Dynamic Selection In Medical Insurance Markets

Thomas G. Koch*
Federal Trade Commission

March 26, 2013

Abstract

Asymmetric information is known to create missing insurance markets, while also protecting individuals from reclassification risk—the uncertainty associated with information, and thus insurance prices, changing over time. This paper measures the importance of present-day and forward looking measures of risk in the employer-provided insurance market. Using data on health, medical spending and labor market outcomes, I find that selection into firms that offer insurance, an often overlooked aspect of the insurance selection literature, is a key component to the ultimate selection into employer-provided insurance. Dynamic considerations, such as the trade-off between buying insurance today and tomorrow, are important aspects of the present-day insurance choice, leading to some advantageous selection. Mental health also appears to be a source of advantageous selection.

*The views expressed in this article are those of the author. They do not necessarily represent those of the Federal Trade Commission, and the work was not prepared as part of the author's work for the FTC. Contact: 600 Pennsylvania Avenue NW, M-8059, Washington DC, 20580. Phone: 512-809-8014. E-mail: tkoch.at.ftc.gov.

1 Introduction

What do people know, and when does it matter? The presence, and asymmetric knowledge, of information has been known to limit the ability of markets to sustain since the seminal work of Akerlof (1970) and Rothschild and Stiglitz (1976). When only one side of the market knows of static and immutable characteristics of a good, markets may fail.

However, if those characteristics may change, then that failure may be efficient, or at least have consequences for the magnitude of market failure in the present day. The creation of information can be destructive, as demonstrated in Hirshleifer (1971) and Chiappori and Gollier (2006), while the changes in insurance prices tied to changes in risk, known as reclassification risk, over time can complicate the efficiency of contracts that are efficient in static settings. (E.g., see Cochrane (1995) and Hendel and Lizzeri (2003).) When the insurance choice is connected to other markets, as is the case with employer-provided medical insurance, the best use of information must be tied to other factors. Madrian (1994) provides a case in point.

Using short panel data on health, medical spending, and labor market outcomes, I measure several important factors that are central to the relationships between present risk, future risk and insurance choice. The novel contributions of this study are two-fold. First, I measure the selection into firms that offer insurance, instead of focusing exclusively on selection into insurance conditional on working at a firm that offers it. This contribution emphasizes the role of dynamics, in that switching jobs may generate larger search or switching costs than simply switching insurance options when working at a firm that offers it.

Second, I study whether projections of future risk are important for determining present-day insurance choices. If future risk is correlated with present-day risk, and there are some frictions to the insurance choice decision, such as search costs or maximum benefit caps, then the knowledge of present risk may exacerbate, or mitigate, the amount of adverse selection due to asymmetric information.

I find that selection into firms that offer insurance is much more limited than into contracts among employees who already work at firms that offer insurance. However, in some specifications, adverse selection can be detected. I also find that future expected medical cost

is negatively correlated with present-day insurance choice, particularly among workers who work at firms that offer insurance. This suggests that, absent search and switching costs, workers make inter-temporal tradeoffs between insurance today, tomorrow, and consumption in both periods. This dynamic appears to limit advserse selection into the present-day insurance contract.

I also present evidence on the sources of this differing selection. The dynamic selection is tied to young children; among them, age predicts much higher future spending than present spending, inducing a dynamic tradeoff. Present-day adverse selection is evident with chronic conditions (high blood pressure and diabetes, e.g.), but not with the medical events often associated with them—hearts attacks and strokes. Similar to the findings of Fang, Keane, and Silverman (2008), low self-reported measures of mental health are both associated with higher medical costs and lower rates of insurance.

Recent empirical work on medical insurance markets has focused on static settings and excluded job switching: provided the assumption of static setting, what are the consequences (adverse selection leading to market failure) of asymmetric information. Instead, I measure how evolving risk characteristics effect the selection into insurance contracts, which may either exaggerate or mitigate the predictions of the static model. In particular, I focus on the employer-provided medical insurance market. This market provides a natural example for study. First, health characteristics are subject to change. Second, differences in switching cost may vary the selection in picking up insurance when working at a firm that offers it (the traditional setting for empirical work), or switching jobs to work at a firm that offers it (a novel contribution provided in this study). Finally, including dynamic elements to insurance choice implies intertemporal tradeoffs. Knowing that he or she will have a heightened need for health insurance in the future relative to day, an individual may forgo insurance today to save for insuance and associated expenses in the subsequent year. Under certain conditions, this may limit adverse selection in the present day.

Future risk impacts present-day decisions on two margins in the employer-provided market: whether to take-up insurance from an employer that offers it, and whether to work at a firm that offers it in the first place. The existing literature focuses on the former question, since much of the data available to researchers is made available from a single employer. I

find that selection into firms that offer insurance is an important component in the ultimate employer-provided insurance outcome, and can be qualitatively different than selection into insurance once employed by a firm that offers it.

Given the measured selection for both present-day and future risk, what are its sources? Finkelstein and McGarry (2006) found evidence that heterogeneity in individual preferences (i.e., risk aversion) is related to selection in the Medigap insurance market. Similarly, the results here suggest that the interaction of preferences with measures of variance is a key component of insurance demand, and dramatically varies the relationship between insurance choice and expected cost.

How does the asymmetry of information in the employer-provided insurance protect employees from reclassification risk? Reclassification risk is uncertainty about future expectations. Using measures of uncertainty related to future expected medical risk, such risk has a limited relationship between insurance choice and expected costs.

