Economic Research Initiative on the Uninsured
CONFERENCE DRAFT

HEALTH INSURANCE, EXPECTATIONS, AND JOB TURNOVER

Draft: Please do not cite or quote without permission.

Prof. Randall P. Ellis
Department of Economics
Boston University
270 Bay State Road;
Boston MA 02215 617-353-2741
ellisrp@bu.edu.

Albert C. T. Ma.

Economic Research Initiative on the Uninsured
University of Michigan
555 South Forest Street, 3rd Floor
Ann Arbor, MI 49104-2531

July, 2004

Acknowledgments

We are grateful for support for this study by the Robert Wood Johnson Foundation's Economic Research Initiative on the Uninsured, funded through the University of Michigan. We are grateful for MEDSTAT for facilitating the use their data and DxCG Inc. for donating programming time to create the analytical files. We would like to thank Pooja Gupta for her programming assistance and Denise Doiron, Jennifer Foster and Elizabeth Savage for helpful comments. Any errors or omissions remain our own.

Introduction

This paper develops a model in which small private employers rationally decide not to offer health insurance to their employees. Our hypothesis is that employers avoid offering health insurance when this is useful as a selection tool for attracting relatively healthy, and more profitable employees. While this selection technique is not attractive to all firms, it is sufficiently attractive for many small firms because of their relatively high variability in expected employee health care costs. It is also relatively attractive for firms with low job specific human capital where high turnover rates are preferred to paying higher wages. The decision not to offer health insurance arises because employers know that young and healthy workers tend to be more mobile and less influenced by whether a firm offers health insurance than older, less healthy workers. High administrative costs for small firms exacerbate the selection problem.

Consider an employer such as a restaurant or retail store, which does not currently offer health insurance. In such firms with relatively low wages and low firm-specific human capital workers, job turnover rates tend to be high. While the existing group of employees may be very healthy and hence appear to be inexpensive to insure, both the employers and insurers know that once workers are offered health insurance, then the average health costs of the firm's employees will tend to worsen. Because insurers are expected to guarantee renewals, and commit to premiums three years in advance, do not price their policies at the current expected costs. Rather, in anticipation of adverse selection, insurers only offer insurance coverage to high turnover firms at premiums that are significantly above their expected costs. In turn, employers where insurance is only valued marginally are discouraged by these high premiums from offering health insurance at all; reinforcing insurers need to maintain high premiums.

Our theoretical model demonstrates that it is not the absolute level of job turnover that matters for the selection problem of interest here, but rather it is the relative turnover rates of high versus low health cost workers. We develop a stylized model of employee selection in which all workers are equally productive, and yet some employers may find it unprofitable to offer health insurance because it lowers their overall labor costs. Their relatively healthy workers undervalue such insurance relative to the premium charged. The model we develop is relevant for both large and small firms, and does not rely on small firm size or an absence of sufficient pooling of risks for the selection problem to emerge. However small firm sizes exacerbate the selection problem because they are better able to use their private knowledge of expected health care costs of their own employees to choose whether or not to insure. A dynamic adverse selection problem emerges in which employers with favorable health risks are reluctant to offer insurance at the offered premium because it will change the average health mix of the employees they attract.

Our stylized model generates several simple empirical hypotheses about the insurance offer decision. Firms not choosing to offer insurance need not have low variability of employee health risks within the firm (as proposed by Bundorf (2002) but rather low average expected health spending, which we operationalize as higher proportions of younger workers or workers in industries with relatively healthy workers. Industries or markets with greater heterogeneity in average age of employees across firms (rather than within firms) are most vulnerable to dynamic selection. Our model highlights the role of expectations rather than realizations. Insurance offer decisions are based on expected health care costs once insured and expected turnover rates, not actual, realized health care spending and job turnover. Empirical models using actual turnover rates introduce measurement error in this key variable, particularly for small firms where estimates of turnover rates have higher variances. Industries with higher turnover rates among young (healthy workers) than among older (higher cost) workers will be less likely to offer insurance. Not offering insurance may persist even in relatively large firms if the form of dynamic selection that we model is important. Also of interest are the predictions about job turnover. Selection of the type that we model is most relevant if insurance lowers job turnover, and if job turnover rates are higher in firms with young employees. Finally, the model predicts that the greater the difference in expected costs of low and high cost employees, the greater the risk that firms will choose not to offer insurance.

Our empirical analysis proceeds in five stages and uses two different data sources. We use the Robert Wood Johnson Foundation's 1997 Employer Health Insurance Survey (EHIS) and MEDSTAT's MarketScan commercially insured health claims and eligibility information. Because our empirical analysis is complex, we provide an overview here.

We first use the EHIS data to examine turnover patterns and their relationship to firm and employee characteristics. We estimate a negative binomial model of firm level turnover, and examine its relationship to insurance, employee age, employee income, firm size, and industry. The EHIS data reveals that firms are very heterogeneous not only in their age distribution and other health-related demographic variables, but also in their job turnover rates. This diversity in job turnover rates has received little attention in this literature on the uninsured, but its presence exacerbates selection problems in health insurance market in this dynamic setting. Firms with higher proportions of young (low cost) workers have higher job turnover rates and lower insurance rates than firms with higher proportions of older (high cost) workers.

We next use the EHIS employer data on firm with 100 or fewer employees to graphically examine patterns of insurance by firm size. Any effort to explain why small firms do or do not offer health insurance must come to grips with the striking pattern whereby 70 percent of firms with one or two employees do not offer insurance, while over 90 percent of firms with 100 employees do offer insurance. We examine how proxies for average employee age, average employee income, within firm heterogeneity in age and income, and turnover are related to firm size. None of the bivariate graphs by firm size provide a convincing story that that explains the striking pattern of insurance coverage by firm size.

We then turn to the MEDSTAT MarketScan data from 1998-99, where we use the individual health care costs of 890,000 adult employees. We develop three predictive models of health care spending. The first one uses only demographic information, the second disaggregated prior year spending, and the third uses prior year diagnoses organized according to the Diagnostic Cost Group (DCG) Hierarchical Condition

Category (DCG/HCC) system *(NB: due to time constraints, results from this third model are not yet included in this draft*). Expected spending from each of the three predictive models is used to examine not only the distribution of covered medical spending but also the distribution across workers in their expected cost. The distribution of expected rather than actual costs is relevant since it is expected costs rather than ex post realized costs that matter for risk selection (Ellis and McGuire, 2004). Since the MEDSTAT data covers primarily large firms with over 1000 employees, we use the data to generate random samples of employees that match as closely as possible the industry, age categories and gender ratio of the employers appearing in the EHIS sample. By repeatedly drawing random samples of firms, we are able to calculate for each firm the probability that it would have a draw of patients sufficiently low expected health care costs that they would on average prefer not to purchase insurance at the imputed premium.

Conventional wisdom leads many policymakers and researchers to expect that pooling of high and low cost employees across small firms is feasible, and that small firms will choose to offer insurance once premiums are close to the average expected costs. We show that risk pooling across firms works less well than conventional wisdom would suggest, since expected firm-level health care spending varies dramatically across firms. The extent of risk pooling is inadequate to induce many small to medium size firms to purchase insurance when they have systematically different age, gender and industry distributions than the average. The threat that it may worsen due to adverse selection induced turnover exacerbates this problem. Our analysis provides a rationale for why many firms may decide not to offer insurance, and is the only one to date that explains the striking pattern by firm size.

To create appropriate expectations, we would like to highlight what this paper does not do. We do not explicitly model the improved labor productivity, the tax subsidy for purchasing insurance, nor the risk aversion of workers, which may motivate some firms to offer insurance to their workers. We have to date only done simulations as if all employees are single, whereas in practice many have families, and will purchase family rather than individual coverage. The EHIS data does not include the key variable the proportion of employees who are family rather than individual contracts, unless the firm offers insurance, hence we are not able to use this information in our analysis.

A Model of Insurance and Expectations

There is a population of workers with heterogeneous healthcare cost characteristics. We normalize the total population to be a mass of one. A randomly drawn worker from this population can have an expected healthcare cost either c_L or c_H, and the proportion of workers in the full population with cost c_L is $\bar{q}$. A worker's cost characteristic is his private information; workers are otherwise similar.

Suppose that in a given period, *t*, a firm has hired a number of workers from the marketplace. Again, each of the selected workers can be one of two types: one with a low expected healthcare cost c_L, or high expected healthcare cost c_H. We assume that the respective probabilities for the low-cost and high-cost types are q_t and $1-q_t$. Notice that q_t is not necessarily $\bar{q}$.

