Are Medicare beneficiaries switching drug plans to save money? The Henry J. Kaiser Family Foundation. Accessed February 2015: https://kaiserfamilyfoundation.files.wordpress.com/2013/10/8501-to-switch-or-not-to-switch1.pdf.

21. Zhou C, Zhang Y (2012) The vast majority of Medicare Part D beneficiaries still don't choose the cheapest plans that meet their medication needs. *Health Affairs* 31(10):2259-2265.

22. Heiss F, Leive A, McFadden D, Winter J (2012) Plan selection in Medicare Part D: Evidence from administrative data. SFB/TR 15 Discussion Paper, No. 384. Accessed January 2015: http://nbn-resolving.de/urn:nbn:de:bvb:19-epub-13170-3.

23. Kearsley CH, Upchurch G, Holmes GM and Weinberger M. 2012. The other yearly checkup: Your Medicare drug plan. *J Am Pharm Assoc* 52: 634-636.

24. Medicare.org. Have you done your Yearly Medicare Plan Review? CMS Product No 11220. Accessed January 23, 2015: https://www.medicare.gov/Pubs/pdf/11220.pdf.

25. Medicare Interactive. Should I review my Medicare Part D plan choice every year? Accessed January 23, 2015: http://www.medicareinteractive.org/page2.php?topic=counselor&page=script&script_id=1497

26. Han J, Ko DW, Urmie JM (2014) The role of satisfaction and switching costs in Medicare Part D choices. *Res Social Adm Pharm* 10(2):398:407.

27. Abaluck J and Gruber J (2011) Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program. *Am Econ Rev* 101(4):1180-1210.

28. Ketcham JD, Lucarelli C, Miravete EJ, Roebuck MC (2012) Sinking, swimming, or learning to swim in Medicare Part D. *American Economic Review* 102(6):2639-2673.

29. Bishop CE, Gilden D, Thomas CP, Kubisiak J (2010) Enrollment of dually eligible beneficiaries in Medicare Part D plans: Autoassignment and choice. Centers for Medicare and Medicaid Services. Accessed January 2015: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/downloads/Bishop_DualsChoice_2010.pdf

30. Zhang Y, Zhou C, Baik SH (2014) A simple change to the Medicare Part D low-income subsidy program could save $5 billion. *Health Affairs* 33(6):940-945.

31. Centers for Medicare & Medicaid Services (2015) CMS releases prescriber-level Medicare data for first time. Centers for Medicare & Medicaid Services. Accessed July 2015: http://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2015-Fact-sheets-items/2015-04-30.html

32. Boos DD, Stefanski LA. Essential Statistical Inference. New York: Springer; 2013: 58-59.

33. Tuerlinckx F, Rijmen F, Verbeke G, De Boeck P (2006) Statistical inference in generalized linear mixed models: A review. *British Journal of Mathematical and Statistical Psychology* 59(2):225-255.

34. Gardiner JC, Luo Z, Roman LA (2009) Fixed effects, random effects and GEE: What are the differences? *Statistics in Medicine* 28:221-239.

35. Yu WW (2006) An approximation of skewed healthcare expenditure distribution using a mixture model. American Statistical Association Section on Survey Research Methods, Online Proceedings.

36. Mihaylova B, Briggs A, O'Hagan A, Thompson SG (2011) Review of statistical methods for analyzing healthcare resources and costs. *Health Economics* 20:897-916.

37. Duan N, Manning Jr WG, Morris CN, Newhouse JP (1984) Choosing between the sample-selection model and the multi-part model. *Journal of Business and Economic Statistics* 1(2):115-126.

38. Tooze JA, Grunwald GK, Jones RH (2002) Analysis of repeated measures data with clumping at zero. *Statistical Methods in Medical Research* 11:341-355.