2 Related Literature

The literature on selection and insurance has taken two distinct approaches. The first, exemplified by Chiappori and Salanié (2000) and Fang, Keane, and Silverman (2008), considers the correlates of having insurance (or the quality of insurance), conditional on the information used in pricing insurance. This literature attempts to find evidence of adverse or advantageous selection, and, if possible, identify its dimensions.

This line of research has focused on a wide variety of insurance contracts: employer-provided medical insurance (e.g., Bundorf (2002), Carlin and Town (2007), and Einav, Finkelstein, and Cullen (2010)), Medigap (Fang, Keane, and Silverman (2008)), long-term care insurance (Finkelstein and McGarry (2006)) car insurance (Chiappori and Salanié (2000), Cohen and Einav (2007), and Cohen (2005)), life insurance, and annuities (Finkelstein, Poterba, and Rothschild (2009)). Here, evidence of selection is evidence of market failure, in the sense originally identified in Akerlof (1970) and Rothschild and Stiglitz (1976). The typical results of these models are complicated by the many potential dimensions of selection, including preferences (see Cutler, Finkelstein, and McGarry (2008) for analyses of

several different insurance markets).

A second literature understands that asymmetric information may act as a commitment mechanism against reclassification risk: the chance that an individual's risk type will change over time, leading to insurance-price uncertainty over time. Hendel and Lizzeri (2003) investigated the potential benefits of asymmetric information as a kind of insurance against reclassification risk in life insurance. Cochrane (1995) studies theoretical solutions to reclassification risk for medical insurance. The protection afforded by non-discriminatory employer-provided insurance is what encourages "job lock"(see Madrian (1994)). Related contributions can be found in de Garidel-Thoron (2005) and Polborn, Hoy, and Sadanand (2006).

This second literature is related to Hirshleifer (1971), which associated the heterogeneity of risk types as a risk itself, faced by agents before types are revealed. Chiappori and Gollier (2006) recast this risk as medical risk, where medical screening provides early revelations of relevant information about future health. In both of these cases, the focus is on the ex ante: how does the uncertainty of the risk that I may face impact the information regime I would choose from behind the veil of ignorance? This analysis is incomplete, because agents may not be either ex ante (completely ignorant of their risk profile) or ex post (having realized their risk). Instead, agents may be ex interim: possessing ex post knowledge about today's realization, but possessing only partial knowledge of the risks to come. The first insurance literature has looked at contract choices over time (see Cohen (2005)), but has tied changes over time to learning about hidden, permanent characteristics, instead of an evolution of hidden characteristics.

These two literatures are studying related problems, but reaching different conclusions. The first literature looks at the cross-section, and measures the selection associated with asymmetric information. The second literature looks at rules on information, and finds that selection a necessary evil to achieve a second-best result: insuring reclassification risk at the cost of present-day insurance. This disconnect was identified as a weakness in a recent review of the literature (Einav, Finkelstein, and Levin (2010)). This is the gap this paper seeks to fill.

3 A Model of Selection with Dynamic Considerations

Suppose that risk averse agents live two periods; in each period they face risk, which may vary between the two periods. Further, agents are heterogeneous in the risk they face. They may buy one-period full insurance contracts against this risk. By rule, these contracts are not priced according to individual risk type, and are offered in a competitive market. The asymmetry of information means that, at most, one contract is offered at price equal to the expected realized risk of agents who select into the contract. For the purposes of exposition, presume that such contracts do exist.

Of particular interest is the relationship between future risk and present-day insurance choice, and its implications for the relationship between present-day risk and insurance choice for the same. There are many possible sources of this interaction. For example, there may be search or switching costs associated with taking up insurance tomorrow, but not today. This would drive a positive relationship between future cost and present insurance. Alternatively, agents may have marginal utility functions that vary with their level of medical need—e.g., if high-need times require large co-pays, or non-medical expenses, then they would have a reason to save present-day premiums for consumption needs tomorrow. That is, to substitute present-day insurance for future insurance. Alternatively, insurance policies may have maximum payout clauses, which limit the total coverage of a plan. If medical needs are known to be more necessary in the future, agents may prefer to save the balance against the maximum for use tomorrow, if medical services today have more substitutes.

All of these stories suggest a dynamic link between future risk and present-day insurance choice. In any of these stories, you can consider a shadow value of future insurance on present-day insurance choice. In the search/option value story, this is positive—choosing insurance today makes insurance tomorrow less expensive. Alternatively, it may be negative, when there are maximum payout clauses or oppotunities for intertemporal substitution. What matters empirically for determining selection into insurance today is how that shadow value is correlated with present-day expected cost.

If this shadow value is negatively correlated with expected cost, then it can contribute to advantageous selection. (Or, at least, mitigate the amount of adverse selection that would

be absent this shadow value.) This corresponds to the relationship between cognitive ability, medical expenditures and Medigap insurance take-up in Fang, Keane, and Silverman (2008). To the extent that these three things are positively correlated, they can lead to advantageous selection.

4 Data and Empirical Framework

Severe demands are placed on the data in order to evaluate the testable implications devised above. First, the data must be panel, in order to calculate future projections from current data. Second, the data must include information on medical spending, as well as relevant determinants of medical spending, such as age, gender, and diagnosed medical conditions. Finally, the data must include information about the employer offerings at a firm: does the firm offer insurance, and, if so, is it taken up. The literature so far has focused on firms that offer insurance, and tests for selection within that firm. Little has been done on the selection into those firms.