Each of the employed workers may leave the firm in a given period with some chance. Workers may search for jobs either actively or passively. A worker's departure rate is dependent on his expected healthcare cost: the probability that a type c_i worker leaves the firm is l_i, $i=L, H$, with $l_L > l_H$. So we assume that those workers with low expected healthcare cost are healthier and hence more mobile.

So at the end of a period, a total of $q_t\ l_L + (1-q_t)\ l_H$ workers will leave the firm. The following period, the firm has to replace these workers, and must therefore hire from the general population. Because we are considering a small firm, we assume that the firm simply gets a random draw from the general population.[1]

[1] A more general model would have each firm drawing from a different distribution of low and high cost workers, however that would add complexity without new insights.

So for the new hires, the expected healthcare cost is simply the expected cost of the worker chosen randomly from the population: $\bar{q}\ c_L + (1-\bar{q})\ c_H$.

Of the workers who have continued from the previous period, $q_t\ (1-l_L)$ of these are type c_L workers, while $(1-q_t)\ (1-l_H)$ are type c_H workers. So the total expected cost of these continuing workers is $q_t\ (1-l_L)\ c_L + (1-q_t)\ (1-l_H)\ c_H$.

After the replacement workers have been hired, the expected healthcare cost of the firm is

(1) $$\left[q_t l_L + (1-q_t) l_H\right]\left[\bar{q}c_L + (1-\bar{q})c_H\right] + \left[q_t(1-l_L)c_L + (1-q_t)(1-l_H)c_H\right]$$

This is therefore the expected healthcare cost for the period $t+1$. We let q_{t+1} be the share of low-cost workers at period $t+1$. So the expression in (1) can be written as $q_{t+1}\ c_L + (1-q_{t+1})\ c_H$.

In a long run, steady state equilibrium, the healthcare cost of the firm will stay constant from period to period. Let the steady-state percentage of low-cost workers in the firm be $\hat{q}$. To solve for theta, we just note that the value of (1) evaluated at $\hat{q}$ must be identical to $\bar{q}c_L + (1-\bar{q})c_H$. Solving this for $\hat{q}$ yields

(2) $$\hat{q} = \frac{\bar{q}}{\bar{q} + (1-\bar{q})\left(\frac{l_L}{l_H}\right)} < \bar{q}$$

where the inequality follows from $l_L > l_H$. In a steady state, the firm will have a workforce with healthcare cost higher than the general population. The intuition is straightforward. In each period, relatively more workers with lower healthcare costs will depart from the firm. So the average healthcare cost of those who remain must rise. The replacement workers, drawn randomly from the population, must counterbalance the cost hike. So the average cost of the workforce must be higher than the population; otherwise, the counterbalance will be ineffective.

We now examine the dynamics. From (1), and the definition of q_{t+1}, we obtain a difference equation:

(3) $$q_{t+1} = [q_t l_L + (1-q_t) l_H]\bar{q} + q_t(1 - l_L)$$

We then subtract $\hat{q}$ on each side of this equation, collect terms and simplify. This yields the following to describe the period-to-period variation of the share of low-cost workers.

(4) $$q_{t+1} - \hat{q} = (q_t - \hat{q})[1 - \bar{q} l_H - (1-\bar{q}) l_L]$$

Because the coefficient of $(q_t - \hat{q})$ is positive and less than 1, the system is stable. From any initial point, eventually q_t will get close to $\hat{q}$.

We have described the employment process and time-paths of a small firm facing workers with different healthcare costs and (correlated) departure rates. How is this related to the firm's insurance provision decision? If fair insurance policies are offered to the firm, the firm will find it advantageous to provide health insurance to workers because of the tax treatment and risk aversion. Whether an insurance policy can be made actuarially fair depends on how well the time path of q_t can be predicted. In practice, an insurer may lack the utilization data to predict the time path of a firm's expected healthcare costs. As well, there are significant costs in changing premium, and adjusting for large cost deviations.

If an insurer either cannot predict the time-path of a firm's expected cost, or find it prohibitively cost to adjust premium charges frequently, how is the market supposed to work? We postulate that an insurer will attempt to price the premium at the long-run, state-state expected cost. Thus, for the above model, an insurer may charge a premium equal to $\hat{q}\, c_L + (1-\hat{q})\, c_H$. This may be reasonably accurate when λ_L and λ_H, workers' departure rates, are small. Nevertheless, for small firms, departure rates may be high.

For an illustration of the failure of fair insurance pricing, suppose now that q_t is much below $\hat{q}$, so that the firm's healthcare cost tends to be very high. An insurer assuming the steady state will choose a premium corresponding to q at $\hat{q}$. This firm will find this policy very attractive, since the premium is below the expected cost. On the other hand, a firm that has a q much above $\hat{q}$ will find it unattractive: it has a workforce with low expected healthcare costs, but it is asked to bear the premium for the steady state. If risk aversion is not severe, the firm may choose not to buy insurance. Now, as firms with low expected costs opt out, what remains in the market are firms with high costs. The insurer therefore raises the premium even more, and a kind of death spiral can occur.

Firms may offer different levels of insurance coverage. It is likely that the extent of insurance coverage will affect workers' turnover. A generous coverage will be more attractive to those workers with high healthcare cost; that is, the value of l_H will become lower. This means that the steady-state share of low-cost workers in the firm will be lower. Again, if the insurer insists on pricing at the long-run equilibrium premium level, the firm may find it too costly to offer insurance with a generous coverage.

Data

We use data from a variety of sources. Our primary file on firm characteristics is the Robert Wood Johnson Foundation 1997 Employer Health Insurance Survey (EHIS). This employer CAT survey has been used for similar analyses by Chu and Trapnell (2002) and others. The survey collects a rich set of information about the firm, including whether or not insurance is offered, establishment size, ten broad industry groups, and most importantly from our perspective, the proportion female and the percent of workers in each of four broad age categories. Starting from the full survey results on 41,432 employers, we excluded results from 8,710 governments, firms with no permanent full time employees, firms with missing values (mostly for income or industry type), and firms with over 5000 employees (who may have self administration and other options available to them). Our final estimation sample includes 20,585 firms. Our simulation modeling focuses on the 18,712 firms with less than 100 full-time, permanent employees. Summary statistics on the variables used for estimation appear in the tables 1.

Empirical analysis of job turnover.

While we have largely completed the empirical analysis for this paper, we have not had time to write it up. Attached are the tables and figures, with relatively complete notes indicating how they were generated. A more complete version of the paper will be distributed at the ERIU conference.

References

1. Ash, A.S., R.P. Ellis, G.C. Pope, J.Z. Ayanian, D.W. Bates, H. Burstin, L.I. Iezzoni, E. McKay, W. Yu, (2000) “Using Diagnoses to Describe Populations and Predict Costs.” Health Care Financing Review, Spring 21(3): 7-28.
2. Blumberg, L.J., and L.M Nichols (2002) “Why are so many Americans Uninsured? A conceptual framework, summary of the evidence, and delineation of the gaps in our knowledge.” ERIU Working Paper 3, University of Michigan.
3. Chernew, and K. Frick (1999) “The impact of managed care on the existence of equilibrium in health insurance markets.” Journal of Health Economics, 18: 573-592.
4. Chernew, M.E., and R. A. Hirth (2002) “Modeling the causes and consequences of lack of health insurance coverage: Gaps in the literature.” ERIU Working Paper 1, University of Michigan.
5. Chernew, M.E., K. Frick, C. McLaughlin (1997) The demand for health insurance coverage by low-income workers: Can reduced premiums achieve full coverage? Health Services Research. 32(1): 102-110.
6. Chu, R.C., G.R. Trapnell. (2002) “Study of the administrative costs and actuarial values of small health plans.” Actuarial Research Corporation, Annandale VA 22003.
7. Congressional Research Service, (1988)
8. Cutler, D. M. and R. J. Zeckhauser (2000). "The Anatomy of Health Insurance", in A. J. Culyer and J. P. Newhouse, eds., Handbook of Health Economics (Elsevier, Amsterdam) Chapter 11.
9. Ellis, R.P., and F. Aragao, (2001) “Death Spirals, Switching Costs, and Health Plan Premium Payments. Boston University ISP Working Paper, February.
10. Ellis, R.P. and T.G. McGuire (2004) “Predictability and Predictiveness in Health Care Spending” Boston University ISP Working Paper, June.
11. Gruber, J. and M. Lettau. (2000) “How elastic is the firms demand for health insurance?” Cambridge MA, National Bureau of Economic Research working paper.
12. Gruber, J. and B. Madrian. (2001) “Health Insurance, Labor Supply, and job mobility: a critical review of the literature.” Unpublished ERIU working paper, U. of Michigan. November.
13. Ma, C.T.A. and M. H. Riordan (2002) Health insurance, moral hazard, and managed care.” Journal of Economics and Management Strategy, 11: 81-107.
14. Miller, N. “Health Benefits and Wages: Minimizing the Total Cost of Compensation.” John F. Kennedy School of Government, Harvard University, Faculty Research Working Paper Series. January 2002.

