An Economic Analysis of Secure Messaging between Patients and Providers

Rodica Gilles

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2012

Reading Committee:
Seik Kim, Chair
Richard C. Hartman
Gregory M. Ellis
Douglas A. Conrad

Program Authorized to Offer Degree:
Economics

University of Washington

Abstract

This dissertation presents an economic analysis of patient-provider communication through secure messaging. Section 1 presents the evolution of secure messaging. Section 2 aims to determine whether secure messages can substitute for face-to-face office visits. To answer this question, two demand equations were estimated and outpatient copayment was used as a key factor. A quasi-conditional maximum likelihood estimator was used as it is robust when there are distributional misspecifications and accounts for individual unobserved heterogeneity. Controlling for individual characteristics such as age, gender, insurance source, health status, and individual unobserved heterogeneity (fixed effects), we found that an increase by $1 in the price of office visits, measured by outpatient copayment, leads to 0.45% fewer office visits and 4.4% more secure messages. This is a strong indication that secure messages and office visits are substitutes. Section 3 aims to determine whether secure messaging reduces provider costs. To answer this question, a cost minimizing model is estimated. An instrumental variable approach was used to correct for the potential correlation of health and secure messaging user measures with the error term. A potential problem of unobserved individual heterogeneity was also taken into account. While we did not find evidence that secure messaging users impact the primary, specialty, inpatient, and emergency care costs, we did find evidence that individual costs on drugs are reduced by 66.4% per quarter. A lack of evidence on the impact of secure messaging on most costs may be due to limitations in the way costs are measured. Finally, to overcome a limitation of the quasi estimator used in Section 2, Section 4 seeks to introduce a new estimator ZI QCMLE aimed at estimating count and continuous models with an excess number of zeros and unobserved individual heterogeneity. This approach is similar to a zero inflated model in the way it deals with excess zeros, but unlike a zero inflated model it is robust to distributional misspecifications. Also, its ability to deal with fixed effects makes this estimator unique.

An Economic Analysis of Secure Messaging between Patients and Providers

Rodica Gilles

Chair of the Supervisory Committee:

Assistant Professor Seik Kim

Economics

TABLE OF CONTENTS

List of Figures

Figure Number Page

List of Tables

Section 1: Defining Messaging between Patients and Providers

Patient-physician messaging is generally considered an application of Telemedicine. Telemedicine has been broadly defined as the use of telecommunication technologies to provide medical information and services to remote populations (Perednia and Allen 1995, Dávalos, et al. 2009). Although online encounters have only existed since the late 1990s, forays into Telemedicine using interactive TV started in the 1960s and 1970s with the first widespread deployments occurring in Norway in the early 1990s. Some of the earliest applications of Telemedicine included teledermatology, telepathology, and teleradiology (Edwards 2011).

Telemedicine can be evaluated in a three-dimensional model (Bashshur, Shannon and Sapci 2005). The three dimensions are: (1) technology, (2) application, and (3) perspective (Figure 1). The next three subsections describe the way these dimensions can be applied to patient physician messaging.

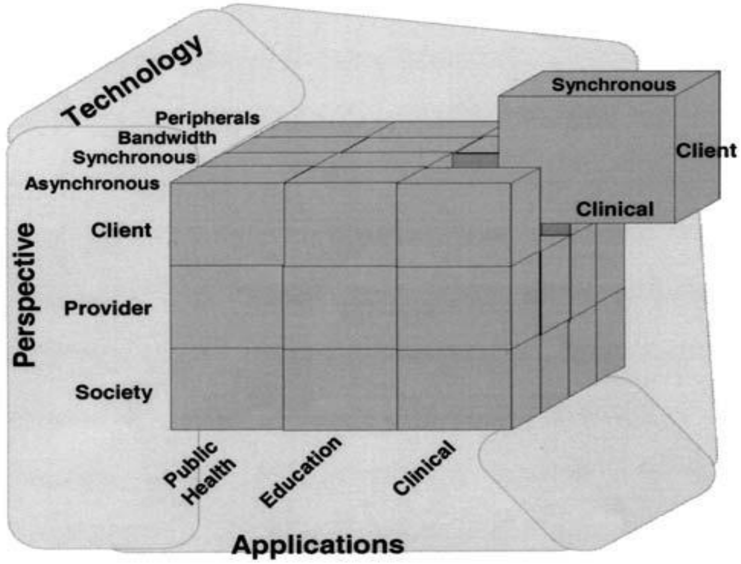

Figure 1. Three Dimensional Framework of Analysis

Dimension 1: Technology

Patient-provider messaging technologies include email and secure messaging through a web portal. The key characteristic of such technologies is that they are asynchronous (also called store and forward messaging or non-real-time messaging). They are in contrast to synchronous technologies that include the analog telephone, voice over IP (VOIP), chat, and video. Both asynchronous and synchronous technologies may be called "online" and "e" technologies because they both rely on the internet. However, they are very different because synchronous technologies require patients and physicians to interact at a common point in time, while asynchronous technologies don't.

Email Messaging

Email was the early choice for patient-provider messaging. This option is relatively easy to implement with few startup costs and is generally a familiar technology for patients. Figure 2 provides the conceptual overview of the first generation patient physician messaging system. The model is simple. The patient sends an email to their provider's email address. Then, the provider reads the email, clicks reply, writes a response, and sends the response back to the patient. Upon receiving the email, the patient may either end the process, or may send a follow-up email.

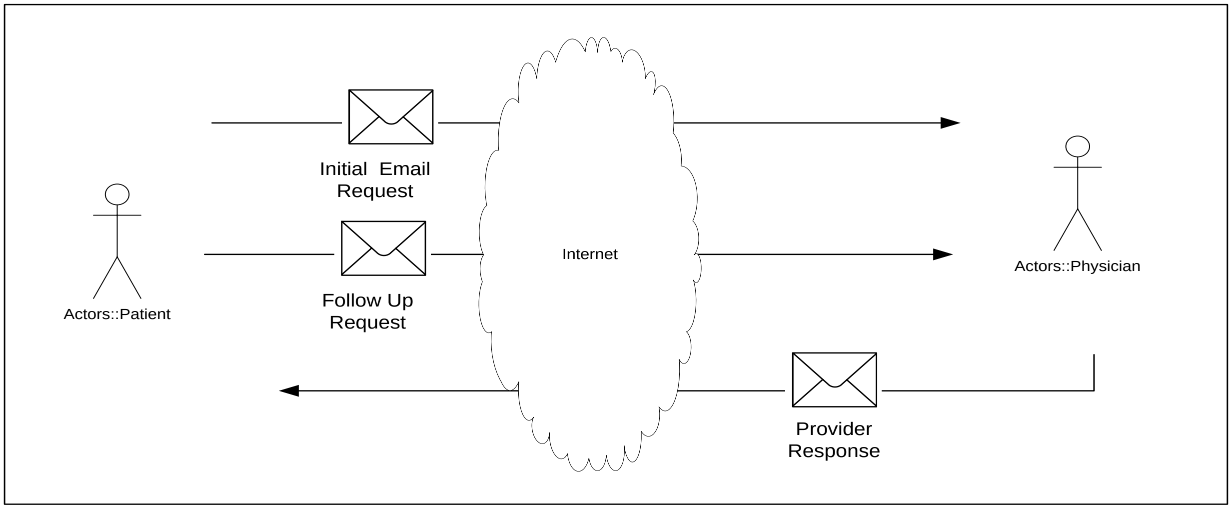

Figure 2. First Generation Patient Provider Messaging

A refinement of this option has been to triage all incoming emails by routing them to a single person who is then responsible for forwarding them to the appropriate recipient, such as the billing clerk for insurance questions or the physician for medical issues. The biggest concern with email is that unless it is encrypted it is often not compliant with the Health Insurance Portability and Accountability Act (HIPAA) privacy and security regulations. However, email is still used by many providers today.

<u>Secure Messaging</u>

Secure messaging (also known as web messaging) has been replacing email as the preferred technology since around the year 2000. Secure messaging generally involves both providers and patients using the same secure website (Figure 3). First, a unique account for the site is given to patients. After they logon, they can generally choose from different functional options which may prompt them for more information. When they are done, a request is stored in the system and someone at the Care Delivery Organization (CDO) is prompted to respond.

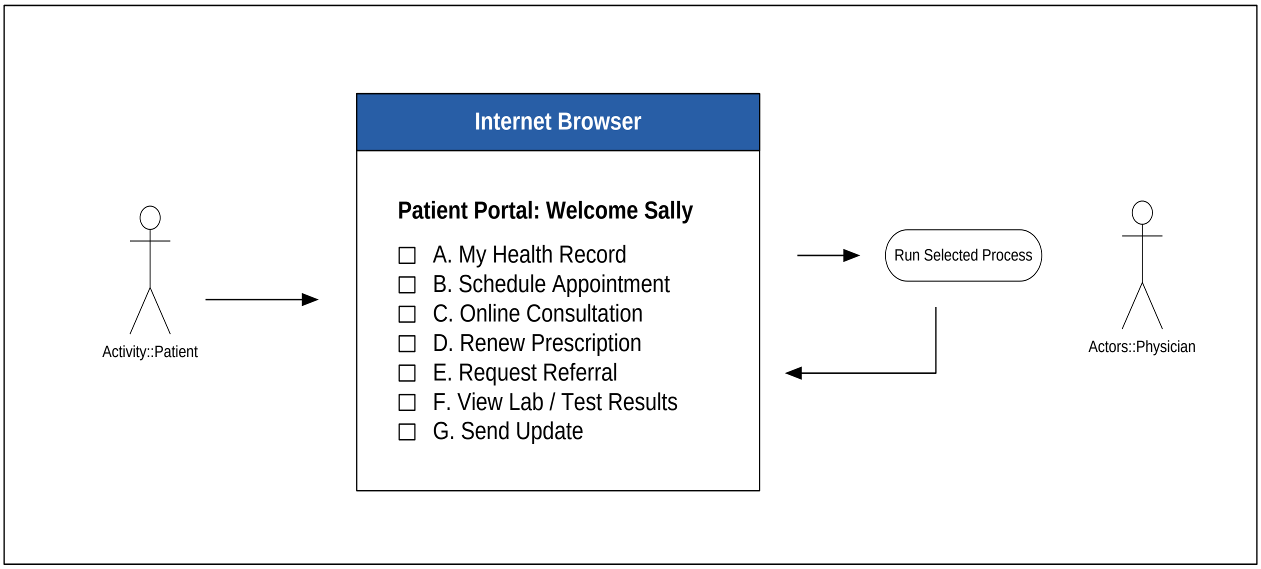

Figure 3. Second Generation Patient Provider Messaging

Secure messaging is generally encrypted and therefore doesn't have the same security and privacy issues as email. In addition, since users must have an authenticated and authorized account, physicians don't have to worry about getting unwanted spam. The implementation costs may be higher than email, but secure messaging systems often offer more functionality and more efficiency.

Furthermore, secure messaging systems generally present the user with multiple functional options from which to choose. For example, there might be an option for requesting an appointment that includes fields for entering availability and preferred times. Or there might be an option for initiating a new evaluation and management (E/M) visit that includes prompts for background history and case history.

Secure Messaging through Patient Portals

Providers who have chosen to use a web-based solution must decide whether to use an in-house or an outsourced solution. Although in many cases CDOs may choose to purchases a single solution that meets all of their needs, in other cases CDOs may make separate implementation decisions for different sets of capabilities.

An outsourced patient portal is a portal that is hosted by a third party. Patient portals generally include patient-provider secure messaging, patient education materials, and access to an electronic health record. Instead of the CDO installing the software in house, it is often installed and managed for them by the software vendor themselves. Examples range from Yahoo and Google email to healthcare specific applications including practice management, electronic medical records, and patient portal applications. For example, MedFusion provides a hosted patient portal solution that is used by the Mayo Clinic which has a proprietary electronic medical record (EMR) solution.

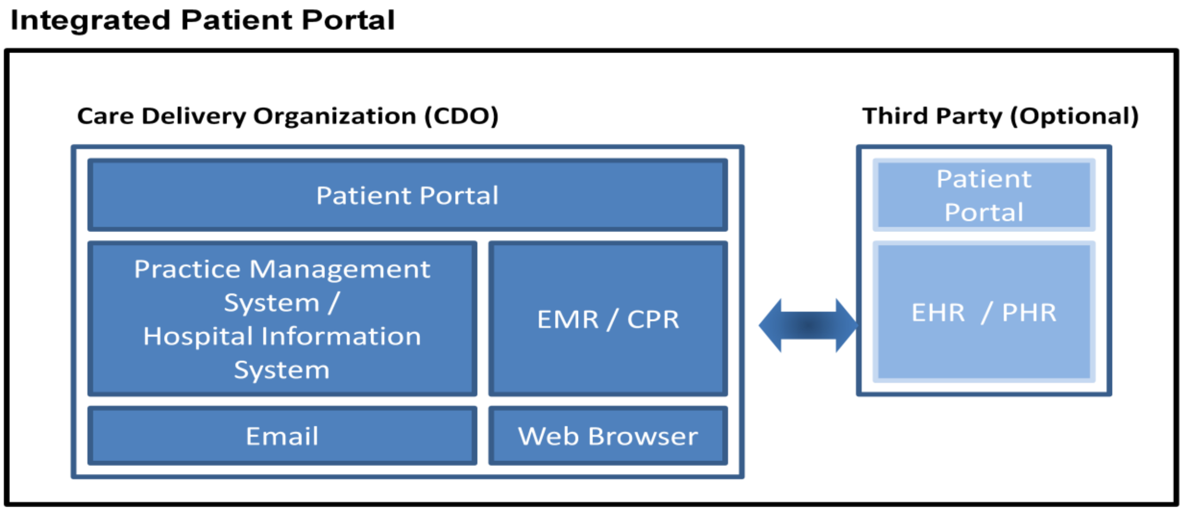

Figure 4. Patient Portal Integration

There are a number of software vendors that provide outsourced patient portal capabilities. The outsourced model generally requires some upfront setup fee, and then an ongoing monthly fee for usage. It may also require a per transaction fee. Outsourcing the patient portal is generally the easiest way for physicians to add secure messaging, including messaging-based E/M visits, ad hoc question and response, electronic prescribing, and similar capabilities. They can often start with limited or no integration with their in house management and electronic medical records system and still gain many benefits. Medfusion, American Well, and RelayHealth are examples of vendors that provide outsourced solutions.