The Medical Expenditure Panel Survey (MEPS) provides such data. The MEPS is a nationally representative survey of American households. It first sampled households in 1996, and has collected data on two-year overlapping panels ever since. A representative of the household is interviewed five times over those two years.

The primary focus of the MEPS is medical spending, by source (e.g., Medicare, private insurer, out-of-pocket) and use (ER, outpatient care, etc.). While data are available on the health-care incident level (e.g., an ER or office-based visit), this paper focuses on the annual medical spending provided in the yearly consolidated files.

In each interview, information is collected on the insurance status and source of individuals in the sample. This information is available at the point of time of the interview. Here, I use this point-in-time, job-level data to be sure to connect the particular job to the insurance availability and choice. If more aggregated measures were used, there may be an incorrect linkage between an insurance plan and an employer, if a household had multiple employers during the sample.

Because insurance is offered for both individuals and families, entire households are included in the sample. Because those offers are often tied to employment within the household, the MEPS includes information on the jobs held by individuals in the households under study (occupation, whether insurance is offered to the particular employee), as well as characteristics of those employers (industry, firm size, whether insurance is offered to any employee).

This last set of information is key to this study. In order to test for selection, the characteristics that insurers use to determine price must be included in set of control variables. Employer-provided medical insurance contracts are group insurance contracts, with strong restrictions on the use of demographic and health information to determine individual prices. While insurance contracts may not discriminate within a firm, they do discriminate across firms. See, for example, Crimmel (2009), which describes patterns of insurance price for firms by firm size.

The focus on this paper is the selection into employer-provided insurance, so we drop those individuals sixty-five years and older for whom Medicare is available. The specifications described below further control for Medicaid and/or SCHIP coverage. Due to changes in survey design, all of the variables described above and enumerated below are not available in every year of the MEPS. Because of this, I use the respondents who entered the survey between 2001 to 2008. For most specifications, this yields a sample size of 72,840. However, when including the measures of risk preference, the sample drops to 64,349.

4.1 Measures of Disease and Health

The MEPS identifies a select group of priority conditions (PC), and specifically asks about prior diagnoses of the PCs. The priority conditions are: sore throat (for children under age 18), diabetes, asthma, high blood pressure, high cholesterol, heart disease (including coronary heart disease, angina, myocardial infarction), stroke, emphysema, joint pain and arthritis.

The MEPS asks the respondents to rate their own health, with the five traditional categorical answers: "Excellent," "Very good," "Good," "Fair," and "Poor." There is a similar

question specifically regarding mental health status, with the same five categorical answers. The height and weight of the survey respondents are solicited; the corresponding BMI was provided in the public release of the data and its log is included in the medical spending equations described below. Finally, the MEPS asks about health-related lifestyle choices. Answers about smoking will be included as predictors of medical risk.

4.2 Attitudes Towards risk

Starting in 2001, the MEPS asked respondents whether or not they were more likely to take risks than the average person. Respondents could choose from five answers: "disagree strongly," "disagree somewhat," "uncertain," "agree somewhat," and "strongly agree." This question was asked in the Self-Administered Questionnaire (SAQ), a paper-and-pencil portion of the survey that collects information about health status and health care quality. It is the same question used in the 1987 National Medical Expenditure Survey.

These answers will stand in for measures of risk aversion. Unfortunately, the MEPS does not ask questions that could be used to precisely calculate individual parameters of risk aversion, as in Fang, Keane, and Silverman (2008); nor is there a level of detail in the insurance contract choice to make a similar inference, as in Cohen and Einav (2007).

There are two potential problems with using answers to this question as proxies for risk aversion: (1) do respondents think this question asks about attitudes towards financial risk, instead of, for example, personal safety; and, (2) does the order from "strongly against" to "strongly for" actually correspond to more refined differences in preferences.

The context of the survey question should allay concern about this first concern. After a long battery of questions about recent health, the SAQ asks questions specifically about health insurance. The first asks the respondent if the he or she needs health insurance. The second, which is the question just prior to the question about attitudes towards risk, asks specifically if he or she thinks health insurance is worth its cost. Because the prior questions specifically ask about insurance and financial trade-offs, the proposed interpretation seems reasonable.

This second concern can be addressed by using separate dummy variables for each an-

swer. This frees estimates from any parametric assumption about intensity. Answers to the question are not evenly distributed across answers. Thirty-eight percent of the 147,718 respondents under age 65 said they strongly disagreed with the statement that they were more likely to take risks than average. Another twenty-three percent said they somewhat disagreed. Thus, over sixty percent say that they are less risk averse than average. The implied skewness of risk preferences is qualitatively consistent with the lognormal distributions of risk aversion estimated in Cohen and Einav (2007).

4.3 Job and Employer Characteristics

The MEPS asks a series of questions about the jobs of all employed members of the household. Standard questions about wages and hours are used to construct monthly earnings. The respondent is asked if he or she has insurance from this job. If not, the respondent is asked if insurance is offered at this job.

They key outcome of concern is whether or not the employee had employer-provided insurance. There may be differences in selection into firms that offer insurance, and selection of taking up insurance once employed by a firm that offers it. Because of this, the empirical analysis that follows will break down the ultimate outcome (does a worker have employer-provided insurance?) into its component parts (does the firm offer it?; if so, is it taken up?)

The survey also asks about characteristics of the firm, such as the size of the firm (later aggregated into eight categories), and whether the firm has more than one location. In order to control for the strong geographic patterns of health care costs, I also include dummies for the four census regions, as well as interactions of those dummies with MSA status. (State level identifiers are not included in the publicly available data due to confidentiality concerns.) There is also an indicator if the job is unionized, and twelve indicators each for the firm's industry, and the occupation of the job.