15. Moran, J.R., M. Chernew, and R. Hirth (2001) Preference Diversity and Breadth of Employee Health Insurance Options." Health Services Research 34(4): 813-837.
16. Rothschild M. and J. Stiglitz (1976). "Equilibrium in Competitive Markets: an essay on the Economics of Imperfect Information", Quarterly Journal of Economics 90:629-649.
17. Short, P.F. (2001) "Counting and Characterizing the Uninsured" Unpublished ERIU working paper, University of Michigan, December.
18. Swartz, K., and D. Garnick. (2000) "Adverse selection and price sensitivity when low income people have subsidies to purchase health insurance in the private market." Inquiry. 27(1): 45-60
19. Van de Ven, W.P.M.M. and R.P. Ellis. (2000) "Risk Adjustment in Competitive Health Plan Markets" Chapter 14 in A.J. Culyer and J.P. Newhouse, (ed.) Handbook in Health Economics, North Holland. pp. 755-845.

Table 1

Sample Means, RWJ 1997 EHIS sample of private employers

	All Firms (N=20585)		Firms offering Insurance (N=13587)		offering insurance (N = 6998)	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
No insurance offered	0.660044	0.4737055	1	0	0	0
turnover rate	0.484095	0.7982073	0.4534	0.7199	0.54379	0.92874
Employee characteristics						
fulltime proportion	0.876163	0.2248601	0.9038	0.1921	0.82254	0.26971
temporary proportion	0.07453	0.1840767	0.0626	0.1629	0.0976	0.21759
union proportion	0.035648	0.1585616	0.0477	0.181	0.01221	0.09761
Employee Age and Gender						
females	0.425421	0.315359	0.4241	0.2984	0.42797	0.34589
employees age <25	0.283452	0.2656942	0.2861	0.2467	0.27821	0.29905
employees age 25-34	0.300202	0.250114	0.3038	0.2247	0.29326	0.29317
employees age 35-44	0.240821	0.2408495	0.2434	0.2165	0.2358	0.28212
employees age 45-54	0.175722	0.2384135	0.1669	0.2111	0.19277	0.28328
Mean age	38.09496	6.7684754	37.921	6.181	38.4331	7.77278
within firm std of age	7.331593	3.6924686	7.9021	3.2636	6.22387	4.1916
between firm std of mean age						
income <$10k	0.058612	0.1832364	0.0329	0.1255	0.10849	0.2538
income $10-14k	0.181171	0.2851524	0.1364	0.2332	0.26813	0.34951
income$14-20k	0.235002	0.2754423	0.2286	0.2519	0.24745	0.31577
income$20-30k	0.255698	0.2753952	0.28	0.2578	0.20858	0.3012
Mean income	22.46356	7.1680783	24.06	6.5092	19.3638	7.37063
within firm std income	4.707952	3.0938258	5.3325	2.7786	3.49538	3.30748
between firm std of mean income						
Industry codes						
agriculture, fisheries, forestry	0.002235	0.0472202	0.002	0.0445	0.00272	0.05204
construction	0.080738	0.2724396	0.0719	0.2583	0.09789	0.29718
manufacturing and mining	0.140685	0.3477047	0.1718	0.3772	0.08031	0.27179
transport, commun, utilities	0.049988	0.2179252	0.057	0.2318	0.03644	0.18739
wholesale trade	0.050279	0.2185261	0.06	0.2375	0.03144	0.17451
retail trade	0.189507	0.3919203	0.1445	0.3517	0.27679	0.44745
financial services	0.17435	0.3794196	0.18	0.3842	0.16348	0.36983
professional services	0.244498	0.4297999	0.2602	0.4387	0.21406	0.4102
Firm size measures						
size=Number of full time employees	60.10838	253.14764	86.104	307.35	9.63647	35.3495
more = 1 if more employees nationwide	0.333641	0.4715246	0.4295	0.495	0.14761	0.35474
1-9 employees at establishment	0.506437	0.4999707	0.3653	0.4815	0.78037	0.41403
10-24 employees at establishment	0.209522	0.406977	0.2404	0.4273	0.14961	0.35672
25-49 employees at establishment	0.114987	0.3190136	0.1496	0.3566	0.04787	0.21351
50-99 employees at establishment	0.072043	0.258565	0.1014	0.3019	0.015	0.12158
100-249 employees at establishment	0.053631	0.2252941	0.079	0.2698	0.00429	0.06534
250+ employees at establishment	0.043381	0.2037184	0.0643	0.2452	0.00286	0.05339
more*1-9 employees at est.	0.113918	0.3177192	0.1209	0.3261	0.10031	0.30044
more*10-24 employees at est.	0.071071	0.2569499	0.0941	0.292	0.02629	0.16002
more*25-49 employees at est.	0.050134	0.2182259	0.0693	0.254	0.01286	0.11268
more*50-99 employees at est.	0.037017	0.1888086	0.0536	0.2252	0.00486	0.06954
more*250+ employees at est.	0.032159	0.1764274	0.0475	0.2127	0.00243	0.04923
more*100-249 employees at est.	0.029342	0.1687667	0.044	0.2051	0.00086	0.02927
max(0,size-5)	55.78309	252.99613	81.468	307.25	5.91455	35.1065
max(0,size-10)	52.98479	252.52386	77.978	306.86	4.45956	34.5905
max(0,size-25)	47.69351	250.59445	70.813	304.96	2.80509	33.2256
max(0,size-50)	42.43911	247.04009	63.313	301.11	1.91155	31.5332
max(0,size-100)	36.39801	239.97593	54.44	293.02	1.36939	28.9198
Simulation results						
Mean firm level predicted medical cost	1560	456	1542	508	1590	421
within-firm std of predicted medical cost	743		604		978	
between-firm std of mean predicted cost	456		508		421	

Table 2

Logit models of the insurance offer decision

		Ordinary logit			**Fixed effects logit**
			Standard	Wald	
Parameter	Estimate	Error	Chi-Square	Pr > ChiSq	
Intercept	0.646	4.351	0.02	0.882	
turnover rate	-0.026	0.023	1.35	0.2461	
Employee characteristics					
fulltime proportion	1.179	0.086	187.92	<.0001	
temporary proportion	0.023	0.099	0.05	0.8153	
union proportion	0.557	0.161	12.02	0.0005	
Employee Age and Gender					
females	-0.016	0.066	0.06	0.8082	
employees age <25	-4.386	4.333	1.02	0.3115	
employees age 25-34	-4.184	4.334	0.93	0.3344	
employees age 35-44	-3.993	4.334	0.85	0.3569	
employees age 45-54	-4.017	4.333	0.86	0.3538	
within firm std of age	0.016	0.006	8.33	0.0039	
income <$10k	-2.310	0.117	390.22	<.0001	
income $10-14k	-2.050	0.083	612.11	<.0001	
income$14-20k	-1.143	0.077	220.28	<.0001	
income$20-30k	-0.501	0.081	38.62	<.0001	
within firm std income	0.108	0.006	311.72	<.0001	
Industry codes					
agriculture, fisheries, forestry	0.448	0.364	1.51	0.2185	
construction	-0.076	0.095	0.65	0.4217	
manufacturing and mining	0.412	0.091	20.33	<.0001	
transport, commun, utilities	0.494	0.114	18.86	<.0001	
wholesale trade	0.703	0.115	37.43	<.0001	
retail trade	0.155	0.080	3.77	0.0522	
financial services	0.447	0.082	29.60	<.0001	
professional services	0.516	0.079	42.54	<.0001	
Firm size measures					
more = 1 if more employees nationwide	0.834	0.296	7.96	0.0048	
1-9 employees at establishment	0.933	0.640	2.12	0.1452	
10-24 employees at establishment	1.001	0.625	2.56	0.1094	
25-49 employees at establishment	0.825	0.592	1.94	0.1633	
50-99 employees at establishment	0.485	0.506	0.92	0.3378	
100-249 employees at establishment	0.621	0.376	2.72	0.099	
more*1-9 employees at est.	0.555	0.301	3.40	0.0652	
more*25-49 employees at est.	0.463	0.312	2.21	0.1373	
more*50-99 employees at est.	0.083	0.328	0.06	0.8009	
more*100-249 employees at est.	0.072	0.372	0.04	0.8477	
size=Number of full time employees	0.295	0.022	174.94	<.0001	
max(0,size-5)	-0.166	0.039	17.95	<.0001	
max(0,size-10)	-0.073	0.025	8.30	0.004	
max(0,size-25)	-0.013	0.014	0.85	0.3572	
max(0,size-50)	-0.024	0.013	3.44	0.0636	
max(0,size-100)	-0.017	0.008	4.23	0.0397	
	N				
-2 Log L	17948				
Number of Observations	20585				