An integrated patient portal is generally run in house and is directly integrated into a CDO's computer-based patient records and administrative systems. The advantages of this configuration generally include tighter integration and more advanced workflows. The key disadvantage is generally the upfront cost and the cost of ongoing maintenance including personnel to run it. As a result of the complexity of operating an externally accessible patient web site, integrated in-house portals are generally only deployed by mid to large size CDO's. Optionally, some systems may also be integrated with third party systems for shared patient health records. Epic Systems, NextGen and Allscripts are examples of vendors that provide in-house solutions.

Dimension 2: Application

The second dimension in the model described above is application. Examples include public health, education, clinical, and administrative applications. This dissertation is focused specifically on applications that involve patient-provider secure messaging for clinical care.

Patient provider clinical encounters generally consist of one or more of the following activities: (1) history, (2) examination, (3) analysis, (4) diagnosis, (5) treatment plan, and (6) treatment. History, analysis, diagnosis, and treatment plan do not generally require physical interaction between the patient and provider. On the other hand, examination and treatment frequently do, although not always. From an application standpoint, patient-provider secure messaging will generally be used for activities that do not require immediate physical proximity.

Dimension 3: Perspective

The third dimension consists of the perspective from which the analysis is being completed. This dimension is not important for the definition of patient-provider secure messaging, but is critical for the actual economic evaluation of its use. Perspectives can be defined for the consumer (or patient), provider, and society (Bashshur, Shannon and Sapci 2005). This dissertation considers the perspective of patients or providers. While the society perspective is important, particularly from a regulatory standpoint, it is not examined here. Specifically, section 2 is concerned with finding the relationship between secure messaging and office visits from the perspective of patients (health care users) and section 3 looks at costs savings from secure messaging use from the perspective of providers in an integrated delivery system where most members are covered by fixed premiums and there are only a few fee for service consumers.

Section 2: Are Secure Messages and Face-to-Face Office Visits between Patients and Providers Substitutes?

Introduction

Now widely used for a broad range of services and activities, the internet is also an increasingly popular tool for individuals to use for communicating with their health care providers. Most people prefer to use the internet, through secure messaging, to contact their health care provider for at least some of their care needs as shown by a 2002 survey that found that 71% of people said they wanted to communicate with their providers online (Kleiner, et al. 2002). However, the opportunity to do so for most Americans remains low. For instance, in 2008 only 6.7% of U.S. physicians routinely used messaging for clinical communications with clients (Boukus, Grossman and O'Malley 2010).

In addition to providing more convenient and greater access to health care providers, the option of communicating with providers through secured messaging has the potential to decrease both out of pocket and opportunity costs for patients. Secure messaging eliminates time spent traveling to clinics, waiting for an appointment and the time spent in direct contact with physicians. In addition, secure messaging at present, is rarely if ever subject to any cost sharing such as copays or co-insurance, and is usually not subject to deductibles so communicating with one's physcian in this way can be signicantly less expensive than a face to face office visit. With Americans facing greater health care premiums and cost sharing in their insurance, secure messaging has the potential to signifigantly reduce total personal expenditures on health care.

While there are studies that describe the population of secure messaging users (Ralston, Rutter, et al. 2009, Houston, et al. 2004), how frequently the system was used by patients (Ralston, Rutter, et al. 2009, Chen, et al. 2009, Tang, Black and Young 2006), how much patients were willing to pay for such a system (Houston, et al. 2004, Anand, et al. 2005, Bergmo, et al. 2005, RedOrbit 2006), patient satisfaction (Anand, et al. 2005, Ralston, Carrell, et al. 2007, Liederman, Lee, et al., Patient-physician web messaging: the impact on message volume and satisfaction 2005, Rosen and Kwoh 2007, Stalberg, et al. 2008), provider satisfaction (Liederman, Lee, et al., Patient-physician web messaging: the impact on message volume and satisfaction 2005), there is limited research regarding the potential of secure messaging to substitute for face-to-face office visits. A study of the Relay Health web messaging system compared the volume of phone calls at a control and a case site and found that the case message volume (phone plus web) was 13.7% lower and fell 5.84 times faster than control (phone only) (Liederman, Lee, et al., Patient-physician web messaging: the impact on message volume and satisfaction 2005). A survey conducted at the Palo Alto Medical Clinic found that 65% of respondents felt that secure messaging with physicians avoided one or more office visits in the previous year (Tang, Black and Young 2006). And at Kaiser Permanente Northwest (KPNW) patient access to the secure messaging system at KP HealthConnect Online was associated with a significant reduction in the number of office visits and phone calls (Zhou, et al. 2007). There is also evidence that suggests that messaging might complement office visits. The same study at KPNW found that in 19.4% of sampled secure message threads physicians recommended additional services including laboratory tests, phone calls, office visits, procedures, and a health education class. Research conducted at a single primary care clinic in Norway found that 10% of secure messages required an office visit follow-up (Bergmo, et al. 2005).

However, none of the studies identified used economic theory to explain and to test whether secure messages substitute for outpatient care such as office visits. Using a panel sample from an integrated delivery system, this study is the first to estimate demand equations and use information on price to determine the relationship between secure messages and office visits.

The results of this study will support more informed decisions by insurers and insurance sponsors when establishing reimbursement and coverage policies by providing guidance about the expected impact of policies related to cost shares for office and virtual visits.

Methods

Setting

The study was conducted at Group Health, which is one of the largest and oldest integrated health care organizations in the U.S. Group Health provides health care and health insurance to more than 650,000 members throughout Washington State and North Idaho. Two-thirds of members receive care from Group Health Physicians, a large multi-specialty group practice that provides exclusive care to Group Health members at 25 owned and operated primary care clinics and four specialty centers located throughout Western Washington and metropolitan Spokane. The remainder of Group Health's members receive care from a statewide network of contracted providers and facilities.

Group Health has a secure messaging system which was first introduced in April 2002 and is powered by software from Epic Systems. It is accessed through Group Health's patient web portal (www.ghc.org). Individuals that have registered on the web portal and have had their identity verified are able to schedule appointments, view their medications, request pharmacy refills, send messages to providers free of charge, view test results, after-visit summaries, and other functions (Ralston, Carrell, et al. 2007). This study focuses specifically on secure messages exchanged between patients and providers.

Primary and specialty care providers at Group Health were salaried and were compensated $5 per message. Also, providers were expected to respond to patient messages within 24 hours (Ralston, Carrell, et al. 2007).

Sample

The study was conducted among all adults age 18 years and older enrolled in the Western Washington component of the integrated group practice. On average there were approximately 350,000 members each quarter in the Western Washington component out of which 10% or approximately 35,000 members were randomly selected. Individuals included in the study were enrolled for at least one quarter (sustained enrollment for at least 2 out of 3 months of the quarter) during the study period from the fourth quarter of 2003 through the fourth quarter of 2005. The use of this study was approved by Group Health institutional review board.

Measures

Secure Messages and Office Visits

Table 1 lists the descriptive statistics for all variables. Two demand equations were estimated with the number of office visits as the dependent variable in one equation and the number of secure messages in the other equation. The number of office visits represents the number of visits that individuals have face-to-face with their primary and specialty care providers in a given quarter. By comparison, the number of secure messages is defined as the number of threads that individuals have with their primary and specialty care providers in a given quarter through the patient web portal. A thread is defined as a set of messages related to an original message by successive replies. Also, a thread could include multiple different strands of conversation between a patient and a set of providers as long as all conversations originated from the same message. A thread was truncated if it had no further message activity for 30 days. Finally, threads are a reasonable approximation of a single episode health related problem and therefore are recommended units of analysis for secure messages (Carrell and Ralston 2005).

Covariates

Sociodemographic Factors: Age is represented by seven dummy variables. The youngest category indicates whether individuals are between 18 and 20 years of age while the oldest category includes individuals 75 years and older. Gender is also a dummy variable with the value of 1 if an individual is female.

Health status: Health is measured by an overall level of morbidity burden based on an individual's expected need for healthcare and was calculated using the John's Hopkins' Adjusted Clinical Group's (ACG) case mix system. Specifically, all the medical conditions are grouped according to the amount of health care that individuals require. The ACG software assigns each individual a level of overall morbidity between 1 (none) and 6 (very high), depending on age, gender, and the number and types of groups populated by the medical conditions over a 12-month period (Ralston, Rutter, et al. 2009). Also, we consider the presence of unobserved individual heterogeneity (fixed effects) in the form of unobserved mental and physical health.

Access to Care: Types of insurance included private, Medicare, Medicaid, Basic Health Plan, and a dummy indicating whether insurance information was missing.

Price: Outpatient copay represents the out of pocket cost that individuals have to pay when they visit their primary or specialty care providers in office. During the study period, between 2003 and 2005, the copay amount was typically the same for primary and specialty care visits. In some instances copay might have been different between the two types of visits, but such instances were rare. Also, the value of copay has been inflation adjusted and reported in fourth quarter 2005 US dollars.

Empirical Specification

Ex ante, we hypothesize that secure messages are substitutes for non-emergent outpatient office visits. The demand equations for office visits and secure messages are derived from the neoclassical approach of consumer demand. Following Meyerhoefer and Zuvekas (2010), we assume that individuals seek to maximize their wellbeing (utility) subject to a budget constraint. Utility *U()* is a function of health H_{it} and a composite commodity of all the other goods C_{it}, where *i* indexes individuals and *t* indexes time. In addition, health is a stock variable which at time *t* depends on the level of health in the previous period *t*-1. The utility maximization model implies the following demand equations:

$$OV_{it} = q_1(copay_{it}, \varepsilon_{it}; Z_{it}, c_i) \quad (1)$$

$$SM_{it} = q_2(copay_{it}, \varepsilon_{it}; Z_{it}, c_i) \quad (2)$$

where OV_{it} is the number of office visits to primary and specialty care providers for individual *i* in quarter *t*; SM_{it} is the number of secure messages with primary and specialty care providers; $copay_{it}$ is the payment amount that individuals incur every time they visit their primary or specialty care provider; Z_{it} are incorporated to control for individual characteristics such as age, gender, health status, and insurance; c_i controls for unobserved individual heterogeneity (fixed effects); and ε_{it} is the stochastic disturbance.

The price variables are denoted by outpatient copayment for office visits. There is no cost sharing for secure messaging and individuals incur only the opportunity cost of time spent writing messages and reading responses. Information on opportunity cost of time is not available; hence there is no price for secure messages in either (1) or (2).

Finally, following Meyerhoefer and Zuvekas (2010), we assume the presence of unobserved individual heterogeneity c_i, namely fixed effects, such as unobserved mental and physical health which vary by individual but which do not vary over time. If left unresolved, the presence of fixed effects could lead to endogeneity issues because these effects might be correlated with covariates and subsequently lead to inconsistent coefficient estimates.

Estimation Method

There are several challenges in estimating the demand equations for office visits and secure messages. First, the infrequent use of health services by at least some part of the population creates a large number of zeros, so that the number of office visits and the number of secure messages are characterized by a zero mass problem. Second, the distribution of health care use is usually highly skewed to the right among users. There are individuals who use health services very frequently, but such occurrences are relatively rare. Third, overdispersion is common defined by the situation when variance exceeds the mean. Finally, the problem of estimation becomes even more challenging if we believe in the presence of unobserved individual heterogeneity.

In the case of count data, the Poisson process is many times the starting model of choice (Cameron and Trivedi 2005). However, the Poisson model is limited because it cannot explain the presence of excess zeros and because of its inherent property of equal mean and variance. In the case of overdispersion, researchers could use models such as generalized Poisson and negative binomial for estimation. If we also consider the problem of excess zeros then models such as hurdle and zero inflated could be used.

However, the presence of fixed effects for count data with excess zeros and over-dispersion strictly limits the choice of estimation methods. For example, if we believe that the data generating process comes from a zero inflated model, then fixed effects estimation is not possible because there is no sufficient statistic for this purpose.

This study uses the Poisson quasi-conditional maximum likelihood estimation (Poisson QCMLE) method (Wooldridge 1999). The quasi estimator is shown to be consistent under relatively weak assumptions and a sufficient condition for estimation is given by:

$$E(y_{it}|\boldsymbol{x_{it}}, c_i) = c_i exp(\boldsymbol{x}'_{\boldsymbol{it}}\boldsymbol{\beta_0}), \tag{3}$$

where $\boldsymbol{\beta_0}$ is a vector of unknown parameters and $\boldsymbol{x_{it}}$ is a vector of covariates. Nothing is required about variance and higher order moments and there is no restriction about the density of $\boldsymbol{y_i}$. Furthermore, for t $\neq r$, $\boldsymbol{y_{it}}$ is allowed to be correlated with $\boldsymbol{y_{ir}}$.