5 Sources of Selection

The data described above allow me to investigate the potential sources of selection. In order to qualify as such, sources must be private (i.e., unpriced) variables that influence insurance demand and are correlated with health expenditure risk.

5.1 Empirical Strategy

This strategy follows Fang, Keane, and Silverman (2008), with two main advantages. First, the MEPS contains all of the data required to implement this strategy. Second, the short panel allows for the projection of next year's outcome on this year's covariates.

I separate the covariates into three categories: demographics, D_i, medical conditions and diagnoses, C_i, and firm characteristics F_i. The medical conditions and diagnoses include indicator variables for all of the diagnosed conditions and self-reported statuses described above. The demographic variables used in the regression are: a fifth-order polynomial in age in months, gender, race, and Hispanic ethnicity.

I adopt Fang, Keane, and Silverman (2008)'s first of two strategies to cope with the endogeneity of medical insurance. The first estimates the prediction equations only on those without employer-provided insurance.[1]

$$\widehat{mx}_{1i} = D_i\hat{\delta}_1 + C_i\hat{\gamma}_1 \quad (1)$$

$$\widehat{mx}^2_{1i} = D_i\hat{\delta}_2 + C_i\hat{\gamma}_2 \quad (2)$$

The predicted variance for each employee is calculated via the rearrangement $\widehat{VAR}_i = \widehat{mx^2}_i - \widehat{mx}_i^2$. Imputed means and variances for those with insurance are constructed using individual demographics, condition profiles and the estimated coefficients of Equations 1 and 2. This is done again with year 2 medical spending for those without medical insurance in that year on present conditions and demographics.

As discussed in (Fang, Keane, and Silverman, 2008), this approach is not ideal. It may be biased if those with employer-provided insurance are healthier in ways not included in the health and condition profile, C_i. If those with insurance are unobservably less healthy

[1] In theory, one might exclude those with public

than those with it, this will understate the medical risk faced by those with insurance.

Despite these conceptual concerns, it is important to note the robust amount of information about health and spending in the MEPS. The imputations described above include a broad set of objective diagnostic measures, as well as qualitative measures of physical and mental health.

6 Sources of Selection

What are the sources of selection into employer-provided medical insurance? First, we must measure the amount of selection, and then decompose it into its component parts. Consider the following equation:

$$\iota_i = \widehat{mx}_i\theta + F_i\lambda_1 + \tilde{u}_i. \tag{3}$$

where ι_i is individual i's insurance outcome; $\widehat{mx}_i$ is his or her expected medical cost; and F_i is a set of firm characteristics of individual i's employer. θ measures the amount of selection. If $\theta > 0$, then insured individuals are higher risk than the uninsured, and there is adverse selection. Alternatively, if $\theta < 0$, the insured are lesser risks than the insured, and there is advantageous selection.

The nature of employer-provided insurance complicates this standard interpretation of θ. First, employer-provided insurance may cover an entire family, not just individuals. To account for this, the expected medical costs (and variances) are added up across individuals within the household to calculate the expected medical costs (and variances) for the insured risk.

Second, θ may measure selection across contracts, instead of into a particular contract, if it is correlated with specific features of the contract. This is an issue with what we would like θ to measure. In the best examples of the empirical literature mentioned above, all price and non-price features of the insurance contract are controlled for in the analogues of Equation 3. With that, θ measures the amount of unpriced selection occurring in the insurance market, ostensibly due to asymmetric information. This is important, because it

is this kind of unpriced selection into insurance contracts that can lead to adverse selection, and the partial or complete collapse of insurance markets.

Unfortunately, the data employed here do not have such fine-grained detail. Thus, differences in the expected costs here may in fact be priced, if there is sorting of workers into firms by expected costs, and these differences are reflected in the price of insurance paid by employers and employees. We may cleanly measure differences in medical cost between the insured and uninsured, but this may not be a precise measure of the importance of asymmetric information in uninsurance.

Table 1 provides the point estimates for θ in Equation 3. Three panels report this coefficient for three different insurance outcomes: having employer provided insurance; having employer provided insurance conditional on working at a firm that offers it; and working at a firm that offers it. The key variable of interest, $\widehat{mx}_i$ is denominated in tens of thousands of dollars. Thus, we interpret the first coefficient reported in Table 1, 0.00556, in the following manner: for every ten thousand dollars increase in expected current-year medical cost, an employee is one half of a percentage point more likely to have employer-provided medical insurance. These regressions only control for the firm characteristics described above. Looking down the first column in each of the three sets of rows, we find that most of this relationship comes from the selection into firms that offer insurance. Put another way, most of the selection into the employer-provided insurance appears to occur while in a job, not while looking for one.

This result does not imply that there is no sorting by risk type across workplaces, nor it is exhaustive of all of the risk sorting that may take place within a workplace that offers insurance. There may be sorting across insurance contracts offered in the same workplace, as documented by Cutler and Reber (1998). More refined data on the contract options and prices is not available in the MEPS used here, so we can only address the selection into having any employer-provided insurance or declining it, even though it is being offered.

Other determinants of demand, such as the variance of risk, are included in the error term, u_i, and may be correlated with expected cost. These other determinants of demand can be added incrementally in the vector O_i. We can then estimate the following equation.

$$\iota_i = \widehat{mx}_i\theta_1 + O_i\omega + F_i\lambda_1 + \tilde{u}_i. \tag{4}$$

The difference between θ and θ_1 is the amount of unconditional selection that can be accounted for by the other factors, O_i. Following Fang, Keane, and Silverman (2008), we incrementally adjust O_i, to determine which portions of θ can be accounted for by each factor of demand in O_i. This is done for each of the three insurance choices of interest: having employer-provided insurance; working at a firm that offers it; and taking it up, condition on it being offered.