Table 3

Negative Binomial model of turnover rates.
ALL FIRMS

No. of Obs	19827
Model chi2(16)	2211.11
Prob>chi2	0
Log Likelihood	-57297.6
Pseudo R2	0.0189

Dependent variable: turnover	**Coefficient**	**Std Error**	**Z**
Impact of insurance on turnover			
Insurance is offered	0.056	0.124	0.452
insure*proportion age <25	-0.097	0.134	-0.723
insure*proportion age 25-34	-0.114	0.139	-0.819
insure*proportion age 35-44	-0.123	0.183	-0.670
Employee characteristics			
fulltime proportion	-0.223	0.050	-4.510
temporary proportion	0.756	0.060	12.602
union proportion	-0.385	0.067	-5.774
income <$10k	0.175	0.061	2.869
income $10-14k	0.441	0.047	9.381
income$14-20k	0.192	0.046	4.204
income$20-30k	0.035	0.052	0.686
Employee Age and Gender			
females	0.029	0.038	0.775
employees age <25	1.169	0.110	10.608
employees age 25-34	0.657	0.122	5.400
employees age 35-44	0.438	0.145	3.014
Industry codes			
agriculture, fisheries, forestry	-0.291	0.215	-1.354
construction	0.278	0.062	4.498
manufacturing and mining	-0.129	0.052	-2.463
transport, commun, utilities	0.002	0.064	0.038
wholesale trade	-0.211	0.076	-2.782
retail trade	0.048	0.049	0.980
financial services	-0.045	0.050	-0.916
professional services	-0.189	0.053	-3.562
Firm size measures			
more = 1 if more employees nation	0.056	0.076	0.735
1-9 employees at establishment	0.194	0.061	3.173
10-24 employees at establishment	0.245	0.064	3.822
25-49 employees at establishment	0.211	0.072	2.929
50-99 employees at establishment	0.233	0.076	3.058
100-249 employees at establishme	0.227	0.086	2.654
more*1-9 employees at est.	0.230	0.089	2.584
more*10-24 employees at est.	0.124	0.085	1.464
more*25-49 employees at est.	0.071	0.086	0.819
more*50-99 employees at est.	0.156	0.100	1.556
more*250+ employees at est.	0.028	0.114	0.245
constant	-1.667	0.138	-12.113
/lnalpha	0.093	0.018	5.213
alpha			

Base for negative binomial is the number of employees

Table 4

Linear regression model of annualized total covered medical spending excluding drugs of employees

MEDSTAT MarketScan Commercially insured adults, aged 18-65

	Demographic only		Demographic with Spline on lagged costs	
R-Square	0.011		0.09135	
Root MSE	7332.497		7027.762	
N	891,857		891,857	
Dependent variable Mean	1,817		1,817	
	Coefficient	t value	Coefficient	t value
Intercept	-387	-1.63	80	0.35
Industry				
Manufact, durable Goods	-124	-4.42	21	0.64
Manufact, nondurable good:	33	0.99	281	6.26
Services	-127	-4.81	295	7.39
Transportation, Communica	-83	-2.91	438	10.33
Missing	.	.	.	.
Plane Type				
Basic/Major Medical	348	3.11	679	6.1
Comprehensive	488	23.49	270	10.79
HMO	-403	-12.69	-425	-12.57
POS	158	4.85	-222	-6.44
POS with Capitation	-266	-9.19	50	1.61
PPO	.	.	.	.
single	-47	-2.84	10	0.63
Age and age splines				
age	45	4.5	17	1.73
max(0,age-30)	-34	-2.41	-12	-0.85
max(0,age-40)	57	4.86	26	2.31
max(0,age-50)	39	2.75	36	2.65
max(0,age-60)	-203	-5.62	-205	-5.95
Discrete age-sex categories				
Female, Age 18-24	725	8.9	506	6.46
Female, Age 25-34	932	10.75	606	7.27
Female, Age 35-44	661	6.56	325	3.36
Female, Age 45-54	513	4.61	242	2.27
Female, Age 55-64	422	3.3	214	1.75
Female, Age 60-64	368	2.33	305	2.01
Male, Age 18-24	.	.	.	.
Male, Age 25-34	-150	-1.72	-81	-0.96
Male, Age 35-44	-22	-0.22	26	0.27
Male, Age 45-54	120	1.08	200	1.87
Male, Age 55-59	247	1.94	273	2.23
Male, Age 60-64	683	4.31	645	4.25
Employee classes				
Salary Non-union			-696	-16.83
Salary Union			-873	-21.06
Salary Other			-978	-20.96
Hourly Non-union			-386	-4.99
Hourly Union			-453	-5.93
Hourly Other			-792	-19.86
Non-union			-243	-9
Union			-177	-1.56
Unknown			.	.

Linear regression model of annualized total covered medical spending excluding drugs
(continued)

	Demographic only		Demographic with Spline on lagged costs	
	Coefficient	t value	Coefficient	t value
Splines using lagged health spending information				
Dummy=1 if any OP$			131	2.67
OP $ in 1998			1.263	2.22
max(0,OP$-100)			-0.042	-0.07
max(0,OP$-500)			-0.112	-0.77
max(0,OP$-1000)			-0.693	-9.19
max(0,OP$-5000)			-0.140	-5.72
max(0,OP$-10000)			0.505	25.78
max(0,OP$-50000)			-0.314	-24.3
Dummy=1 if any IP$			-218	-0.22
IP$ in 1998			1.013	0.99
max(0,IP$-1000)			-1.221	-1.18
max(0,IP$-5000)			0.468	8.25
max(0,IP$-10000)			-0.044	-1.68
max(0,IP$-50000)			-0.139	-21.04
Dummy=1 if any drug$			-146	-3.79
Drug$ in 1998			1.132	2.08
max(0,drug$-1000)			1.074	1.63
max(0,drug$-5000)			0.470	1.21
max(0,drug$-10000)			-1.772	-5.27

Notes: Regressions used MEDSTAT Marketscan Commercially insured data using
only full time active employees, aged 18 through 64.
Dependent variable is annualized 1999 covered inpatient plus outpatient health care costs
Spending by people eligible for only part of 1999 were annualized by dividing by the
fraction of the year eligible in that year
OP$ stands for covered outpatient spending in 1998, IP$ stands for covered
inpatient spending in 1998; drug$ stands for covered drug spending in 1998.

Table 5

Percentile distributions of predicted spending from 250 simulations of each firm, by firm size categories

Percentiles	1-9 employees	10-24 employees	25-49 employees	50-99 employees	100-249 employees	250+ employees	All Firms
100%	31442	10042	5440	3199	3056	2690	31442.021
99%	5310	3663	2900	2623	2470	2535	4391.507
95%	3205	2529	2281	2145	2107	2055	2769.165
90%	2615	2180	2022	1940	1925	1964	2319.001
75%	1954	1792	1726	1711	1719	1766	1828.324
50%	1480	1478	1470	1504	1541	1618	1494.586
25%	1090	1220	1285	1338	1366	1454	1215.199
10%	788	1033	1140	1204	1255	1320	946.114
5%	622	922	1067	1140	1194	1247	770.827
1%	267	767	915	1009	1083	1162	417.872
0%	0	483	743	875	952	1081	0
N	10403	4292	2348	1469	1100	893	20505
Mean predicted cost	1666	1579	1546	1552	1572	1629	1619
Std. Dev	1143	595	399	315	285	255	878
Administrative cost multiplier	1.33	1.25	1.25	1.20	1.16	1.05	1.27
Imputed premium	2154	2024	2024	1943	1878	1700	2062

Note: This table was generated by combining RWJ 1997 EHIS and MEDSTAT 1998-99 data. 250 random samples of employees from the MEDSTAT data matching the age-gender-industry intervals were drawn for each firm appearing in the EHIS data. Average predicted spending per employee is shown where predictions are generated using demographic and lagged spending splines as in Model 2 of Table 1. Imputed premiums were calculated assuming fully community rating. Imputed premiums are the grand mean average cost ($1619) increased by average administrative cost percentages based on Chu and Trapnell (2002). The percentiles from the simulated distribution for each firm size category are shown above.