The Poisson QCMLE method is appealing because it solves the problem of fixed effects and because the distribution of $\boldsymbol{y_i}$ is not restricted and therefore allows distributions with lots of zeros, overdispersion and right skewness. However, the quasi estimator cannot tackle the problem of excess zeros in the same way as a zero inflated or hurdle model can. Specifically, it does not estimate separately the probability of someone using medical services. This is a limitation and we acknowledge that bias can arise in this case. But the choice to estimate a zero inflated model that deals with excess zeros and which does not control for fixed effects comes at a greater cost because of endogeneity issues as the error term becomes correlated with regressors. Also, zero inflated models are highly sensitive to distributional misspecifications (Staub and Winkelmann 2011) while the proposed quasi approach is robust to misspecifications in the data generating process.

Furthermore, it is important to note that the Poisson QCMLE method returns marginal effects in percentage terms. Specifically, if $\widehat{m_j}$ is the estimated effect of the j-th covariate, then a unit increase in the value of the j-th covariate increases the conditional expectation $E(\widehat{y_{it}|\boldsymbol{x_{it}}}, c_i)$ by 100*$\widehat{m_j}$%.

Finally, we use the marginal effects of copay to assess the relationship between secure messages and office visits. If the marginal effect of copay on office visits is negative $\widehat{m_{copay}^{OV}}$<0 and the marginal effect of copay on secure messages is positive $\widehat{m_{copay}^{SM}}$>0, then these two goods are substitutes. On the other hand, if $\widehat{m_{copay}^{OV}}$<0 and $\widehat{m_{copay}^{SM}}$<0, then the goods are complements.

Results

Regression results for estimates of the demand for office visits and secure messages are presented in Table 2.

<u>Increasing Copay Reduces Office Visits and Increases Secure Message Use</u>

The estimated coefficient of copay in model (2) is 0.044 which is positive and statistically significant with a p-value < 0.01. This means that a $1 increase in copay induces individuals to use approximately 4.4% more secure messages. By contrast, the estimated coefficient of copay in model (1) is -0.0045 which is negative and statistically significant with a p-value < 0.1. In this case a $1 increase in copay induces individuals to use 0.45% fewer office visits. The price variable copay has opposite effects on secure messages and office visits. The results indicate that when office visits become more expensive in terms of copay individuals not only decrease the number of times they visits their doctors in office but also increase the number of times they communicate with their providers through messaging. While it was not possible to control for other costs such as travel time to office or deductibles, these are nontheless promising results that suggest that secure messages and office visits are substitutes.

Furthermore, to obtain marginal effects in unit rather than percentage terms we multiply the estimated coefficients of copay from Table 2 with the respective mean utilization rates of office visits and secure messages from Table 1. Hence, a $1 increase in copay induces individuals to use approximately 0.0056 more secure messages and 0.005 fewer office visits. Similarly, for every copay induced reduction of 1.1 office visits there is an increase in 1 messaging thread. This finding does not imply that one messaging thread substitutes for 1.1 office visits, as members may be alternatively substiting phone calls or inaction when not visiting the doctor, but it does show a very high level of overall substitution.

Sicker Individuals Use More Office Visits and More Secure Messages

The estimated coefficients on morbidity are all positive and statistically significant with p-value < 0.01. As individuals display greater morbidity relative to the reference category of no morbidity they tend to use more office visits and more secure messages. Also, the coefficients on morbidity are highest in the demand for office visits which indicate that individuals are more sensitive to morbidity when they choose to visit their providers in office relative to the case of secure messaging.

Older Age Does Not Always Mean More Office Visits

In the regression with secure messaging statistically significant coefficients on age are for individuals of 45 years and older. The positive coefficients indicate that these individuals use more secure messaging relative to young individuals between 18 and 20 years. By contrast, the coefficients on age in the regression with office visits are many times statistically significant and indicate that the number of office visits decreases with age. This result might seem counterintuitive. To investigate further, Table 3 shows that older age is associated with a higher number of office visits when none of the factors including health status are accounted for. However, by properly controlling for health status, a lower utilization rate of office visits might include other factors associated with age such as preference and cost.

Conclusions

The goal of this section was to identify whether secure messages act as substitutes for physician office visits. Based on the neoclassical theory of consumer behavior, demand equations for office visits and secure messages were estimated. Outpatient copay was used as the key variable that individuals consider when deciding to seek care from their providers in office or through secure messaging. By using the Poisson QCMLE method for estimation, it was found that individuals substitute away from office visits towards more secure messaging when office visits become more expensive in terms of copay. This is a strong indication that secure messages and office visits are substitutes.

The strength of the paper lies with the design and estimation method. Individuals have preferences towards the consumption of various goods, but they also respond to changes in price. No other paper used information on price to determine the relationship between secure messages and office visits. This is also the first paper to account for fixed effects in answering the substitution question. Finally, the use of the Poisson QCMLE method is advantageous to the use of a negative binomial, zero inflated, or hurdle models because the latter methods are sensitive to distributional misspecifications. In contrast, the quasi estimator is based on relatively weak assumptions and is robust to distributional misspecifications. Hower, it can not address the zero mass problem in the way a zero inflated can. This is a limitation that Section 4 is trying to address.

Limitations also exist with respect to data and results. First, the budget constraint condition implies that we control for individual disposable income in both demand equations. But income is not tracked by Group Health. This is an important variable that could affect the behavior of individuals. However, if some individuals have relatively stable income over time, the income effect is eliminated through the fixed effects approach. Second, the data sets for prices are not complete. There are other out of pocket costs that individuals incur when they visit their physicians in office or through messaging. For example, going to the doctor's office involves costs to the patient including the opportunity cost of time to schedule an appointment, travel costs, visit copay, any deductibles, and the time spent during the visit. Ideally, one would construct a composite price that includes all the costs mentioned above. However, data limitations preclude doing so. Furthermore, the goal of the paper is to assess whether secure messages and office visits are substitutes. To answer this question it is enough to show the relationship between the two goods when at least one cost varies by individuals over time. If individuals decrease their consumption of office visits and increase their consumption of secure messaging when copay increases, holding everything else constant, then these two goods are substitutes. Conversely, if individuals decrease their consumption of both office visits and secure messaging when copay increases, holding everything else constant, then these two goods are complements. This is precisely what we attempted to test in this study.

Finally, the findings of this paper imply cost savings to health plans and to capitated providers that receive a fixed payment per patient. If the messaging system is relatively less costly to implement and run, providers could save costs because individuals show willingness to switch to receiving care online for at least some medical services. Health Plans benefit when they pay on a per visit basis. Conversely, this paper helps explain the lack of use of messaging by fee-for-service providers; by using messaging the overall number of encounters per patient is reduced, although productivity improvements not addressed in this paper could offset this effect.

Table 1. Descriptive Statistics: All Adults, Q4 2003 – Q4 2005

Variable	Total Number	Mean	Stdv	Min	Max
Secure Messages (Threads)	239330	0.129	0.770	0	57
Office Visits					
Primary and Specialty	239330	1.116	1.739	0	40
Age (years)					
18 – 20	239330	0.051	0.219	0	1
21 – 34	239330	0.178	0.382	0	1
35 – 44	239330	0.168	0.373	0	1
45 – 54	239330	0.227	0.419	0	1
55 – 64	239330	0.184	0.388	0	1
65 – 74	239330	0.095	0.293	0	1
75 years and older	239330	0.098	0.298	0	1
Gender					
Female	239330	0.551	0.497	0	1
Morbidity					
None	224742	0.132	0.339	0	1
Very Low	224742	0.066	0.248	0	1
Low	224742	0.162	0.369	0	1
Moderate	224742	0.462	0.499	0	1
High	224742	0.122	0.327	0	1
Very High	224742	0.056	0.230	0	1
Insurance					
Private	239330	0.860	0.347	0	1
Medicare	239330	0.075	0.264	0	1
Medicaid	239330	0.005	0.072	0	1
Basic Health Plan	239330	0.025	0.155	0	1
Missing	239330	0.035	0.184	0	1
Outpatient Copay	219938	13.063	5.195	0	37.18
Secure Messages (Threads)	239330	0.129	0.770	0	57

Table 2. Estimated Marginal Effects on Secure Messages and Office Visits (Q4 2003 – Q4 2005)

	Poisson QCMLE (Robust SE)	
	Secure Messages	Office Visits: Primary and Specialty
Age		
18 – 20	Reference Category	Reference Category
21 – 34	0.677	-0.085
	(0.61)	(0.089)
35 – 44	0.936	-0.197
	(0.68)	(0.125)
45 – 54	1.764**	-0.269*
	(0.72)	(0.14)
55 – 64	2.356***	-0.314**
	(0.73)	(0.15)
65 – 74	3.146***	-0.398**
	(0.78)	(0.16)
75 years and older	4.478***	-0.390**
	(0.84)	(0.18)
Gender	-	-
Morbidity		
None	Reference Category	Reference Category
Very Low	0.764***	5.049***
	(0.22)	(0.53)
Low	1.246***	5.617***
	(0.2)	(0.53)
Moderate	1.795***	6.250***
	(0.19)	(0.53)
High	2.125***	6.651***
	(0.2)	(0.53)
Very High	2.419***	6.892***
Insurance		
Private	Reference Category	Reference Category
Medicare	-0.472	-0.011
	(0.42)	(0.09)
Medicaid	-	-0.287
		(0.42)
Basic Health Plan	0.093	0.197
	(0.57)	(0.15)
Missing	0.899**	-0.339
	(0.44)	(0.22)
Outpatient Copay	0.044***	-0.0045*
	(0.01)	(0.0027)
Observations	11,428	77,119
Individuals	4,933	33,734

Standard errors are in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 3. Average Number of Primary and/or Specialty Care Office Visits and Secure Messages by Age, Gender, Health Status, and Insurance (Age ≥ 18 years, Q4 2003 – Q4 2005)

	Office Visits: Primary and Specialty	Office Visits: Primary	Office Visits: Specialty	Secure Messages (Threads)
Age (years)				
18-20	0.55	0.4	0.15	0.02
21-34	0.82	0.54	0.28	0.09
35-44	0.85	0.58	0.28	0.12
45-54	0.98	0.64	0.34	0.15
55-64	1.19	0.72	0.47	0.19
65-74	1.65	0.95	0.7	0.14
≥ 75	2.06	1.19	0.87	0.08
Gender				
Female	1.25	0.78	0.47	0.15
Male	0.95	0.59	0.36	0.11
Morbidity				
None	0.001	0.0006	0.0004	0.006
Very Low	0.28	0.25	0.03	0.03
Low	0.45	0.35	0.13	0.05
Medium	1.16	0.74	0.42	0.15
High	2.42	1.39	1.03	0.28
Very High	3.41	1.93	1.45	0.34
Insurance				
Private	1.07	0.67	0.4	0.13
Medicare	1.78	1.04	0.74	0.12
Medicaid	1.48	1.04	0.44	0.05
Basic Health Plan	0.92	0.64	0.28	0.08

Section 3. Does Secure Messaging Reduce Provider Costs in an Integrated Delivery System?

Introduction

The internet is an increasingly popular tool for individuals to use for communicating with their health care providers. Most people prefer to use the internet, through secured messaging, to contact their health care provider for at least some of their care needs. Fo example, a 2002 survey that found that 71% of respondents wanted to communicate with their providers online (Kleiner, et al. 2002). However, the opportunity to do so for most Americans remains low. In 2008 only 6.7% of U.S. physicians routinely used messaging for clinical communications with clients (Boukus, Grossman and O'Malley 2010).

In addition to being convenient and providing greater access to health care, the option of communicating with providers through secure messaging has the potential to decrease both out of pocket and opportunity costs for patients. Secure messaging eliminates time spent traveling to clinics, waiting for an appointment and the time spent in direct contact with physicians. In addition, unlike face to face office visits, secure messaging is frequently exempt from cost sharing such as copays or co-insurance, and is usually not subject to deductibles. Therefore, secure messaging has the potential to signifigantly reduce total personal expenditures on health care.

While there are studies that describe the population of secure messaging users (Ralston, Rutter, et al. 2009, Houston, et al. 2004), how frequently such systems are used by patients (Ralston, Rutter, et al. 2009, Chen, et al. 2009, Tang, Black and Young 2006), how much patients are willing to pay for such a system (Houston, et al. 2004, Anand, et al. 2005, Bergmo, et al. 2005, RedOrbit 2006), patient satisfaction (Anand, et al. 2005, Ralston, Carrell, et al. 2007, Liederman, Lee, et al., Patient-physician web messaging: the impact on message volume and satisfaction 2005, Rosen and Kwoh 2007, Stalberg, et al. 2008), and provider satisfaction (Liederman, Lee, et al., Patient-physician web messaging: the impact on message volume and satisfaction 2005), there is limited research regarding the impact of secure messaging on costs.

One of the few studies to specifically address cost was done with the RelayHealth web messaging system. It found a $1.71 reduction in spending for doctor office visits and a $0.12 reduction in laboratory services spending per month per member after the introduction of web messaging system (Baker, et al. 2005). However, Baker's study has several limitations. The biggest limitation is that the number of observed messages is fairly small at 215. Also, the number of patients communicating with their providers was not reported. By contrast, this study uses 13,055 secure messages and 40,371 individuals. And while Baker's study was done from the perspective of the insurer, the current study examines the savings in costs from the perspective of the provider.

Methods

Setting

The study was conducted at Group Health, which is one of the largest and oldest integrated health care organizations in the U.S. Group Health provides health care and health insurance to more than 650,000 members throughout Washington State and North Idaho. Two-thirds of its members receive care from Group Health Physicians, a large multi-specialty group practice that provides exclusive care to Group Health members at 25 owned and operated primary care clinics, one hospital, and four specialty centers located throughout Western Washington and metropolitan Spokane. The remainder of Group Health's members receive care from a statewide network of contracted providers and facilities.