The relationship between the various individual demand factors in O_i and their corresponding $\hat{\theta}_1$s is not just a matter of accounting for the unconditional selection captured by θ. As described by Einav and Finkelstein (2011), the relative shape of the demand curve, the marginal and average cost curves determine the welfare consequences of asymmetric information.

For example, if agents varied in their expected cost, but had the same constant preferences and variance, then the demand curve would (up to the Arrow-Pratt approximation) lie parallel and above the marginal cost curve. Alternatively, if variance and preferences decreased with expected cost, then a decreasing demand curve could (depending on the magnitudes of the varying factors) correspond to an increasing marginal cost curve. The correlation between expected cost and other demand factors is important for understanding the general shapes of the demand and marginal cost curves, which are important for understanding the impact of policy interventions, such as tax subsidies or mandates for insurance.

Table 1 presents the estimates for θ_1 for different controls in O_i. The first addition of interest is the expected value of next year's medical spending. With this addition, the coefficient on present-year medical spending is nearly three times larger than when not controlling for next year's expected cost. The negative correlation between present-day insurance and future spending, as well as the positive correlation between present-day and future spending, leads to an increase in the present-day spending coefficient.

The coefficient for next-year's medical spending is negative and similar in magnitude to the coefficient for this year's medical spending. Holding present-day spending constant, fu-

ture medical spending makes insurance today less likely. This suggests strong inter temporal trade offs for employees making health insurance choices. This suggestion is reinforced by the fact that the negative relationship between future spending and present-day choice is drawn from workers at firms that offer insurance (Column 2 of the second group). This inter temporal trade off is more salient when there are smaller switching costs. When switching costs are large, as is the case when switching jobs, this trade off is not apparent (Column 2 of the third group).

The third column of each group adds the controls for risk preferences, while the fourth column interacts the risk preference dummies with the imputed variances of medical risk. Unlike Fang, Keane, and Silverman (2008), including risk preferences and imputed variance makes the coefficient on present-day expected cost less positive. This is because the imputed variance of medical spending is positively correlated with insurance and positively correlated with expected cost. This pattern is consistent with standard expected utility theory.

The magnitudes of θ_1 are particularly responsive to the interaction of the imputed variances with risk preferences. The statistical significance of the expected cost coefficients for the overall insurance choice diminishes, though is marginal for take-up. The coefficients themselves are attentuated, but the standard errors double in magnitude. This suggests that the selection into employer-provided insurance contracts is not simply a matter of sorting according to high- vs. low expected cost employees. Rather, it depends upon many factors of demand, including expected cost, the variance of risk, risk preferences, and how the three are correlated in the population.

The fifth columns include the variance of the fitted value and the variance of the forecast for the next year's expected medical cost. These controls proxy for reclassification risk: how variable or uncertain is an employees future expected cost. These variances are also interacted with the risk preference dummies. Including these regressors does increase the point estimate of θ_1.

With these additional controls, the coefficients on expected medical spending for the take-up regressions are further attenuatead. This is reflected in the overall insurance choice, while these additional controls have a mixed effect on the correlation between a worker's risk profile and whether he or she works at a job that offers insurance. This suggests, unsurprisingly,

that demand for the option to buy insurance is largely tied up in the non-expected cost portion of the demand for health care. Once those other drivers of demand are accounted for, present expected cost becomes more salient for the present-day insurance choice.

Across the entire population, selection appears to be dominated by the within-firm choices: whether to take up insurance or not. The coefficients on the decision to work at a firm that offers insurance are all an order of magnitude smaller than those for the choice to take it up, conditional on it being offered, and not statistically significant with one exception.

6.1 Subpopulations

The results thus far glossed over one particular reality of the employer-provided insurance market: the decision is made at the household level, and the household might have multiple employers offering insurance. The former complication was addressed by adding up the individual-level risk components within the household. To address the second, we repeat the above regressions for households with only one employed member.

The results for this subpopulation are reported in Panel A of Table 2. This table also repeats the results from Table 1 to aid in comparison. These are reported in Panel A.

In both of these populations, the correlation between employer-provided insurance and expected medical spending is larger than in the general population. When the controls for future medical risk are added, much of the correlation between present-day expected spending and present-day choice are attentuated, while the future medical spending coefficients are both large and statistically significant.

In one-job households, these patterns vary according to the level of the decision. For the decision with smaller switching costs (taking up insurance), the correlation is strongest between present risk and take-up; future expected medical spending is much less correlated without any additional controls. This pattern withstands the addition of the dummy variables measuring risk preferences. As in the general population, the statistical significance of the correlations diminishes when the measures of variance and reclassification risk are added to the set of controls, suggesting the importance of these characteristics when it comes to

taking up insurance.

The correlations are different in an expected way when looking the outcome is working at a firm that offers insurance. Here, the future medical spending is strongly correlated with working at a firm that offers insurance; future medical risk should play a larger role when decisions are costlier to change.

This appears to be due to present-day's expected medical spending's correlation with other factors of demand, such as risk preferences and variance. When controls for these factors are included, this negative correlation dissipates. Future expected cost is also correlated with these other factors, but the magnitude of the correlation persists more strongly for future risk than present risk when these additional controls are added.