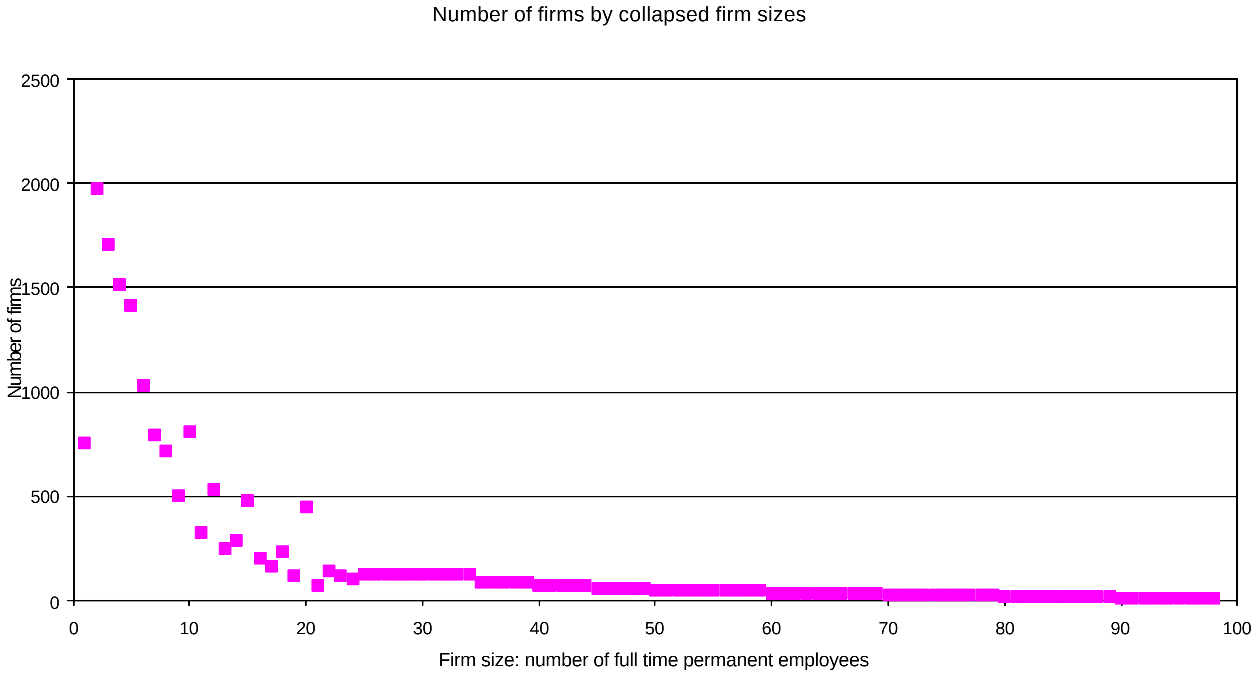

Figure 1 Number of firms by collapsed firm sizes, RWJ 1997 EHIS data, firm sizes less than 100. N=19,712. Firms with sizes between 25 and 49 inclusive were collapsed into five categories of size to ensure that at least 100 firms were in each collapsed category, while firms with 50 or more enrollees were grouped into firm size intervals of ten.

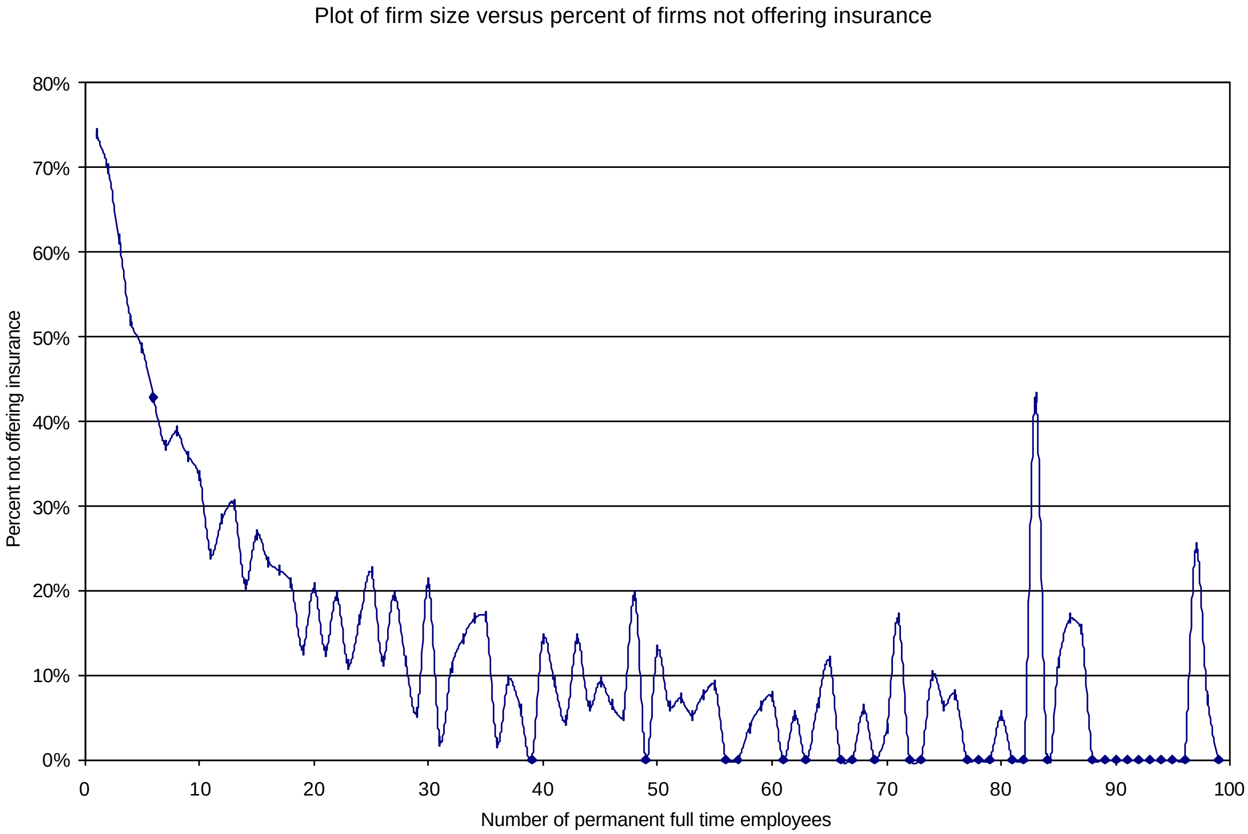

Figure 2. Plot of firm size versus percent of firms not offering health insurance. Plot uses RWJ 1997 EHIS data, for firm sizes less than 100. N=19,712. Firms with sizes between 25 and 49 inclusive were collapsed into five categories of size to ensure that at least 100 firms were in each collapsed category, while firms with 50 or more enrollees were grouped into firm size intervals of ten. Data points are sample proportions for given firm size.