Group Health has a secure messaging system which was first introduced in April 2002 and is powered by software from Epic Systems. It is accessed through Group Health's patient web portal (www.ghc.org). Individuals that have registered on the web portal and have had their identity verified are able to schedule appointments, view their medications, request pharmacy refills, send messages to providers free of charge, view test results, after-visit summaries, and other functions (Ralston, Carrell, et al. 2007). This study focuses specifically on secure messages exchanged between patients with hypertension and their providers.

Primary and specialty care providers at Group Health are salaried and are compensated $5 per message. Also, providers are expected to respond to patient messages within 24 hours (Ralston, Carrell, et al. 2007).

Sample

The study was conducted among adults identified with hypertension who were at least 18 years old and who were enrolled in the Western Washington component of the integrated group practice. On average there were approximately 350,000 members each quarter in the Western Washington component out of which 10% or approximately 35,000 members were randomly selected. Out of the quartely 10% random sample only adult members identified with hypertension were considered. Individuals with hypertension were defined as members who satisfied at least one of the following three criteria: they were diagnosed with hypertension by their physician, were taking drugs for hypertension, or had uncontrolled blood pressure measures. Approximately 80% of quartely members were identified as adults out of which 27% had hypertension resulting in a final sample size of 7,220 individuals per quarter or 40,371 individuals over the entire study period. Furthermore, individuals included in the study were enrolled for at least one quarter (sustained enrollment for at least 2 out of 3 months of the quarter) during the study period from the first quarter of 2004 through the fourth quarter of 2005. The use of this study was approved by Group Health institutional review board.

Measures

Costs

Table 4 lists the descriptive statistics for all variables. This paper uses six different measures of costs including total, primary, specialty, drugs, inpatient, and emergency care costs. Total costs were calculated as the number of Relative Value Units (RVUs) generated by a given member in a given quarter multiplied by the cost per RVU. Cost per RVU is calculated on the aggregate level by summing up all the costs that Group Health incurs in treating its members including costs on personnel and staff, office space, administrative, technology, and others and dividing these costs by the total number of RVUs generated by all memebers in a given quarter of time. Furthermore, primary, specialty, inpatient and emergency care costs are calculated in a similar way except that only RVUs related to primary, specialty, inpatient, and emergency care accordingly are taken into account. Lastly, costs on drugs are a function of the number of drugs consumed multiplied by the price that Group Health spent to acquire them and including any administrative and personnel costs involved in dispensing drugs. All costs have been inflation adjusted and reported in fourth quarter 2005 US dollars.

Covariates

Control: Individuals with hypertension are considered in good condition if their blood pressure (BP) measures are under control. Specifically, if systolic BP (SBP) is above 140 mm Hg or diastolic BP (DBP) is above 90 mm Hg then individuals were considered to have uncontrolled BP measures (Fishman, et al. 2011). An individual's mean arterial pressure (MAP) was also used as an indicator of whether BP was under control. MAP is defined as the sum of two thirds of DBP and one third of SBP. The variable Control was assigned the value of 1 if MAP was below 106 mm Hg and 0 otherwise. If there was no information on BP measures, variable Control was still 1 if an individual was consuming drugs for hypertension. Finally, to avoid individuals with hypotension (characterized by low BP measures), we excluded those who had MAP $\leq$ 70 mm Hg.

Morbidity: The level of morbidity burden is based on an individual's expected need for healthcare and was calculated using the John's Hopkins' Adjusted Clinical Group's (ACG) case mix system. Specifically, all the medical conditions are grouped according to the amount of health care that individuals require. The ACG software assigns each individual a level of overall morbidity between 1 (none) and 6 (very high), depending on age, gender, and the number and types of groups populated by the medical conditions over a 12-month period (Ralston, Rutter, et al. 2009). Also, we consider the presence of unobserved individual heterogeneity (fixed effects) in the form of unobserved mental and physical health.

Secure Messaging User (*SMU*)**:** SMU is a dummy variable with the value of 1 if a given member in a given quarter had any secure messaging activity and 0 if there were no secure messages. Furthermore, the number of secure messages was defined as the number of threads that individuals had with their primary and specialty care providers in a given quarter through the patient web portal. A thread was defined as a set of messages related to an original message by successive replies. Also, a thread could include multiple different strands of conversation between a patient and a set of providers as long as all conversations originated from the same message. A thread was truncated if it had no further message activity for 30 days (Carrell and Ralston 2005).

Price: Outpatient copayment represents the out of pocket cost that individuals have to pay when they visit their primary or specialty care providers in office. Outpatient copayment is one of three instrumental variables used for secure messaging user status. The value of copay was inflation adjusted and reported in fourth quarter 2005 US dollars.

Sociodemographic Factors: Age is represented by seven dummy variables. The youngest category indicates whether individuals are between 18 and 20 years of age while the oldest category includes individuals 75 years and older. Gender is also a dummy variable with the value of 1 if individuals are female.

Access to Care: Types of insurance included private, Medicare, Medicaid, Basic Health Plan, and a dummy indicating whether insurance information was missing.

Clinics: Although Group Health owns 25 primary care clinics, only 20 of such clinics are located in the Western Washington component of integrated group practice. Therefore, this study uses twenty dummy variables to control for clinics fixed effects. For instance, Clinic 1 has the value of 1 if individuals attend this clinic and 0 if they go to a different clinic. All the other dummies are defined in a similar way.

Theoretical Considerations

Ex ante, we hypothesize that secure messaging users generate lower costs to Group Health relative to non users of secure messaging. This hypothesis is based on the work in Section 2 which found that secure messages serve as substitutes for office visits. Because the use of office visits generally involve costs in terms of office space, staff, technology, and others, then substituting them with relatively lower cost secure messages should be associated with lower costs at Group Health.

This study uses a cost minimization model. For each of its members, Group Health receives a fixed payment amount per year. Given that revenue per member is fixed, Group Health then has an incentive to minimize the costs of treating its members subject to achieving a certain health level:

$$min\,Cost(treating\ member\ i)$$

s.t. treatment of i = f (medical services) = $\overline{H}$

More formally, for each member *i* at each period of time *t*, the goal is to:

$$\min_{K,L} C_{it}(K_{it}, L_{it}, r_t, w_t) = r_t K_{it} + w_t L_{it}$$

s.t. $\overline{H_{it}} = f\big(m_{1it}(K_{it}, L_{it}), m_{2it}(K_{it}, L_{it}), \dots m_{git}(K_{it}, L_{it})\big) = f(K_{it}, L_{it}),$

where K_{it} is capital such as office space and medical equipment used to treat individuals; r_t is the rental rate of capital; L_{it} represents labor such as medical doctors, nurses, and other staff; w_t are wages paid for labor; m_{git} are intermediate products in terms of medical services that produce health such as office visits, secure messages, drugs, and others; and $\overline{H_{it}}$ is the target health level for hypertensive individuals.

Solving the cost minimization problem implies an optimum amount of labor and capital that depend on the prices of labor and capital and the amount of health (final output) produced:

$$L_{it}^* = L_{it}^*(w_t, r_t, \overline{H_{it}})$$

$$K_{it}^* = K_{it}^*(w_t, r_t, \overline{H_{it}})$$

By substituting the optimum L_{it}^* and K_{it}^* into the cost function the minimum cost is derived as:

$$C_{it}^* = C_{it}^*(w_t, r_t, \overline{H_{it}})$$

Estimation Method

Based on the cost minimization model and the availability of data this paper estimates the following specification:

$$C_{it} = C_{it}(\, SMU_{it}, Control_{it}, Morbidity_{it}, Age_{it}, Gender_{i}, Insurance_{it}, Clinic_{it}), \quad (4)$$

where SMU_{it} indicates whether individual i had any messaging activity in period t, $Control_{it}$ indicates whether blood pressure is under control, $Morbidity_{it}$ controls for morbidity levels, Age_{it} controls for age, $Gender_{i}$ controls for any differential effect based on gender, $Insurance_{it}$ indicates various insurance sources, and $Clinic_{it}$ is clinic fixed effect and controls for any differential effect that arises from attending or being assigned to a particular primary care clinic.

There are several challenges in estimating the cost function (4). First, the covariates on blood pressure being under control and secure messaging user status may be correlated with the error term. For example, there are other unaccounted contemporaneous health factors such as cholesterol levels that often move together with BP levels (U 2012) that need to be controlled to consider someone healthy. Also, the dummy variable *Control* does not take into account different levels of blood pressure, hence they are in the error term and may affect someone's decision for secure messaging activity.

A traditional approach for solving the endogeneity problem is to use an instrumental variable approach. This paper uses the Arellano-Bond dynamic panel generalized methods of moments (GMM) estimator available in Stata as xtabond2 command (Arellano and Bond, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations 1991, Arellano and Bover, Another look at the instrumental variables estimation of errorcomponents models 1995, Blundell and Bond 1998, Roodman 2009). Specifically, the task is to identify variables that are correlated with *Control* and *SMU* accordingly but uncorrelated with the error term. One of the instruments identified for secure messaging user status is outpatient copayment for office visits. It was shown in Section 2 that outpatient copayment was a statistically significant predictor of secure messaging use. At the same time, the amount of copayment that individuals have to pay while visiting their provider in office is a predetermined value at the time when the insurance contract is signed and does not change with contemporaneous changes in blood pressure or cholesterol levels, hence it is uncorrelated with the error term. Furthermore, the panel structure of the data provides the opportunity to use lag variables as instruments. The lag of secure messaging user status is one instrument used. The correlation between *SMU* and lag *SMU* is 0.4 which excludes the possibility of the lag *SMU* being a weak instrument. Similarly, the lag value of *Control* is used as an instrument for *Control*. The correlation between the latter two variables is 0.13. Finally, in order to test the validity of instruments the Arellano-Bond dynamic panel GMM method reports the Hansen J statistic under the null hypothesis that "instruments as a group are exogenous." Table 2 presents p-values > 0.7 for the Hansen J statistic in all regressions. Hence, we cannot reject the hypothesis that the instruments used are jointly valid.

Second, the cost minimizing approach implies the existence of labor wages and rental rates of capital that currently are not controlled in (4) due to the lack of such data. Fortunately, the panel structure of the data provides the opportunity to use lagged information. Specifically, if we use the lag of the dependent variable as one regressor we could potentially control for factors that affect costs for which there is no available data.

Furthermore, there could be unobserved individual factors (fixed effects) such as family history or mental and physical states that are unchanged over time but correlated with regressors such as blood pressure under control (Blood pressure and cholesterol 2012), secure messaging user status, or lagged dependent variable. This paper uses the orthogonal deviations approach available as an option in Stata's xtabond2 command to control for individual fixed effects. The approach is to subtract the average of all future available observations of a variable. The benefit of using orthogonal deviations as opposed to first differencing is that lagged variables do not enter the formula; hence they can be used as instruments. Also, the gaps in unbalanced panels are not magnified (Roodman 2009). The orthogonal deviations approach transforms variables in the following way:

$$w_{it+1} = c_{it}\left(w_{it} - \frac{1}{T_{it}}\sum_{s>t} w_{is}\right)$$

where $c_{it} = \sqrt{\frac{\mathrm{T_{it}}}{\mathrm{T_{it+1}}}}$ is a scale factor, the sum is taken over future observations, and $\mathrm{T_{it}}$ is the sum of such observations.

Finally, the orthogonal deviations approach makes the transformed lagged dependent variable $c_{it-1}\left[y_{it-1} - \frac{1}{T_{it-1}}(y_{it} + y_{it+1} + \cdots)\right]$ correlated with the transformed error term $c_{it}\left[\varepsilon_{it} - \frac{1}{T_{it}}(\varepsilon_{it+1} + \varepsilon_{it+2} \ldots)\right]$ because y_{it} is correlated with ε_{it}. A natural instrument for the transformed lagged dependent variable is the untransformed y_{it-1}. This instrument is correlated with the transformed variable because it is part of its formulation and uncorrelated with ε_{it} in the transformed error term because no autocorrelation of level 1 was detected. Specifically, Table 2 reports p-values > 1 from testing the hypothesis of no AR(1) in levels.

Results

Regression results for estimated impact of secure messaging on the costs of care at Group Health are presented in Table 5.

Insignificant Impact on Total, Primary, Specialty, Emergency, and Inpatient Care Costs

The estimated coefficients of secure messaging user status are negative in all 5 columns of Table 5. However, except for column 4, they are not significantly different from zero. Contrary to the underlined hypothesis we do not find evidence that secure messaging users impact the costs of care at Group Health. The Discussion section below outlines possible explanations for this counterintuitive result.

Secure Messaging Users Have Lower Costs on Drugs

In contrast, the fourth column of Table 5 shows that the estimated coefficient of *SMU* is negative \$223.7 and is statistically significant with a p-value < 0.05. We find that individuals that engage in one or more secure messaging activity experience lower drugs costs, perhaps because of reduced utilization rates of drugs. If we divide the estimated coefficient by the average costs on drugs of \$336.69 from Table 4 we obtain 66.4% reduction in costs on drugs per quarter by secure messaging users.

Sicker Individuals Generate Higher Costs

The estimated coefficients on morbidity are most of the time positive and statistically significant for total, primary, and specialty care costs. For inpatient and emergency care costs the most significant morbidity levels are high and very high. As individuals display greater morbidity relative to the reference category of no morbidity they tend to generate higher costs of care at Group Health perhaps because of higher overall utilization and/or intensity rates of services. Also, the only statistically significant coefficients of variable Control is for the regressions with primary care and drugs costs. Individuals who are healthier in terms of keeping their blood pressure levels under control experience -$190.4 or approximately 84.7% reduction in quartely primary care costs and -$196.9 or approximately 58.5% reduction in quartely drugs costs.