Similar patterns appear, though less strikingly, when looking at one-person households. The sample shrinks so the standard errors grow accordingly, though the magnitude of the coefficients is also smaller. Where there does appear to be correlation between expected medical risk and insurance options is in working at a firm that offers insurance, particuarly for future expected cost.

These results demonstrate the difficulty of assessing the correlation between risk and insurance, and its sources, when households have options. When households have more limited options, the correlations are more likely to appear as we would traditionally expect: present-risk dominates the easily switched decisions, while future risk plays a large role in the decisions that are costlier to change.

7 Potential Mechanisms

Given the novelty of the results above, it is useful to describe potential mechanisms, or channels, that are driving them. For a channel to be consistent with the findings so far, a demographic or health characteristic (condition or some other measurement) needs to be positively correlated with medical spending risk in the present day, but moreso in the future. For there to be advantageous selection, a characteristic must be negatively correlated with insurance choice in the present day. These correlations need to be true even when controlling for the firm-level characteristics that proxy for insurance price.

The dynamic selection could be due to the conditions themselves. Table 3 presents the point estimates for such correlations: regressing present-day insurance choice on the health and demographic information used to construct the present and future expected medical spending imputations. Also reported are the point estimates from those imputations, which are based on the uninsured. This is an imperfect measure of the potential channels, as the characteristics used in the regression are only those of the worker, and not summary measures for the entire family.

Chronic conditions, such as high blood pressure, asthma and diabetes, are positively correlated with having employer-provided insurance and larger expected medical expenditures. These conditions would ostensibly drive much of the positive correlation between expected medical cost and present-day insurance choice. Their associations with future medical spending are also larger than they are for the present-day.

The most striking exception to this is angina, with little increases for present-spending but future increases on the order of $4,000. Conditional on all of the other health and demographic information, there is not a statistically significant correlation with present-day health insurance. Health events, like myocardial infarction (heart attack) or strokes, have similarly sized increases in present and future spending. These events also tend to be either not correlated, or negatively correlated with present-day insurance choice. Because the exact timing of these events can be thought of as unanticipated, it should not be surprising that they are not strongly correlated with insurance choice, as the chronic conditions are. The negative correlation may be due to differences in access to care tied to insurance, and the deterioration of health associated with a lack of insurance.

There is suggestive evidence of this in the softer health measurements. The self-reported overall health and mental health measures do appear to be a way to characterize some advantageous selection, and may help explain the correlation between insurance and the events. Self-reported health measures are included as indicator variables, with the highest reports ("excellent") as the excluded category. Thus, the relative size of point estimates for the self-reported indicators informs the relationship between health, spending and insurance choice. For the self-reported (overall) health, there is evidence of advantageous selection: those who rate their health status as fair or poor are less likely to have employer-provided

insurance. They are also expected to have increased medical costs, though the predicted future costs are similar to the predicted present costs.

These relationship between self-reported health and insurance is strongest for the mental health measure, especially for the lowest level, poor. All other coefficients for the mental health measures on spending are limited in thei magnitude, but, holding other things constant, individuals with poor mental health cost $1,700 more in medical spending on the present year, and $6,500 more in the next year. All of the other coefficients are $200 or less. This dramatic difference in spending occurs in spite of a strongly negative correlation between poor mental health and insurance choice. Those with poor mental health are eight percentage points less likely to have employer-provided medical insurance than those with excellent medical health; the difference between good and excellent mental health is just two percent. There is a similar pattern for self-reported general health, with a strong, though less dramatic, difference at the lowest levels of self-reported health. Though this is a more loosely defined measure of mental acuity than in Fang, Keane, and Silverman (2008), this result is suggestively similar.

8 Conclusion

In a static setting, information symmetry is essential for insurance markets to function. When that information changes over time, the benefits of a series of well-functioning, fairly-priced (spot) markets may be diminished by its adverse consequences. The tension and trade offs between the static and dynamic considerations has been under-studied in the recent literature.

This paper assesses the importance of risk dynamics and preferences in the employer-provided insurance markets. Employer-provided insurance markets are a natural setting for this kind of investigation, as non-discrimination rules limit the use of information in this market, while the costs of job search induce option value for future contracts.

Following Fang, Keane, and Silverman (2008), I decompose the sources of selection into employer-provided insurance. I find that future expected medical spending is negatively correlated with present-day insurance choice. Because present and future expectations are

positively correlated, the dynamic aspects of the present-day insurance choice mitigate the consequences of asymmetric information. This effect is the most noticeable when the insurance choices involve the fewest switching costs—i.e., looking at the take-up choices of workers at firms that offer medical insurance. Other determinants of demand, such as risk preferences and the variance of medical risk, also mitigate the relationship between expected medical cost and present-day medical insurance choices.

I also investigate the selection into firms that offer employer-provided medical insurance, instead of focusing only on selection into insurance among those already employed at a firm that offers medical insurance. I find that selection into such firms is an important component of the final selection into employer-provided insurance at all. Further, the dynamic aspects of selection (future medical spending versus present-day insurance choice) may be qualitatively different within firms that offer insurance vs. across firms that do or do not offer it.

These patterns are robust across a variety of subgroups, and controlling for different aspects of firm heterogeneity tied to employer-provided insurance. Further, the potential mechanisms of selection investigated, such as young children, chronic conditions and mental health, reinforce the intuition of the results.

References

Akerlof, G. A. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 84(3), 488–500.