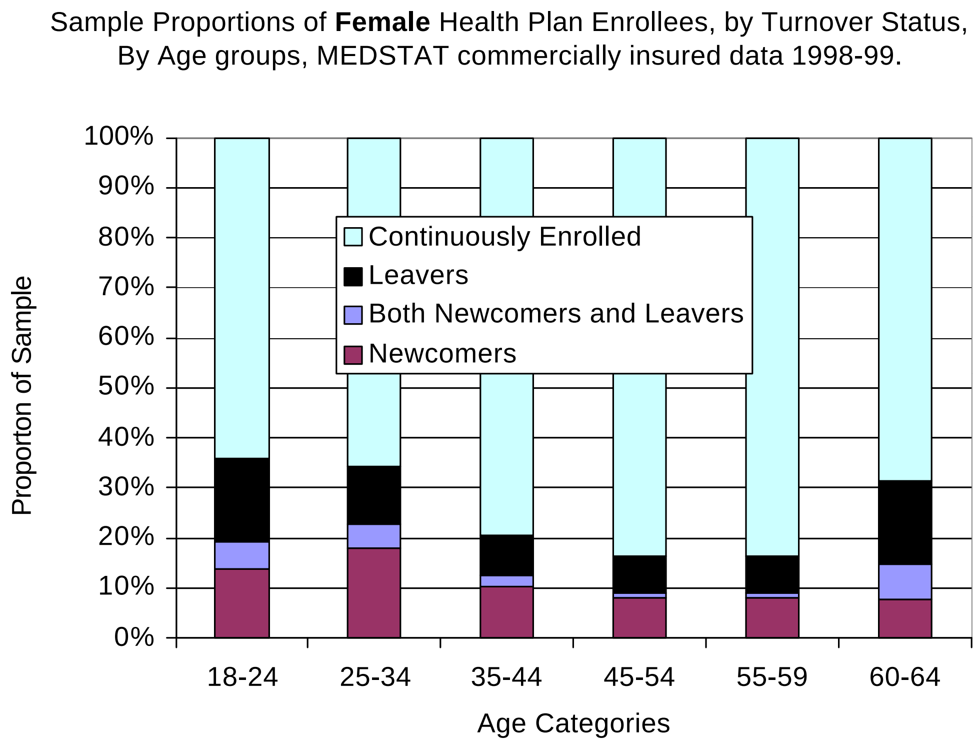

Figure 3. Sample proportions of female health plan enrollees, by turnover status, by age groups. MEDSTAT MarketScan commercially insured data 1998-1999 (N=891,857). Continuously reenrolled sample includes those eligible for all 12 months of 1999. Newcomers were eligible for less than 12 months in 1999 and not at all in 1998. Leavers were eligible for at least month in each of 1998 and 1999. Both newcomers and leavers were eligible for less than 12 months in 1999 and not enrolled at all in December 1999. Plan changes reflect both employment changes and switches between plans, which cannot be distinguished with this data. Since plan switches are relatively rare, it is plausible that the majority is due to job turnover.

Sample Proportions of **Male** Health Plan Enrollees, by Turnover Status,
By Age groups, MEDSTAT commercially insured data 1998-99.

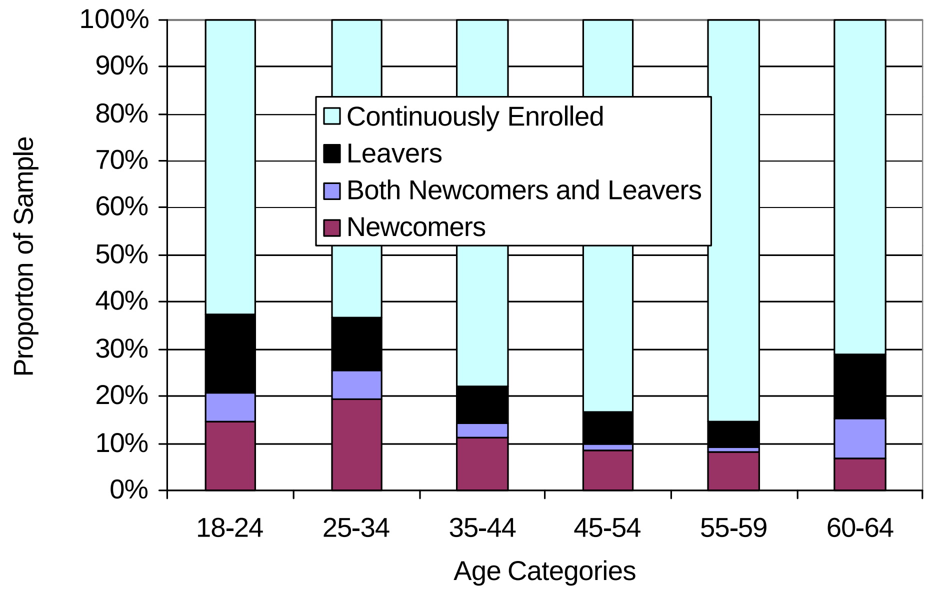

Figure 4 Sample proportions of male health plan enrollees, by turnover status, by age groups, using MEDSTAT MarketScan commercially insured data 1998-1999 (N=891,857). Continuously enrolled cohort includes those eligible for all 12 months of 1999. Newcomers were eligible for less than 12 months in 1999 and not at all in 1998. Leavers were eligible for at least month in each of 1998 and 1999. Both newcomers and leavers were eligible for less than 12 months in 1999 and not enrolled at all in December 1999. Plan changes reflect both employment changes and switches between plans, which cannot be distinguished with this data. Since plan switches are relatively rare, it is plausible that the majority is due to job turnover.

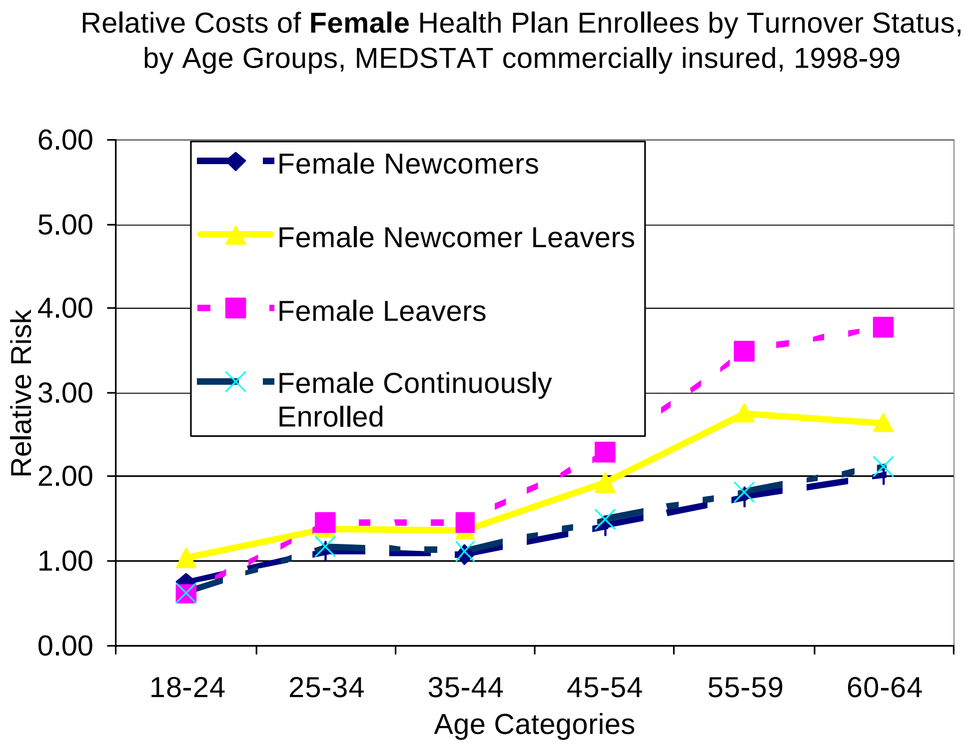

Figure 5. Relative costs of female health plan enrollees by turnover status, by age groups. MEDSTAT MarketScan commercially insured sample, 1998-99. Sample means of annualized health care spending for each cohort were divided by the overall sample mean to convert costs into relative risk. A score of one corresponds to the mean cost for the entire insured population including children.

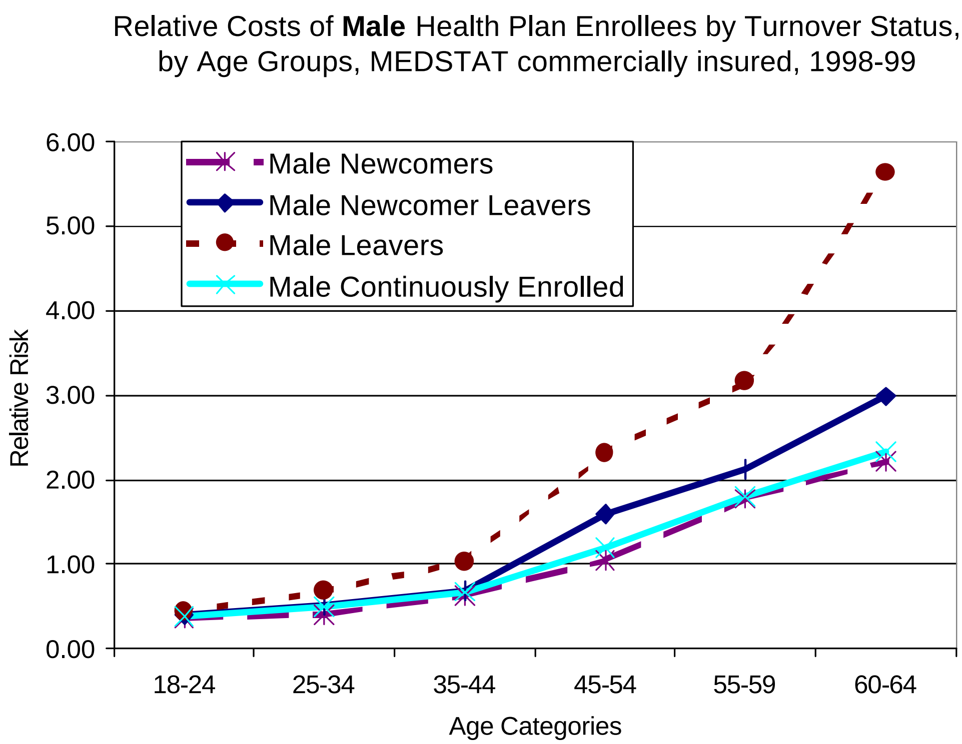

Figure 6. Relative costs of male health plan enrollees by turnover status, by age groups, MEDSTAT MarketScan commercially insured sample, 1998-99. Sample means of annualized health care spending for each cohort were divided by the overall sample mean to convert costs into relative risk. A score of one corresponds to the mean cost for the entire insured population including children.