Conclusions

The goal of this study was to determine whether the use of secure messaging leads to lower costs of care in an integrated group practice. Based on a cost minimizing model, cost equations were constructed and estimated. Secure messaging user status was used as the key variable to assess the impact on costs. By using an instrumental variable approach via the Arellano-Bond dynamic panel GMM with fixed effects method, this study did not find evidence that secure messaging use impacts total, primary, specialty, emergency, or inpatient care costs. However, we did find that the use of secure messaging lead to a 66.4% reduction in individual drugs costs per quarter.

A significant limitation of the study, which may be obscuring actual costs savings for primary and specialty care, is the way that Group Health calculates and reports member costs. Actual care costs are a function of the required physician and non-physician time to diagnose and treat each member multiplied by the associated costs per unit of time, along with supplies and overhead costs such as costs for office space and insurance. But rather than capturing actual physician and non-physician time per member, Group Health assigns a standardized Current Procedural Terminology (CPT) code for each encounter between patients and providers. Then, the member costs are calculated as the number of RVU values assigned to each CPT code multiplied by an average cost per RVU.

As a result, if the amount of time required to deliver a given CPT code is reduced because of efficiencies gained from the substitution of secure messaging for less efficient work, such a phone-based or face-to-face follow-up visits, the actual cost would be reduced but the reported costs would stay the same. In fact a study in 2005 found that physicians using secure messaging were able to increase their productivity by increasing the overall number of RVUs delivered by 10% per day (Liederman, Lee, et al., The impact of patient-physician Web messaging on provider productivity 2005). The current study, which relies on Group Health per member cost calculations, does not account for such potential costs saving due to increased provider productivity. Other benefits, such as the potential for more flexible work allocation (Hobbs, et al. 2003), are also not accounted for.

Another limitation of the study was the sole use of blood pressure and morbidity levels as controls for member health. While these are important measures, they do not necessarily paint a complete picture of a member's health. Future research could use additional health indicators, such as measures for blood sugar and cholesterol level.

The one area where significant cost savings were identified was in the use of drugs. Nationally, the Center for Disease Control and Prevention (CDC) has found that the number of people with high blood pressure that are treated with drugs was 71.6% in 2007-2008 (Yoon, Ostchega and Louis 2010). Since blood pressure was specifically controlled for as the key measure for health in the study, it's not surprising that this is one area where a statistically significant reduction in costs was identified. Further research is needed to investigate whether these savings were from a reduction in dosage or a change in medication and whether these costs were primarily related to hypertension drugs or for the treatment of other conditions.

Finally, from the time the data for this study was collected, Group Health has seen significant increases in the use of patient provider messaging. Therefore, given the growing importance of messaging in member care, and given the existing study's limitations related to member cost measurement and health level measurement, additional research that addresses these issues is needed to better understand the impact that messaging is having on per member costs at Group Health.

Table 4. Descriptive Statistics: Adults with Hypertension, Q1 2004 – Q4 2005

Variable	Observations	Mean	Stdv	Min	Max
Costs					
Total	57576	1614.39	6315.56	0	373769.2
Primary Care	57576	224.76	394.4	0	16501.85
Specialty Care	57576	199.92	501.23	0	21242.46
Emergency Care	57576	93.72	616.41	0	26986.29
Inpatient Care	57576	631	5803.49	0	373769.2
Drugs	57576	336.69	740.46	0	23808.87
Secure Messaging User (SMU)	57576	0.09	0.29	0	1
Outpatient Copayment	52812	12.67	4.56	0	37.18
Health					
Control: 1 if healthy	57191	0.88	0.32	0	1
Morbidity					
None	55814	0.03	0.18	0	1
Very Low	55814	0.02	0.14	0	1
Low	55814	0.07	0.26	0	1
Moderate	55814	0.52	0.50	0	1
High	55814	0.21	0.41	0	1
Very High	55814	0.14	0.34	0	1
Age (years)					
18 – 20	57576	0.00	0.05	0	1
21 – 34	57576	0.02	0.15	0	1
35 – 44	57576	0.06	0.24	0	1
45 – 54	57576	0.19	0.39	0	1
55 – 64	57576	0.27	0.44	0	1
65 – 74	57576	0.20	0.40	0	1
≥ 75	57576	0.25	0.43	0	1
Gender					
Female	57576	0.55	0.50	0	1
Insurance					
Private	57576	0.80	0.40	0	1
Medicare	57576	0.17	0.37	0	1
Medicaid	57576	0.00	0.05	0	1
Basic Health Plan	57576	0.02	0.12	0	1
Missing	57576	0.02	0.13	0	1
Clinics					
Clinic 1	57576	0.05	0.21	0	1
Clinic 2	57576	0.09	0.28	0	1
Clinic 3	57576	0.02	0.13	0	1
Clinic 4	57576	0.05	0.22	0	1
Clinic 5	57576	0.08	0.27	0	1
Clinic 6	57576	0.04	0.19	0	1
Clinic 7	57576	0.14	0.35	0	1
Clinic 8	57576	0.07	0.25	0	1
Clinic 9	57576	0.05	0.23	0	1
Clinic 10	57576	0.05	0.23	0	1
Clinic 11	57576	0.04	0.20	0	1
Clinic 12	57576	0.06	0.24	0	1
Clinic 13	57576	0.03	0.18	0	1
Clinic 14	57576	0.04	0.19	0	1
Clinic 15	57576	0.02	0.13	0	1
Clinic 16	57576	0.05	0.21	0	1
Clinic 17	57576	0.02	0.14	0	1

Table 4 continued

Variable	Observations	Mean	Stdv	Min	Max
Clinic 18	57576	0.03	0.18	0	1
Clinic 19	57576	0.04	0.21	0	1
Clinic 20	57576	0.02	0.14	0	1

Table 5. Coefficient Estimates (Adults with Hypertension, Q1 2004 – Q4 2005)

	Arellano-Bond Dynamic Panel GMM with Fixed Effects (Robust SE)					
	Total Costs	Primary Care Costs	Specialty Care Costs	Drugs Costs	Inpatient Care Costs	Emergency Care Costs
Lag Costs	-0.124***	0.0428	0.0200	1.069***	-0.149***	-0.00524
	(0.0211)	(0.0525)	(0.0685)	(0.159)	(0.0173)	(0.00856)
SMU	-748.3	-8.654	-41.48	-223.7**	-348.5	-7.997
	(778.4)	(100.4)	(115.6)	(110.2)	(778.2)	(36.50)
Control	67.85	-190.4**	113.9	-196.9*	150.5	-130.8
	(410.4)	(88.56)	(117.7)	(117.2)	(631.3)	(102.8)
Morbidity						
None	Ref Cat	Ref Cat	Ref Cat	Ref Cat	Ref Cat	Ref Cat
Very Low	82.09	27.95**	9.945	-6.522	74.16	-11.45
	(109.6)	(13.28)	(13.88)	(20.29)	(125.9)	(15.30)
Low	224.3*	29.62**	42.03**	-14.65	139.9	-0.218
	(123.6)	(13.76)	(17.55)	(19.65)	(128.9)	(18.87)
Moderate	741.3***	126.0***	140.7***	-5.040	127.4	4.517
	(132.2)	(15.44)	(21.68)	(29.07)	(152.9)	(14.19)
High	2122***	255.1***	288.4***	8.890	703.1***	71.45***
	(219.2)	(23.78)	(31.85)	(60.84)	(229.8)	(19.61)
Very High	6186***	347.0***	440.9***	-79.33	4194***	316.6***
	(886.9)	(32.04)	(49.16)	(106.9)	(927.3)	(54.79)
Age (years)						
18 – 20	Ref Cat	Ref Cat	Ref Cat	Ref Cat	Ref Cat	Ref Cat
21 – 34	502.1	50.51	78.09	46.36	468.3	-128.4
	(414.1)	(49.13)	(64.58)	(73.56)	(345.7)	(99.38)
35 – 44	72.29	20.94	26.47	-16.38	200.2	-110.7
	(361.4)	(43.92)	(57.52)	(72.97)	(310.8)	(97.48)
45 – 54	402.8	22.11	48.18	2.510	569.0	-93.80
	(419.7)	(43.08)	(56.43)	(74.53)	(416.8)	(96.98)
55 – 64	140.2	32.66	79.81	42.27	84.81	-94.44
	(335.4)	(43.39)	(57.82)	(72.14)	(283.8)	(96.92)
65 – 74	59.29	23.94	83.09	39.20	241.0	-99.19
	(379.6)	(44.08)	(57.73)	(74.84)	(329.4)	(96.72)
≥75	-162.0	40.22	81.83	15.76	10.41	-78.44
	(461.7)	(43.81)	(58.28)	(73.94)	(423.9)	(98.20)
Gender	199.9	16.56	-2.643	-8.454	123.1	-1.923
	(202.1)	(10.89)	(12.42)	(19.10)	(200.8)	(15.11)
Insurance						
Private	Ref Cat	Ref Cat	Ref Cat	Ref Cat	Ref Cat	Ref Cat
Medicare	472.4	15.86	4.070	-51.62*	295.6	12.75
	(382.1)	(18.18)	(21.26)	(30.72)	(360.6)	(22.12)
Medicaid	-986.5***	-118.9**	-159.1***	-102.8	-497.5*	-51.85
	(353.8)	(51.40)	(50.41)	(111.4)	(273.2)	(40.40)
Basic Health Plan	581.6	20.79	47.51	-9.397	478.6	78.65
	(674.6)	(36.72)	(51.78)	(57.29)	(550.7)	(106.2)
Missing	-274.8	-74.04	-91.12*	75.48	-460.0	-35.04
	(599.9)	(80.04)	(47.02)	(210.9)	(432.8)	(68.47)
Clinic 1	Ref Cat	Ref Cat	Ref Cat	Ref Cat	Ref Cat	Ref Cat
Clinic 2	191.9	-9.585	54.87	16.20	313.6	-61.55
	(745.0)	(19.36)	(34.91)	(38.10)	(785.2)	(39.10)

Table 5 continued

	Arellano-Bond Dynamic Panel GMM with Fixed Effects (Robust SE)					
	Total Costs	Primary Care Costs	Specialty Care Costs	Drugs Costs	Inpatient Care Costs	Emergency Care Costs
Clinic 3	574.5	33.65	43.17	56.09	340.6	-15.09
	(444.8)	(29.90)	(38.06)	(47.48)	(421.4)	(70.05)
Clinic 4	457.3	37.68	59.98*	44.10	389.8	-20.74
	(597.4)	(31.34)	(30.89)	(45.07)	(516.8)	(47.39)
Clinic 5	89.21	-19.83	86.66**	-3.950	16.57	-74.84**
	(298.7)	(25.85)	(33.92)	(42.20)	(283.2)	(37.98)
Clinic 6	-140.0	9.037	38.64	39.15	-149.6	-37.51
	(327.5)	(28.46)	(38.98)	(40.02)	(303.7)	(46.04)
Clinic 7	8.188	11.15	12.90	1.416	-20.20	-82.80**
	(300.4)	(18.89)	(26.22)	(46.76)	(300.4)	(35.15)
Clinic 8	-196.8	58.76*	12.81	45.09	-194.1	-56.14
	(309.2)	(35.72)	(31.20)	(44.53)	(316.8)	(37.77)
Clinic 9	91.24	31.71	19.42	92.68**	-155.2	-25.81
	(306.1)	(21.63)	(29.62)	(40.17)	(289.5)	(47.60)
Clinic 10	730.1	26.30	17.79	41.55	559.8	-26.15
	(539.7)	(28.01)	(41.35)	(60.11)	(503.9)	(54.39)
Clinic 11	1370	34.11	53.32	31.99	1592*	-14.79
	(865.2)	(28.83)	(44.42)	(88.18)	(890.6)	(59.87)
Clinic 12	115.9	21.90	13.92	12.29	-84.65	-15.20
	(338.4)	(24.90)	(33.56)	(39.79)	(299.8)	(72.39)
Clinic 13	471.8	22.99	61.67	35.32	356.2	-29.13
	(884.0)	(26.82)	(39.29)	(49.43)	(842.9)	(51.58)
Clinic 14	-321.6	-1.147	-18.91	26.80	-158.3	-53.71
	(370.0)	(24.21)	(35.71)	(40.74)	(341.9)	(39.53)
Clinic 15	-112.8	0.930	48.50	17.61	-24.83	-73.21*
	(356.0)	(27.02)	(45.12)	(40.40)	(325.5)	(38.13)
Clinic 16	52.64	27.68	49.61	89.51	-251.2	37.98
	(320.5)	(24.27)	(49.66)	(61.55)	(285.0)	(63.98)
Clinic 17	20.87	33.47	20.18	114.8*	-106.8	-89.96**
	(394.5)	(36.94)	(64.17)	(59.31)	(341.7)	(41.43)
Clinic 18	466.4	-10.41	55.71	40.40	493.6	-39.91
	(671.2)	(24.25)	(35.97)	(47.62)	(637.5)	(44.37)
Clinic 19	544.1	71.69*	-11.68	48.07	402.0	-45.52
	(611.6)	(39.82)	(32.94)	(51.10)	(563.8)	(42.01)
Clinic 20	547.6	30.57	40.98	51.28	300.7	37.10
	(418.2)	(32.43)	(61.31)	(41.32)	(383.3)	(53.22)
Constant	-448.1	142.4	-208.6*	149.8	-642.5	264.9*
	(677.1)	(90.62)	(124.0)	(125.4)	(694.7)	(141.7)
Observations	5380	5380	5380	5380	5380	5380
Individuals	4819	4819	4819	4819	4819	4819
Hansen J statistic H_0: exogenous (p-value)	0.734	0.786	0.677	0.965	0.919	0.968
H_0: no AR(1) in levels (p-value)	0.28	0.35	0.553	0.3	0.38	0.356

Section 4: Quasi Likelihood Estimation of Zero Inflated Models in the Presence of Unobserved Heterogeneity

Introduction

Researchers who want to want to work with health care data are usually faced with several econometric challenges. First, the infrequent use of health services by at least some part of the population creates excess zeros. As a result, continuous data such as health care costs and count data such as the number of office visits to providers are characterized by a zero mass problem. Second, the distribution of health care use is usually highly skewed to the right. There are individuals who use health services very frequently but such occurrences are relatively rare. Third, overdispersion is common defined by the situation when variance exceeds the mean. The implication is that a typical Poisson process can not fit count data very well. Finally, the problem of estimation becomes even more challenging if researchers believe in the presence of unobserved heterogeneity, random effects, or population average effects.