Bundorf, M. K. (2002): "Employee demand for health insurance and employer health plan choices," *Journal of Health Economics*, 21(1), 65–88.

Carlin, C., and R. Town (2007): "Adverse selection, welfare and optimal pricing of employer-sponsored health plans," *U. Minnesota Working Paper*.

Chiappori, P., and B. Salanié (2000): "Testing for asymmetric information in insurance markets," *Journal of Political Economy*, pp. 56–78.

Chiappori, P.-A., and C. Gollier (eds.) (2006): *Competitive Failures in Insurance Markets: Theory and Policy Implications*chap. 3. MIT Press.

Cochrane, J. H. (1995): "Time-Consistent Health Insurance," *Journal of Political Economy*, 103(3), 445–73.

Cohen, A. (2005): "Asymmetric information and learning: Evidence from the automobile insurance market," *Review of Economics and Statistics*, 87(2), 197–207.

Cohen, A., and L. Einav (2007): "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, 97(3), 745–788.

Crimmel, B. L. (2009): "Premiums, Employer Costs, and Employee Contributions for Private Sector Employer-Sponsored Health Insurance, Single Coverage by Firm Size, 19962006," Statistical Brief 231, Agency for Healthcare Research and Quality.

Cutler, D. M., A. Finkelstein, and K. McGarry (2008): "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance," *American Economic Review*, 98(2), 157–62.

Cutler, D. M., and S. J. Reber (1998): "Paying For Health Insurance: The Trade-Off Between Competition And Adverse Selection," *The Quarterly Journal of Economics*, 113(2), 433–466.

de Garidel-Thoron, T. (2005): "Welfare-Improving Asymmetric Information in Dynamic Insurance Markets," *Journal of Political Economy*, 113(1), 121–150.

Einav, L., and A. Finkelstein (2011): "Selection in Insurance Markets: Theory and Empirics in Pictures," Working Paper 16723, National Bureau of Economic Research.

Einav, L., A. Finkelstein, and M. Cullen (2010): "Estimating Welfare in Insurance Markets Using Variation in Prices," *Quarterly Journal of Economics*, 125(3), 877–921.

Einav, L., A. Finkelstein, and J. Levin (2010): "Beyond Testing: Empirical Models of Insurance Markets," *Annual Review of Economics*, 2, 311–336.

Fang, H., M. Keane, and D. Silverman (2008): "Advantageous selection in the Medigap insurance market," *Journal of Political Economy*, 116(2), 303–350.

Finkelstein, A., and K. McGarry (2006): "Multiple dimensions of private information: evidence from the long-term care insurance market," *American Economic Review*, 96(4), 938–958.

Finkelstein, A., J. Poterba, and C. Rothschild (2009): "Redistribution by insurance market regulation: Analyzing a ban on gender-based retirement annuities," *Journal of Financial Economics*, 91(1), 38–58.

Hendel, I., and A. Lizzeri (2003): "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance," *Quarterly Journal of Economics*, 118(1), 299–327.

Hirshleifer, J. (1971): "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61(4), 561–574.

Madrian, B. (1994): "Employment-Based Health Insurance and Job Mobility: Is there Evidence of Job-Lock?," *Quarterly Journal of Economics*, 109(1), 27–54.

Polborn, M. K., M. Hoy, and A. Sadanand (2006): "Advantageous Effects of Regulatory Adverse Selection in the Life Insurance Market," *Economic Journal*, 116(508), 327–354.

Rothschild, M., and J. E. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 90(4), 630–49.

Table 1: Sources of Selection

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Has Employer-Provided HI				
Present $\widehat{mx}_i$	0.00973***	0.0277***	0.0301***	0.0235	-0.0343
	(0.00361)	(0.00881)	(0.00937)	(0.0202)	(0.0297)
Future $\widehat{mx}_i$		-0.0239**	-0.0268**	-0.0261	0.0730***
		(0.0106)	(0.0113)	(0.0223)	(0.0272)
Observations	72,840	72,840	64,349	64,349	64,349
R-squared	0.229	0.230	0.227	0.227	0.231
	Has Employer-Provided HI — Firm Offers				
Present $\widehat{mx}_i$	0.0130***	0.0445***	0.0455***	0.0412*	0.00273
	(0.00390)	(0.00903)	(0.00949)	(0.0216)	(0.0316)
Future $\widehat{mx}_i$		-0.0416***	-0.0416***	-0.0464**	0.0283
		(0.0110)	(0.0115)	(0.0237)	(0.0288)
Observations	49,982	49,982	44,905	44,905	44,905
R-squared	0.057	0.058	0.057	0.057	0.061
	Firm Offers HI				
Present $\widehat{mx}_i$	0.00262	-0.000305	0.00149	0.00194	-0.0254
	(0.00337)	(0.00808)	(0.00849)	(0.0183)	(0.0272)
Future $\widehat{mx}_i$		0.00390	-0.000920	-0.00244	0.0554**
		(0.00973)	(0.0102)	(0.0201)	(0.0248)
Observations	72,840	72,840	64,349	64,349	64,349
R-squared	0.258	0.258	0.258	0.258	0.259
Preferences			X	X	X
$\widehat{Var}_i$				X	X
Future $Var(\widehat{mx}_i)$					X

Note: Estimates for linear probability models, with robust standard errors in parentheses. This table reports the point estimate of expected medical spending, in tens of thousands of dollars. Each column represents a different set of controls for a regression, while each panel corresponds to a different dependent variable. Specifications with variances also include the interaction of those variances with preferences.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2: Subpopulation Estimates