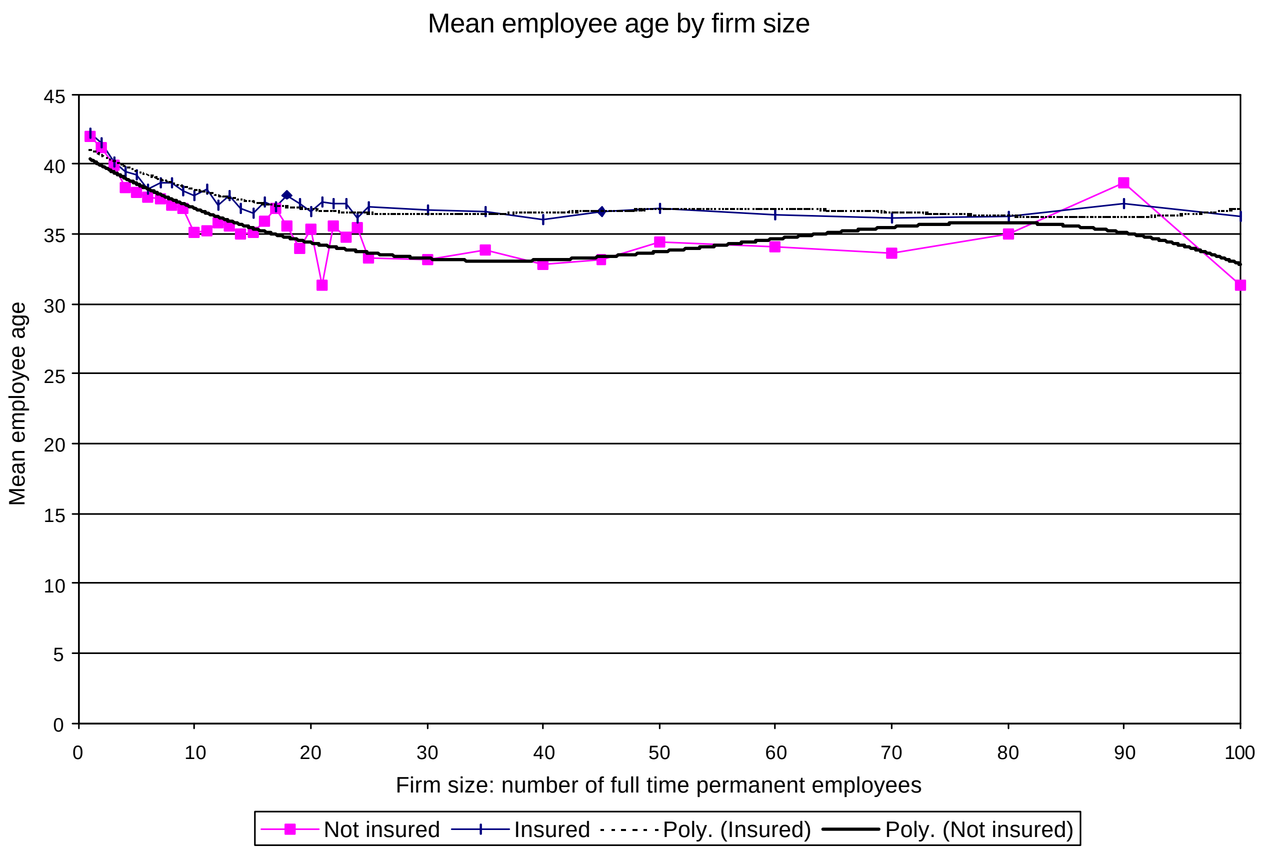

Figure 7. Mean employee age, by collapsed firm size. RWJ 1997 EHIS data, N=18,712. Average age was calculated for firms offering and not offering insurance using five age interval proportions, (18-24, 25-34, 35-44, 45-54, 55-65) using midpoints for each interval. Collapsed firm sizes were generated as previously discussed. Points shown are averages for each collapsed firm size. Polynomial trend lines are bold for uninsured firms, dashed for insured firms.

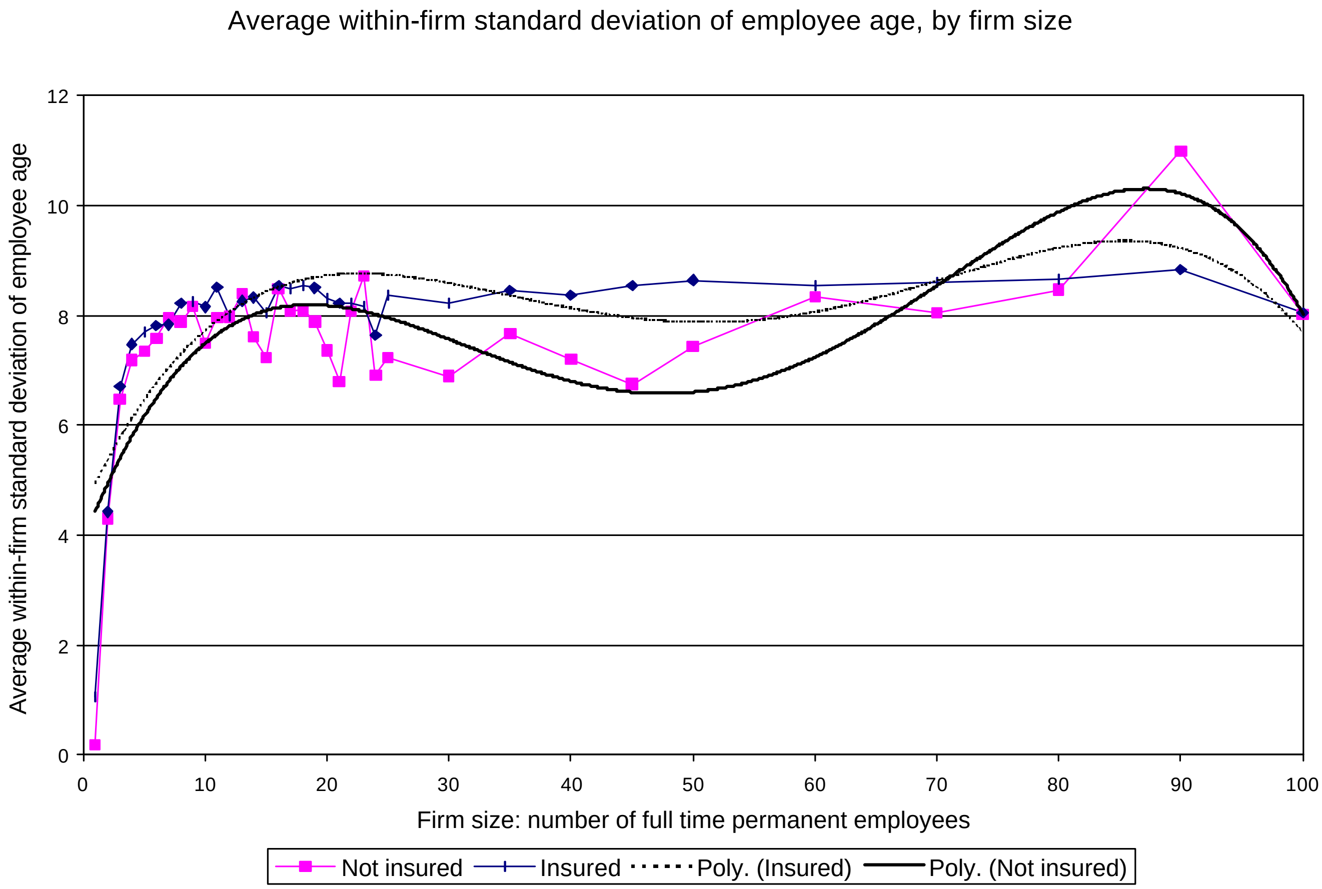

Figure 8. Average within-firm standard deviation of employee age, by collapsed firm size. RWJ 1997 EHIS data, N=18,712. Average age was calculated for firms offering and not offering insurance using five age interval proportions, (18-24, 25-34, 35-44, 45-54, 55-65) using midpoints for each interval. Standard deviations around this firm-specific average were calculated using simple weighted average of deviations using same midpoints. Collapsed firm sizes were generated as previously discussed. Points shown are averages for each collapsed firm size. Polynomial trend lines are bold for uninsured firms, dashed for insured firms.