In the case of count data, the Poisson regression is usually the starting model of choice (Cameron and Trivedi 2005). However, the Poisson model is limited because it cannot explain the presence of excess zeros and because of its inherent property of equal mean and variance. In case of overdispersion, researches could use models such as generalized Poisson and negative binomial for estimation. If we also consider the problem of excess zeros then models such as hurdle and zero inflated could be used. Specifically, hurdle models relax the assumption that the zeros and the positive outcomes come from the same distribution. A logit/probit model is used to predict the probability of zero outcomes and a zero truncated Poisson or negative binomial explains the distribution of positive outcomes. Furthermore, a zero inflated model allows the zeros to come from two distributions. A logit/probit model is used to predict the probability of someone having only zero outcomes and a Poisson or negative binomial explains the distribution of outcomes for someone who has both zero and positive outcomes.

Applications using zero inflated models can be found in manufacturing (Lambert 1992), dental care (Böhning, et al. 1999), horticulture (Hall 2000), foetal and postnatal somatic growth (Cheung 2002), apple roots (Ridout, Hinde and Demetrio 2001), species occurrence and abundance (Wenger and Freeman 2008), and others. The use of hurdle models can be found in explaining doctor visits (d'Uva 2006), inequalities in the use of general practitioners and specialist services (Doorslaer, Koolman and Jones 2004), and others.

The presence of fixed effects for count panel data with excess zeros and overdispersion strictly limits the choice of estimation methods. While it is possible to condition out the fixed effects in a Poisson, negative binomial regression framework (Hausman, Brownwyn and Griliches 1984, Allison and Waterman 2002), or hurdle model (Winkelmann 2008), up until recently it was not possible to do so in a zero inflated model because there was no sufficient statistic for this purpose. There is a recent working paper that developed a fixed effects approach in a zero inflated Poisson framework (Majo and Soest 2011). However, such models are not robust to distributional misspecifications because they rely on rather restrictive assumptions.

The goal of this study is to estimate the parameters of a count model in the presence of excess zeros and unobserved individual heterogeneity (fixed effects) when the true generating process in unknown. The quasi conditional framework cannot consistently estimate such models unless the conditional mean assumption is correctly specified (Wooldridge 1999). Current statistical packages such as Stata and R have implemented a somewhat limited version of the quasi approach titled Poisson quasi conditional maximum likelihood estimator (Poisson QCMLE). Provided that the conditional mean is Poisson or negative binomial, this method consistently estimates a variety of processes with fixed effects while ignoring the amount of zeros that cannot be explained by a traditional count framework such as Poisson.

This study extends the Poisson QCMLE method used in current statistical packages to include the ability to estimate zero inflated (ZI) models with fixed effects. Specifically, we propose a new estimator titled ZI QCMLE that would be consistent for ZI models in addition to Poisson and negative binomial models. This estimator is an improvement over Poisson QCMLE because it is consistent under a larger class of data generating processes.

Specific Aim

The purpose of this study is to test the performance and robustness of the new ZI QCMLE method when count data are characterized by the presence of a zero mass problem and unobserved heterogeneity (fixed effects). In order to address this question, two different processes are generated for count data commonly found in the health economics literature including zero inflated negative binomial (ZINB) and zero inflated Poisson (ZIP). Both of these processes are designed to have a zero mass problem in one case with 10% zero inflation and in another case with 50% zero inflation. Then 100 Monte Carlo simulations are performed to compare parameters and marginal effects predicted by theory with parameters and marginal effects estimated via the ZI QCMLE method. For comparison purposes, we also estimate marginal effects via Poisson QCMLE and Ordinary Least Squares (OLS).

Significance

The proposed new estimator is useful for researchers who work with count data characterized by overdispersion/underdispersion, skewness, zeros mass problem, and unobserved individual heterogeneity. In addition, this method is especially valuable if the researcher is not sure about the true data generating process because it is robust to distributional misspecification. Specifically, only the mean has to be correctly specified and no restrictions exist about the true data generating process.

Innovation

Only one paper was identified with research agenda close to this study. Others have looked at estimating zero inflated processes via a quasi method but without unobserved individual component (Staub and Winkelmann 2011). However, no paper so far has looked at a quasi conditional method for different types of count data in cases when they are characterized by both a zero mass problem and unobserved individual heterogeneity.

The QCMLE: Theoretical Background

Let $\{(\boldsymbol{y_i}, \boldsymbol{x_i}, c_i): i = 1,2, \dots\}$ be a sequence of iid random variables, where $\boldsymbol{y_i}$ is Tx1, $\boldsymbol{x_i}$ is TxK, and c_i is the unobserved individual characteristic. It has been shown that the QCMLE estimator is consistent under relatively weak assumptions (Wooldridge 1999). It is enough to assume that:

$$E(y_{it}|\boldsymbol{x_i}, c_i) = c_i\mu(\boldsymbol{x_{it}}, \boldsymbol{\beta_0}), \text{ t} = 1,\dots\text{T}, \tag{5}$$

where $\boldsymbol{\beta_0}$ is a vector of unknown parameters and is estimated by maximizing a log likelihood function. Specifically, the log likelihood for observation i is:

$$l_i = \sum_{t=1}^{T_i} y_{it} \log\left[p_t(x_i, \beta)\right]$$
$$p_t(x_i, \beta) \equiv \frac{\mu(x_{it}, \beta)}{\sum_{r=1}^{T} \mu(x_{ir}, \beta)}$$

where the unobserved component c_i is cancelled out in the expression for p_t.

Furthermore, this method does not require any assumptions about variance and higher order moments and there is no restcriction about the density of $\boldsymbol{y_i}$. In addition, for t $\neq r$, $\boldsymbol{y_{it}}$ is allowed to be correlated with $\boldsymbol{y_{ir}}$. It is also worth noting that this method does not require $\boldsymbol{y_i}$ to be a vector of counts. The dependent variable could be binary or continuous with nonnegative values.

This method proves to be very useful for various applications especially if the researcher is not sure about the true generating process of $\boldsymbol{y_i}$. And even if the researcher strongly believes that the generating process is hurdle or zero inflated, it is not possible to condition out the effect c_i in which case the QCMLE method can be used. Finally, this method can be applied for cases when $\boldsymbol{y_i}$ is characterized by a zero mass problem, precisely because the distribution of $\boldsymbol{y_i}$ is not restricted and therefore can allow distributions that generate lots of zeros.

Zero Inflated Model

A data generating process is defined to have excess zeros if the number of zeros exceeds the amount predicted by a traditional model such as Poisson. The ZI model was precisely designed to overcome this problem. Specifically, there are two groups of values considered. One group includes only zero values and is called the always zero group. The other group includes both zero and positive values. For example, in the case of medical services, an individual would be in the always zero group if he/she does not go to doctors (not a user). If, however, an individual is a user of medical services but sometimes visits and sometimes does not visit providers then he/she is in the group with both zero a positive values.

Two processes are involved in the estimation of a ZI model. The logit/probit model is used to predict whether a person is in the always zero group. If the person belongs to the group of both zero and positive values, then such values are usually Poisson or negative binomial distributed. The conditional mean of the ZIP/ZINB model is defined as:

$$E(y_{it}|\boldsymbol{x_{it}}, \boldsymbol{z_{it}}, c_i) = (1 - \varphi_{it})\lambda_{2it} = \frac{\lambda_{2it}}{1 + \lambda_{1it}} = c_i \frac{\exp{(\boldsymbol{x}'_{it}\boldsymbol{\beta})}}{1 + \exp{(\boldsymbol{z}'_{it}\boldsymbol{\gamma})}} \quad (6)$$

where $\varphi_{it} = \frac{\lambda_{1it}}{1+\lambda_{1it}}$ is the probability of being a nonuser of medical services with $\lambda_{1it} = \exp{(\boldsymbol{z}'_{it}\boldsymbol{\gamma})}$ derived from a logit process, and $\lambda_{2it} = c_i \exp{(\boldsymbol{x}'_{it}\boldsymbol{\beta})}$ is the mean of a Poisson or negative binomial process. The unknown parameters to be estimated are $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

Extension from Poisson QCMLE to ZI QCMLE

Poisson QCMLE is a special case in that the conditional assumption (5) comes from a Poisson or negative binomial process:

$$E(y_{it}|\boldsymbol{x_i}, c_i) = c_i exp(\boldsymbol{x}'_{it}\boldsymbol{\beta_0}), \quad (7)$$

Consistency of $\boldsymbol{\beta_0}$ is assured provided that (7) holds. However, if data is characterized by an excess of zeros then the estimated parameters might be inconsistent as the conditional mean assumption (7) may no longer hold true. Instead, form (6) might be the correct specification.

The new ZI QCMLE method assumes that (6) is the correct conditional mean specification. The log-likelihood for observation *i* for the ZI QCMLE approach has the form:

$$l_i = \sum_{t=1}^{T_i}\left(y_{it}\log\widetilde{\lambda_{it}} - \widetilde{\lambda_{it}}n_i\right), \tag{8}$$

where

$$\widetilde{\lambda_{it}} = \frac{\frac{\lambda_{2it}}{1+\lambda_{1it}}}{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}}, \quad n_i = \sum_{t=1}^{T_i} y_{it}$$

The problem of fixed effects disappears as the component c_i is cancelled out in the expression for $\widetilde{\lambda_{it}}$.

Observe that the second term in (8) is safely added as it reduces to a constant and does not impact the maximization process:

$$\sum_{t=1}^{T_i}\widetilde{\lambda_{it}}n_i = \frac{\sum_{t=1}^{T_i}\frac{\lambda_{2it}}{1+\lambda_{1it}}}{\sum_{r=1}^{T_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}}n_i = n_i$$

The first order conditions from maximizing (8) are derived as:

$$\frac{\partial l_i(\beta,\gamma)}{\partial\beta} = \sum_{t=1}^{T_i}\left(y_{it} - n_i\widetilde{\lambda_{it}}\right)x_{it} = 0$$

and

$$\frac{\partial l_i(\beta,\gamma)}{\partial\gamma} = \sum_{t=1}^{T_i}\left(-y_{it} + n_i\widetilde{\lambda_{it}}\right)\frac{\lambda_{1it}}{1+\lambda_{1it}}z_{it} = 0$$

The full derivation can be found in Appendix A. These first order conditions are plain orthogonality conditions between residuals and regressors with the second condition weighted by the probability of being in the always zero group $\varphi_{it} = \frac{\lambda_{1it}}{1+\lambda_{1it}}$. There are no identification restrictions and regressors ***x*** and ***z*** can be independent, correlated, or equal to each other.

Simulation Study

Monte Carlo Design

Currently, there are no methods implemented in statistical packages that would estimate ZI models with fixed effects. For comparison purposes, we present the estimates of the new ZI QCMLE along with Poisson QCMLE, and Ordinary Least Squares (OLS) with fixed effects. In the experimental design, the count variable y is drawn from a ZI process:

$$y_{it} = \begin{cases} 0, & \text{with probability } \varphi_{it} \\ y_{it}^{*}, & \text{with probability } 1 - \varphi_{it} \end{cases}$$

where $\varphi_{it} = \frac{\lambda_{1it}}{1+\lambda_{1it}}$, $y_{it}^{*} \sim Poisson(\lambda_{2it})$ if y_{it} is from a ZIP process and $y_{it}^{*} \sim Negative\ Binomial(\lambda_{2it}, \theta)$ if y_{it} is from a ZINB process. Furthermore, λ_{1it}=exp($\gamma_0 + \gamma z_{it}$), λ_{2it} =exp($\beta_0 + \beta x_{it} + c_i$) with c_i as the unobserved component (fixed effect). The parameter $\theta = 0.5$ is the dispersion parameter in the negative binomial process. The degree of zero inflation is measured by probability φ_{it} and can vary as a function of γ_0, γ, and z_{it}. The Stata code developed to estimate the ZI QCMLE method (see Appendix B) cannot yet estimate γ correctly when $\gamma_0 \neq 0$. Therefore, $\gamma_0 = 0$ in all simulations and the degree of zero inflation is manipulated by changing z_{it}.

There are four experimental setups. First, the count variable y_{it} is drawn from a ZIP process, $x_{it} = \sqrt{0.1}v1_i + \sqrt{0.15}v3_{it}$, $c_i = \sqrt{0.4}v1_i + \sqrt{0.6}v2_i$ and is correlated with x_{it} with all three $v1_i, v2_i, v3_{it} \sim N(0,1)$, and $\boldsymbol{z} \neq \boldsymbol{x}$ with $z_{it} \sim N(0,1)$ for 50% zero inflation and $z_{it} \sim N(-2.5,1)$ for 10% zero inflation. The second setup is similar to the first except that $\boldsymbol{z} = \boldsymbol{x}$ with $v3_{it} \sim N(0,1)$ for 50% zero inflation and $v3_{it} \sim N(-5.7,1)$ for 10% zero inflation. Finally, the last two setups are identical to the first two except that y_{it} is drawn from a ZINB process.