	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
	Has Employer-Provided HI					Has Employer-Provided HI				
Present $\widehat{mx}_i$	0.0298***	-0.00576	-0.00149	-0.0633	0.0166	0.0444***	-0.0481	-0.0205	-0.0362	0.0153
	(0.00627)	(0.0143)	(0.0154)	(0.0391)	(0.0568)	(0.0156)	(0.0350)	(0.0362)	(0.0788)	(0.109)
Future $\widehat{mx}_i$		0.0505***	0.0503***	0.126***	0.0538		0.120***	0.0859**	0.115	0.0703
		(0.0183)	(0.0194)	(0.0434)	(0.0515)		(0.0410)	(0.0423)	(0.0842)	(0.0986)
Observations	19,914	19,914	18,358	18,358	18,358	7,526	7,526	7,116	7,116	7,116
	Has Employer-Provided HI — Firm Offers					Has Employer-Provided HI — Firm Offers				
Present $\widehat{mx}_i$	0.0297***	0.0353***	0.0404***	-0.0304	0.0495	0.0253*	0.0148	0.0251	0.0703	0.167*
	(0.00601)	(0.0125)	(0.0133)	(0.0376)	(0.0540)	(0.0138)	(0.0254)	(0.0264)	(0.0686)	(0.0937)
Future $\widehat{mx}_i$		-0.00771	-0.00920	0.0693*	0.00219		0.0133	-0.00780	-0.0537	-0.116
		(0.0167)	(0.0176)	(0.0417)	(0.0494)		(0.0324)	(0.0333)	(0.0739)	(0.0855)
Observations	14,079	14,079	13,060	13,060	13,060	5,844	5,844	5,537	5,537	5,537
	Firm Offers HI					Firm Offers HI				
Present $\widehat{mx}_i$	0.0142**	-0.0353***	-0.0324**	-0.0267	0.0120	0.0322**	-0.0657**	-0.0412	-0.0673	-0.102
	(0.00599)	(0.0136)	(0.0146)	(0.0364)	(0.0536)	(0.0145)	(0.0334)	(0.0340)	(0.0696)	(0.0993)
Future $\widehat{mx}_i$		0.0703***	0.0675***	0.0693*	0.0295		0.127***	0.105***	0.146*	0.154*
		(0.0173)	(0.0183)	(0.0404)	(0.0488)		(0.0380)	(0.0389)	(0.0751)	(0.0894)
Observations	19,914	19,914	18,358	18,358	18,358	7,526	7,526	7,116	7,116	7,116
Preferences			X	X	X			X	X	X
$\widehat{Var}_i$				X	X				X	X
Future $Var(\widehat{mx}_i)$					X					X

Note: Estimates for linear probability models, with robust standard errors in parentheses. This table reports the point estimate of expected medical spending, in tens of thousands of dollars. Each column represents a different set of controls for a regression, while each panel corresponds to a different dependent variable and subpopulation. Specifications with variances also include the interaction of those variances with preferences.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Sources of Selection: Health Conditions and Status

	(1) Insurance	(2) Present Spending	(3) Future Spending
High blood pressure	0.0180***	599.9***	551.2***
	(0.00489)	(81.81)	(118.8)
Angina	0.0144	787.1***	2,804***
	(0.0209)	(295.4)	(444.4)
Miocardial Infarction	-0.00959	2,038***	820.6**
	(0.0179)	(259.0)	(389.3)
Other heart disease	-0.0175**	1,034***	1,414***
	(0.00892)	(163.2)	(242.5)
Stroke	-0.0413**	1,323***	1,263***
	(0.0199)	(279.6)	(442.6)
Emphysema	-0.00172	2,553***	1,649***
	(0.0241)	(302.5)	(441.1)
Arthritis	-0.00106	533.8***	718.0***
	(0.00539)	(93.75)	(136.0)
Daily Limitations	0.0158***	1,358***	683.8***
	(0.00386)	(63.43)	(89.30)
Smoker	-0.0300***	-261.3***	-212.7**
	(0.00418)	(59.76)	(87.03)
Asthma	0.0105*	-25.69	-40.92
	(0.00607)	(108.1)	(154.4)
Diabetic	0.0149*	1,629***	1,558***
	(0.00832)	(124.7)	(183.4)
Self-reported Health			
Very good	0.0107**	109.5	-62.89
	(0.00482)	(80.36)	(116.2)
Good	-0.00537	214.8**	41.70
	(0.00547)	(84.19)	(119.8)
Fair	-0.0274***	320.7***	360.0**
	(0.00786)	(107.6)	(153.4)
Poor	-0.0325**	2,174***	590.5**
	(0.0156)	(176.4)	(241.7)
Self-reported mental health			
Very good	-0.0112**	-109.4	-9.377
	(0.00447)	(72.80)	(104.8)
Good	-0.0171***	-143.3*	-131.6
	(0.00523)	(77.41)	(110.0)
Fair	-0.0587***	82.83	32.67
	(0.0101)	(127.6)	(184.5)
Poor	-0.0845***	1,107***	1,438***
	(0.0266)	(254.0)	(354.5)
Observations	64,327	24,958	9,453
Demographic Controls	X	X	X
Firm Controls	X		
Exclude insured		X	X

Note: Estimates for linear probability models (1) and linear regressions (2 and 3), with robust standard errors in parentheses. The first column reports the relationship between health characteristics and present-day insurance outcomes. The second and third columns report the point estimates from the expected medical spending imputation described in the text.
* significant at 10%; ** significant at 5%; *** significant at 1%.