The primary focus is on the estimation of β and γ which are both set equal to 1. However, Poisson QCMLE and OLS cannot estimate β and γ separately. Rather, they estimate marginal effects in either unit or percentage terms. Specifically, when $\boldsymbol{z} \neq \boldsymbol{x}$ marginal effects in percentage terms are theoretically equal to β:

$$\frac{\partial E(y_{it}|x_{it}, z_{it}, c_i)/E(y_{it}|x_{it}, z_{it}, c_i)}{\partial x_{it}} = \beta$$

When $\boldsymbol{z} = \boldsymbol{x}$:

$$\frac{\partial E(y_{it}|x_{it}, z_{it}, c_i)/E(y_{it}|x_{it}, z_{it}, c_i)}{\partial x_{it}} = -\varphi_{\mathrm{it}}\gamma + \beta$$

Therefore, along with estimated $\hat{\beta}$ and $\hat{\gamma}$ this paper also reports estimated marginal effects from the ZI QCMLE method to be compared with those by Poisson QCMLE and OLS.

Results

The results of the first two experimental setups are presented in Table 7 and the results of the last two setups are presented in Table 8. Each table is divided into two sections. The first section corresponds to the case when x is independent of z and the second section presents the results when there is 100% correlation $\boldsymbol{z} = \boldsymbol{x}$. Then, each section is further subdivided to present the cases of 10% and 50% zero inflation. The main entries are the mean of estimated $\hat{\beta}$, $\hat{\gamma}$, and marginal effects over 100 replications and the numbers in parentheses are standard deviations.

Case when z is independent of x

In both tables the ZI QCMLE estimates of β are very close to the true value of 1 regardless of sample size and whether the degree of zero inflation is 10% or 50%. As the degree of zero inflation goes up there is less information to estimate β. Consequently standard errors are slightly larger, although the effect is negligible. Also, since Poisson QCMLE and OLS have marginal effects in percentage terms equal to β, their estimates are very close to ZI QCMLE estimates. Again, since there is less information in the smaller sample case the standard errors are slightly larger.

Furthermore, while the estimates of γ are very close to the true value of 1 when the degree of zero inflation is 50%, it is biased downward in the 10% zero inflation case because there is less information to estimate γ. Between the two tables the bias ranges from 3.3-10.2% in the small sample case and 0.4-2% in the large sample case.

Finally, while all three methods estimate marginal effects (equal to β) free of bias, standard errors are generally the smallest for ZI QCMLE and the largest for OLS. Also, the advantage of ZI QCMLE lies in the ability to estimate γ when neither Poisson QCMLE nor OLS can do that.

Case when $z = x$

Previously we considered the case when the regressors in the parent distribution are independent from the regressors in the logit part. While this is an interesting case, in reality such regressors are most likely correlated with each other. It is hard to imagine the case that there are no overlapping factors that both affect the probability of medical care use and the amount of medical care consumed. In this subsection we consider the extreme scenario that all the factors in the logit part that affect the probability of medical care use are the same as the factors in the parent distribution which affect the amount consumed.

The estimation of such a model requires a considerably larger sample size and 1,000 observations in each period of time proved too small to consider. Therefore, the small sample size was increased to 5,000 per time period. Even so, the ZI QCMLE estimates are notably worse than in the previous case although considerably better than either Poisson QCMLE or OLS. However, the quasi estimation may not be a problem if more regressors are available since variation from more regressors can help estimate the parameters more precisely (Staub and Winkelmann 2011).

In both tables, while the estimated parameters $\hat{\beta}$ and $\hat{\gamma}$ are slightly downward biased when the degree of zero inflation is 50%, the estimated marginal effects are very close to the true value of 0.5. At the same time Poisson QCMLE and OLS have much worse estimates for marginal effects. This is because they are not designed to take into account the logit part and γ becomes an important parameter in the expression for theoretical marginal effects when x and z are correlated or equal to each other. However, when the degree of zero inflation is 10% the logit part becomes less important and both Poisson QCMLE and OLS estimates are closer to the true values although still worse than ZI QCMLE. At the same time the estimates $\hat{\beta}$ and $\hat{\gamma}$ are less precise with $\hat{\gamma}$ affected the most as there is less information for its estimation in the 10% zero inflation case. At last, all estimates become closer to true values as the sample size gets larger regardless of the degree of zero inflation.

Finally, when the degree of zero inflation is large enough ZI QCMLE provides good estimates for β, γ, and marginal effects and this method is doing better than both Poisson QCMLE and OLS. Another advantage is that ZI QCMLE is able to estimate β and γ separately, while the other two methods cannot do so.

Conclusions

The goal of this study was to test the performance and robustness of the ZI QCMLE method for two count data processes characterized by unobserved individual heterogeneity with different degrees of zero inflation and different sample sizes. In addition, this method was compared with Poisson QCMLE and OLS. Monte Carlo simulations show that the ZI QCMLE method is the best among available methods. Even when the regressors used to predict the probability of zeros are independent from the regressors used to predict the amount of medical services (or any other goods) consumed, the ZI QCMLE has the advantage that it is able to estimate all parameters separately, while Poisson QCMLE and OLS can't.

The advantage of the ZI QCMLE to a ZI model with fixed effects (Majo and Soest 2011) is that it is robust to distributional misspecifications. The ZI QCMLE method estimates in the same way whether the underlining data generating process comes from a ZIP, ZINB, or other model. In fact the only assumption required is that the conditional mean is correctly specified. Higher order moments are not considered. This is a very useful property as in practice it is hard to know the exact nature of the data generating process. If a ZIP/ZINB method with fixed effects is used, even slight misspecifications in the data generating process can lead to considerable biases (Staub and Winkelmann 2011). Therefore, the ZI QCMLE method may be a better choice in estimating zero inflated models with fixed effects.

A limitation of this study is that the Stata code developed works for only the case when the logit part that explains the probability of zeros has no constant term. To be even more useful, the code should be enhanced to include the ability to estimate with a constant term in the logit part.

Table 6. ZIP: Monte Carlo Results from 100 Simulations

	ZIP, FE: x ≠ z, $\boldsymbol{\beta} = \boldsymbol{\gamma} = 1$, Marginal Effects = 1					
Estimator	**10% zero-inflation**			**50% zero-inflation**		
	$\hat{\beta}$	$\hat{\gamma}$	$\widehat{Marg\,Eff}$	$\hat{\beta}$	$\hat{\gamma}$	$\widehat{Marg\,Eff}$
N=1,000; T=9						
ZI QCMLE	1.0066	0.9667	1.0066	1.0000	1.0111	1.0000
	(0.0304)	(0.2619)	(0.0304)	(0.0746)	(0.0767)	(0.0746)
Poisson QCMLE	1.0064	-	1.0064	0.9997	-	0.9997
	(0.0311)		(0.0311)	(0.0784)		(0.0784)
OLS	1.0069	-	1.0069	0.9963	-	0.9963
	(0.0460)		(0.0460)	(0.0866)		(0.0866)
N=30,000; T=9						
ZI QCMLE	1.0001	0.9962	1.0001	1.0005	1.0005	1.0005
	(0.0057)	(0.0338)	(0.0057)	(0.0131)	(0.0141)	(0.0131)
Poisson QCMLE	1.0000	-	1.0000	1.0003	-	1.0003
	(0.0059)		(0.0059)	(0.0144)		(0.0144)
OLS	0.9996	-	0.9996	1.0007	-	1.0007
	(0.0080)		(0.0080)	(0.0164)		(0.0164)
	ZIP, FE: x = z, $\boldsymbol{\beta} = \boldsymbol{\gamma} = 1$					
	Marginal Effects = 0.898			**Marginal Effects = 0.5**		
N=5,000; T=9						
ZI QCMLE	0.7876	0.3443	0.8799	0.9765	0.9550	0.4992
	(0.2901)	(0.7922)	(0.0357)	(0.1672)	(0.3175)	(0.0286)
Poisson QCMLE	-	-	0.8485	-	-	0.4294
			(0.0334)			(0.0395)
OLS	-	-	0.8436	-	-	0.4276
			(0.0373)			(0.0391)
N=30,000; T=9						
ZI QCMLE	0.9590	1.3036	0.8873	0.9910	0.9821	0.5
	(0.0841)	(0.7634)	(0.0145)	(0.0672)	(0.1229)	(0.0133)
Poisson QCMLE	-	-	0.8508	-	-	0.4310
			(0.0137)			(0.0163)
OLS	-	-	0.8486	-	-	0.4294
			(0.0153)			(0.0162)

Table 7. ZINB: Monte Carlo Results from 100 Simulations

	ZINB, FE: x ≠ z, $\boldsymbol{\beta} = \boldsymbol{\gamma} = 1$, Marginal Effects = 1					
Estimator	**10% zero-inflation**			**50% zero-inflation**		
	$\hat{\beta}$	$\hat{\gamma}$	$\widehat{Marg\,Eff}$	$\hat{\beta}$	$\hat{\gamma}$	$\widehat{Marg\,Eff}$
N=1,000; T=9						
ZI QCMLE	1.0056	0.8978	1.0056	0.9923	1.0170	0.9923
	(0.1059)	(0.8232)	(0.1059)	(0.1522)	(0.1292)	(0.1522)
Poisson QCMLE	1.0054	-	1.0054	0.9933	-	0.9933
	(0.1062)		(0.1062)	(0.1549)		(0.1549)
OLS	1.0090	-	1.0090	0.9907	-	0.9907
	(0.1112)		(0.1112)	(0.1576)		(0.1576)
N=30,000; T=9						
ZI QCMLE	1.0024	0.9798	1.0024	0.9999	1.0016	0.9999
	(0.0202)	(0.1242)	(0.0202)	(0.0280)	(0.0303)	(0.0280)
Poisson QCMLE	1.0022	-	1.0022	1.0017	-	1.0017
	(0.0201)		(0.0201)	(0.0279)		(0.0279)
OLS	1.0021	-	1.0021	1.0019	-	1.0019
	(0.0211)		(0.0211)	(0.0300)		(0.0300)
	ZINB, FE: x = z, $\boldsymbol{\beta} = \boldsymbol{\gamma} = 1$					
	Marginal Effects = 0.892			**Marginal Effects = 0.5**		
N=5,000; T=9						
ZI QCMLE	0.8656	1.2501	0.8760	0.9783	0.9629	0.4972
	(0.3079)	(1.2594)	(0.0483)	(0.4840)	(0.9276)	(0.15)
Poisson QCMLE	-	-	0.8494	-	-	0.3918
			(0.0530)			(0.1939)
OLS	-	-	0.8481	-	-	0.3912
			(0.0551)			(0.1946)
N=30,000; T=9						
ZI QCMLE	0.9372	1.2530	0.8802	0.9946	0.9742	0.5076
	(0.1143)	(0.9524)	(0.0198)	(0.2395)	(0.4513)	(0.0536)
Poisson QCMLE	-	-	0.8503	-	-	0.4308
			(0.0226)			(0.0650)
OLS	-	-	0.8471	-	-	0.4292
			(0.0235)			(0.0656)

Bibliography

Allison, PD, and R Waterman. "Fixed-effects negative binomial regression models." *Sociological Methodology* 32, no. 1 (2002): 247-265.

Anand, SG, MJ Feldman, DS Geller, A Bisbee, and H Bauchner. "A content analysis of e-mail communication between primary care providers and parents." *Pediatrics* 115, no. 5 (May 2005): 1283-1288.

Arellano, M, and O Bover. "Another look at the instrumental variables estimation of errorcomponents models." *Journal of Econometrics* 68 (1995): 29-51.

Arellano, M, and S Bond. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies* 58 (1991): 277-297.

Baker, L, J Rideout, P Gertler, and K Raube. "Effect of an internet-based system for doctor-patient communication on health care spending." *J Am Med Inform Assn* 12, no. 5 (2005): 530-536.

Bashshur, R, G Shannon, and H Sapci. "Telemedicine evaluation." *Telemed J E Health*, 2005: 296-316.

Bergmo, TS, PE Kummervold, D Gammon, and LB Dahl. "Electronic patient-provider communication: will it offset office visits and telephone consultations in primary care?" *Int J Med Inform* 74, no. 9 (Sep 2005): 705-710.

Blundell, R, and S Bond. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics* 87 (1998): 11-143.

Böhning, D, E Dietz, P Schlattmann, L Mendonça, and U Kirchner. "The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology." *Journal of the Royal Statistical Society Series A* 162, no. 2 (1999): 195-209.

Boukus, ER, JM Grossman, and AS O'Malley. "Physicians slow to e-mail routinely with patients." *Issue Brief Cent Stud Health Syst Change*, no. 134 (Oct 2010): 1-5.

Cameron, AC, and PK Trivedi. *Microeconomics: methods and applications.* Cambridge: Cambridge University Press, 2005.

Carrell, D, and JD Ralston. "Messages, strands and threads: measuring use of electronic patient-provider messaging." *AMIA Annu Symp Proc*, 2005: 913.

Chen, C, T Garrido, D Chock, G Okawa, and L Liang. "The Kaiser Permanente electronic health record: transforming and streamlining modalities of care." *Health Affair* 28, no. 2 (2009): 323-333.

Cheung, YB. "Zero-inflated models for regression analysis of count data: a study of growth and development." *Stat Med* 21, no. 10 (May 2002): 1461-9.

Dávalos, ME, MT French, AE Burdick, and SC Simmons. "Economic evaluation of telemedicine: review of the literature and research guidelines for benefit-cost analysis." *Telemed JE Health*, 2009: 933-948.

Doorslaer, E, X Koolman, and AM Jones. "Explaining income-related inequalities in doctor utilisation in Europe." *Health Economics* 13, no. 7 (June 2004): 629-647.

d'Uva, TB. "Latent class models for utilisation of health care." *Health Economics* 15, no. 4 (2006): 329-343.

Edwards, J. "Hype cycle for telemedicine, 2011." *Gartner.* July 28, 2011. http://my.gartner.com/portal/server.pt?open=512&objID=260&mode=2&PageID=34607

02&resId=1754914&ref=QuickSearch&content=html#h-N65801 (accessed April 6, 2012).

Fishman, PA, et al. "Accuracy of blood pressure measurements reported in an electronic medical record during routine primary care visits." *J Clin Hypertens (Greenwich)* 13, no. 11 (Nov 2011): 821-828.

Hall, DB. "Zero-inflated Poisson and binomial regression with random effects: a case study." *Biometrics* 56 (Dec 2000): 1030-1039.

Hausman, J, HH Brownwyn, and Z Griliches. "Econometric models for count data with an application to the patents-R&D relationship." *Econometrica* 52, no. 4 (July 1984): 909-938.

Hobbs, J, et al. "Opportunities to enhance patient and physician e-mail contact." *Int J Med Inform*, 2003: 1-9.

Houston, TK, DZ Sands, MW Jenckes, and DE Ford. "Experiences of patients who were early adopters of electronic communication with their physician: satisfaction, benefits, and concerns." *Am J Manag Care* 10, no. 9 (Sep 2004): 601-608.

Kleiner, KD, R Akers, BL Burke, and EJ Werner. "Parent and physician attitudes regarding electronic communication in pediatric practices." *Pediatrics* 109, no. 5 (May 2002): 740-744.

Lambert, D. "Zero-inflated Poisson regression, with an application to defects in manufacturing." *Technometrics* 34, no. 1 (Feb 1992): 1-14.

Liederman, EM, JC Lee, VH Baquero, and PG Seites. "Patient-physician web messaging: the impact on message volume and satisfaction." *J Gen Intern Med* 20, no. 1 (Jan 2005): 52-57.

Liederman, EM, JC Lee, VH Baquero, and PG Seites. "The impact of patient-physician Web messaging on provider productivity." *J Healthc Inf Manag*, 2005: 81-86.

Majo, MC, and AV Soest. *The fixed-effects zero-inflated Poisson model with an application to health care utilization.* July 25, 2011. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1899582 (accessed April 4, 2012).

Meyerhoefer, CD, and SH Zuvekas. "New estimates of the demand for physical and mental health treatment." *Health Econ*, 2010: 297-315.

Perednia, DA, and A Allen. "Telemedicine technology and clinical applications." *JAMA*, 1995: 461-462.

Ralston, JD, CM Rutter, D Carrell, J Hecht, D Rubanowice, and GE Simon. "Patient use of secure electronic messaging within a shared medical record: a cross-sectional study." *J Gen Intern Med* 24, no. 3 (Mar 2009): 349-355.

Ralston, JD, D Carrell, R Reid, M Anderson, M Moran, and J Hereford. "Patient web services integrated with a shared medical record: patient use and satisfaction." *J Am Med Inform Assn* 14, no. 6 (2007): 798-806.

RedOrbit. "Few patients use or have access to online services for communicating with their doctors, but most would like to." *RedOrbit.* Sep 22, 2006. http://www.redorbit.com/news/health/667200/few_patients_use_or_have_access_to_online_services_for/ (accessed Mar 12, 2012).

Ridout, M, J Hinde, and CGB Demetrio. "A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives." *Biometrics* 57, no. 1 (Mar 2001): 219-223.

Roodman, D. "How to do xtabond2: an introduction to "difference" and "system" GMM in Stata - working paper 103." *Center for Global Development.* 12 06, 2006. http://www.cgdev.org/content/publications/detail/11619 (accessed Mar 12, 2012).

Rosen, P, and CK Kwoh. "Patient-physician e-mail: an opportunity to transform pediatric health care delivery." *Rosen P, Kwoh CK* 120, no. 4 (Oct 2007): 701-706.

Stalberg, P, M Yeh, G Ketteridge, H Delbridge, and L Delbridge. "E-mail access and improved communication between patient and surgeon." *Arch Surg* 143, no. 2 (2008): 164-169.

Staub, KE, and R Winkelmann. "Consistent estimation of zero-inflated count models." *University of Zurich.* Aug 2011. http://www.econ.uzh.ch/faculty/staub/zeroinflated.pdf (accessed Mar 22, 2012).

Tang, PC, W Black, and CY Young. "Proposed criteria for reimbursing evisits: content analysis of secure patient messages in a personal health record system." *AMIA Annu Symp Proc*, 2006: 764-768.

U. "Blood pressure and cholesterol." *Learn. Genetics. The University of Utah.* 2012. http://learn.genetics.utah.edu/content/health/history/bp/ (accessed Mar 12, 2012).

Wenger, SJ, and MC Freeman. "Estimating species occurrence, abundance, and detection probability using zero-inflated distributions." *Ecology* 89, no. 10 (Oct 2008): 2953-2959.

Winkelmann, R. *Econometric analysis of count data.* 5. Springer, 2008.

Wooldridge, JM. "Distribution-free estimation of some nonlinear panel data models." *Journal of Econometrics* 90, no. 1 (May 1999): 77-97.

Yoon, SS, Ostchega Y, and T Louis. "Recent trends in the prevalence of high blood pressure and its treatment and control, 1999–2008." *NCHS Data Brief* 48 (Oct 2010).

Zhou, YY, T Garrido, HL Chin, AM Wiesenthal, and LL Liang. "Patient access to an electronic health record with secure messaging: impact on primary care utilization." *Am J Manag Care*, 2007: 418-424.

Appendix A: Derivation of First Order Conditions for ZI QCMLE

The log-likelihood for observation *i* for the ZI QCMLE approach has the form:

$$l_i = \sum_{t=1}^{T_i} \left(y_{it} \log \widetilde{\lambda_{it}} - \widetilde{\lambda_{it}} n_i \right)$$

where

$$\widetilde{\lambda_{it}} = \frac{\frac{\lambda_{2it}}{1+\lambda_{1it}}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}}, \quad n_i = \sum_{t=1}^{T_i} y_{it}$$

$$\frac{\partial \widetilde{\lambda_{it}}}{\partial \beta} = \frac{\frac{\lambda_{2it}}{1+\lambda_{1it}}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}} x_{it} - \frac{\frac{\lambda_{2it}}{1+\lambda_{1it}}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}} \frac{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}} x_{ir}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}} = \widetilde{\lambda_{it}} x_{it} - \widetilde{\lambda_{it}} \frac{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}} x_{ir}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}}$$

$$= \widetilde{\lambda_{it}} \left(x_{it} - \frac{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}} x_{ir}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}} \right)$$

$$\frac{\partial \widetilde{\lambda_{it}}}{\partial \gamma} = -\frac{\frac{\lambda_{2it}}{1+\lambda_{1it}}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}} \frac{\lambda_{1it}}{1+\lambda_{1it}} z_{it} + \frac{\frac{\lambda_{2it}}{1+\lambda_{1it}}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}} \frac{\sum_{r=1}^{N_i} \frac{\lambda_{1ir}\lambda_{2ir}}{(1+\lambda_{1ir})^2} z_{ir}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}}$$

$$= -\widetilde{\lambda_{it}} \frac{\lambda_{1it}}{1+\lambda_{1it}} z_{it} - \widetilde{\lambda_{it}} \frac{\sum_{r=1}^{N_i} \frac{\lambda_{1ir}\lambda_{2ir}}{(1+\lambda_{1ir})^2} z_{ir}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}}$$

$$= \widetilde{\lambda_{it}} \left(-\frac{\lambda_{1it}}{1+\lambda_{1it}} z_{it} + \frac{\sum_{r=1}^{N_i} \frac{\lambda_{1ir}\lambda_{2ir}}{(1+\lambda_{1ir})^2} z_{ir}}{\sum_{r=1}^{N_i} \frac{\lambda_{2ir}}{1+\lambda_{1ir}}} \right)$$

Then the first order conditions are:

$$\frac{\partial l_i(\beta,\gamma)}{\partial\beta} = \sum_{t=1}^{T_i}\left[y_{it}\frac{1}{\widetilde{\lambda_{it}}}\frac{\partial\widetilde{\lambda_{it}}}{\partial\beta} - n_i\frac{\partial\widetilde{\lambda_{it}}}{\partial\beta}\right]$$

$$= \sum_{t=1}^{T_i}\left[y_{it}\frac{1}{\widetilde{\lambda_{it}}}\widetilde{\lambda_{it}}\left(x_{it} - \frac{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}x_{ir}}{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}}\right) - n_i\widetilde{\lambda_{it}}\left(x_{it} - \frac{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}x_{ir}}{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}}\right)\right]$$

$$= \sum_{t=1}^{T_i}\left[(y_{it} - n_i\widetilde{\lambda_{it}})\left(x_{it} - \frac{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}x_{ir}}{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}}\right)\right] = \sum_{t=1}^{T_i}x_{it}y_{it} - n_i\sum_{t=1}^{T_i}\widetilde{\lambda_{it}}x_{it}$$

$$= \sum_{t=1}^{T_i}(y_{it} - n_i\widetilde{\lambda_{it}})x_{it}$$

$$\frac{\partial l_i(\beta,\gamma)}{\partial\gamma} = \sum_{t=1}^{T_i}\left[y_{it}\frac{1}{\widetilde{\lambda_{it}}}\frac{\partial\widetilde{\lambda_{it}}}{\partial\gamma} - n_i\frac{\partial\widetilde{\lambda_{it}}}{\partial\gamma}\right]$$

$$= \sum_{t=1}^{T_i}\left[y_{it}\frac{1}{\widetilde{\lambda_{it}}}\widetilde{\lambda_{it}}\left(-\frac{\lambda_{1it}}{1+\lambda_{1it}}z_{it} + \frac{\sum_{r=1}^{N_i}\frac{\lambda_{1ir}\lambda_{2ir}}{(1+\lambda_{1ir})^2}z_{ir}}{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}}\right)\right.$$

$$\left. - n_i\widetilde{\lambda_{it}}\left(-\frac{\lambda_{1it}}{1+\lambda_{1it}}z_{it} + \frac{\sum_{r=1}^{N_i}\frac{\lambda_{1ir}\lambda_{2ir}}{(1+\lambda_{1ir})^2}z_{ir}}{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}}\right)\right]$$

$$= \sum_{t=1}^{T_i}\left[(y_{it} - n_i\widetilde{\lambda_{it}})\left(-\frac{\lambda_{1it}}{1+\lambda_{1it}}z_{it} + \frac{\sum_{r=1}^{N_i}\frac{\lambda_{1ir}\lambda_{2ir}}{(1+\lambda_{1ir})^2}z_{ir}}{\sum_{r=1}^{N_i}\frac{\lambda_{2ir}}{1+\lambda_{1ir}}}\right)\right]$$

$$= -\sum_{t=1}^{T_i}\frac{\lambda_{1it}}{1+\lambda_{1it}}z_{it}y_{it} + n_i\sum_{t=1}^{T_i}\frac{\lambda_{1it}}{1+\lambda_{1it}}\widetilde{\lambda_{it}}z_{it} = \sum_{t=1}^{T_i}(-y_{it} + n_i\widetilde{\lambda_{it}})\frac{\lambda_{1it}}{1+\lambda_{1it}}z_{it}$$

Appendix B: Stata Code for Estimation of ZI QCMLE

```
clear all
global MY_panel id

program qcmlezi
        args todo b lnf

        tempvar xb zb lamda1 lamda2 lamda last suml sumlam sum1y sumy sumql

        mleval `xb'=`b', eq(1)
        mleval `zb'=`b', eq(2)

    local y "$ML_y1"
    local i $MY_panel
        sort `i'

    quietly {

* Calculate the log-likelihood

    gen double `lamda2' = exp(`xb')
        gen double `lamda1' = exp(`zb')
    by `i': gen double `last' = _n==_N
        by `i': gen double `suml' = sum(`lamda2'/(1+`lamda1'))
        by `i': gen double `sumlam' = `suml'[_N]
        gen double `lamda' = (`lamda2'/(1+`lamda1'))/ `sumlam'
    by `i': gen double `sum1y' = sum(`y')
    by `i': gen double `sumy' = `sum1y'[_N]
    by `i': gen double `sumql' = sum(`y'*ln(`lamda') - `lamda'*`sumy')

        mlsum `lnf' = `sumql' if `last'==1

        if (`todo' == 0 | `lnf'==.) exit
    }
end

* Estimate QcmleZI model
ml model d0 qcmlezi (y = x, nocons) (y = z, nocons)
ml search
ml maximize
```

VITA

Rodica Gilles was born in Chisinau, Moldova. Currently, Rodica resides in Seattle with her husband Greg and their two children Nicole and Andrew. In 2003 Rodica earned a Bachelor of Science degree in Applied Mathematics at Moldova State University. After moving to Seattle, Rodica earned a Bachelor of Arts degree in Economics at Seattle University. Also, in 2012 she earned a Doctor of Philosophy in Economics at the University of Washington. Rodica plans to use her analytical skills to build a prolific career in the Greater Seattle area.