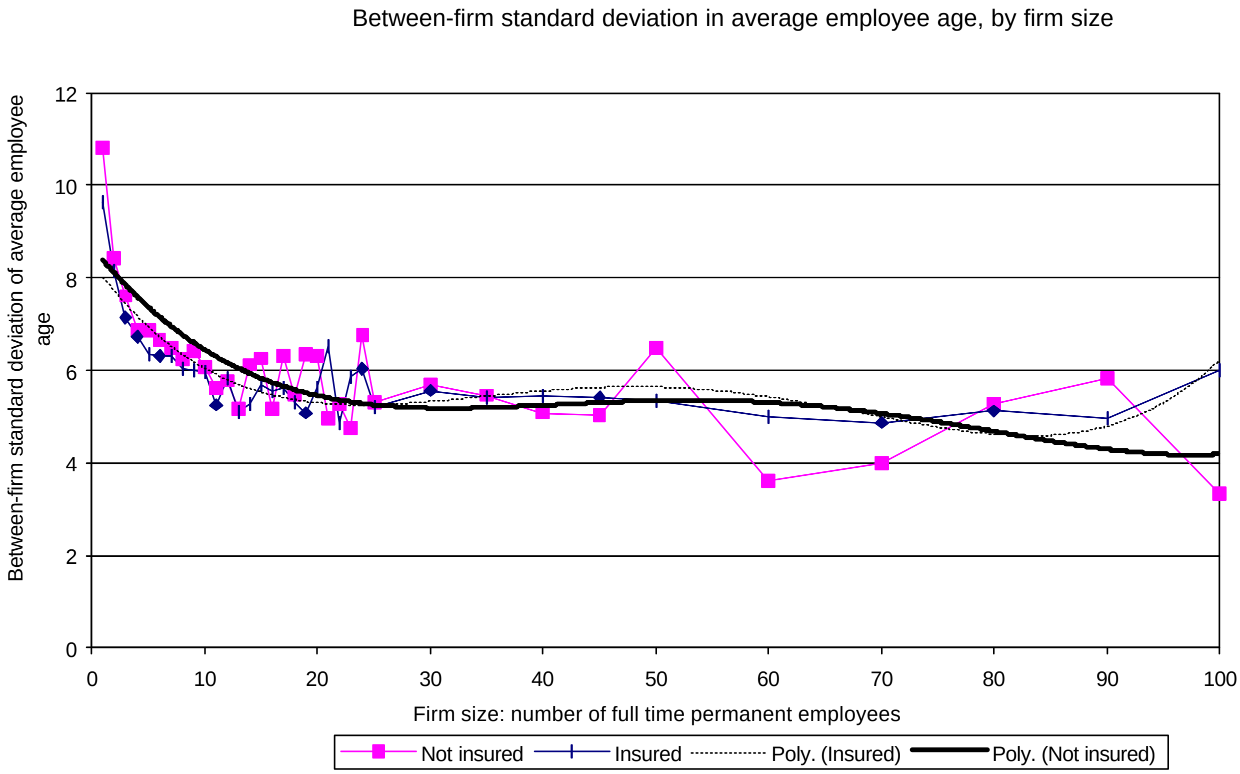

Figure 9. Between-firm standard deviation in average employee age, by collapsed firm size. RWJ 1997 EHIS data, N=18,712. Average age for each firm was calculated using five age interval proportions, (18-24, 25-34, 35-44, 45-54, 55-65) using midpoints for each interval. Collapsed firm sizes were generated as previously discussed. Standard deviations were calculated for each collapsed firm size separately for firms offering and not offering insurance. Polynomial trend lines are bold for uninsured firms, dashed for insured firms.

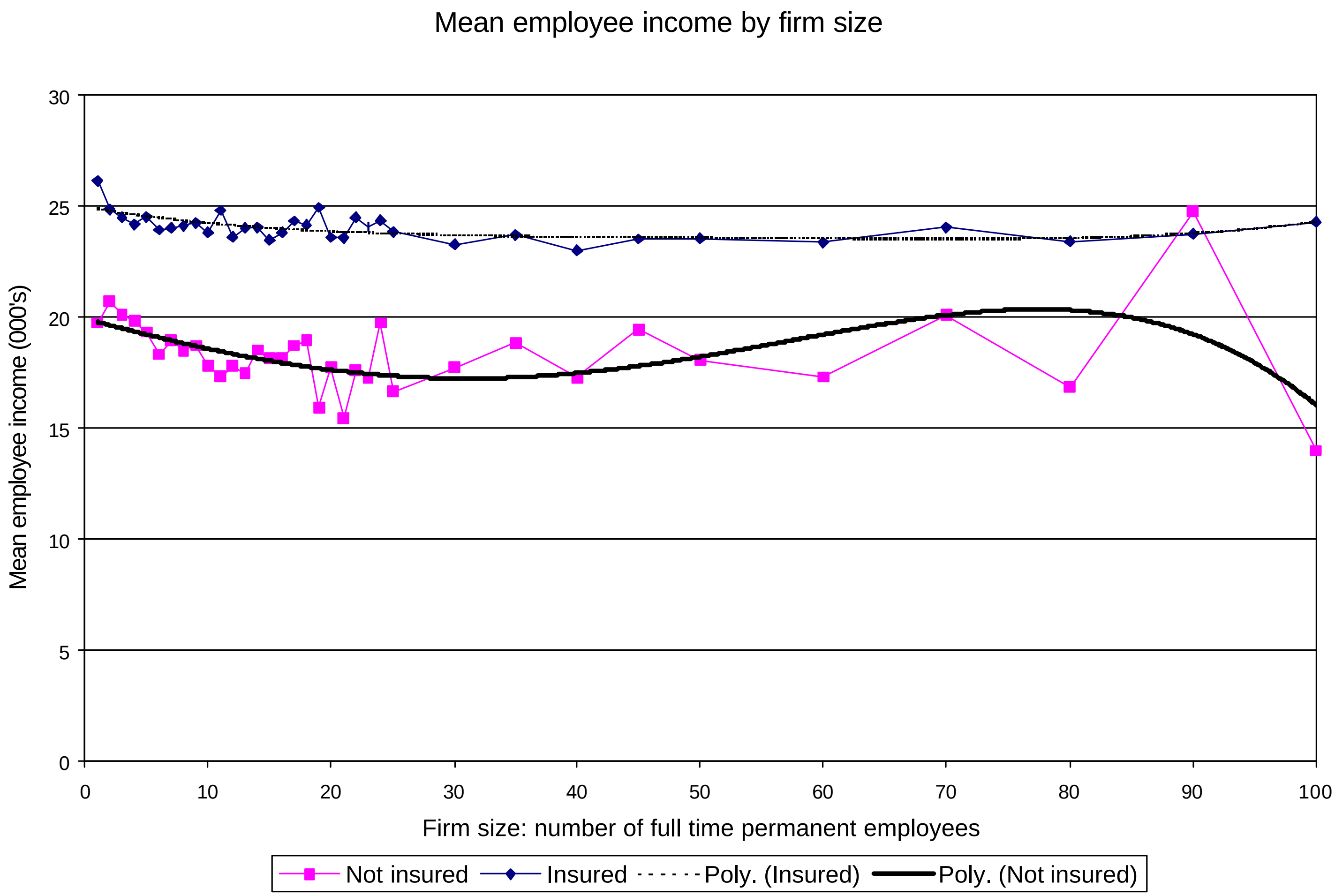

Figure 10. Mean employee income by collapsed firm size. RWJ 1997 EHIS data, N=18,712. Firm mean employee income was calculated using income proportions for each income interval (<10k, 10-14k, 14-20k, 20-25k, 25-35k, 35k+) multiplied by midpoints (8, 12, 17, 22.5, 30, 40). Collapsed firm sizes were generated as previously discussed. These firm specific means were then averaged for each collapsed firm size separately for firms offering and not offering insurance. Polynomial trend lines are bold for uninsured firms, dashed for insured firms.

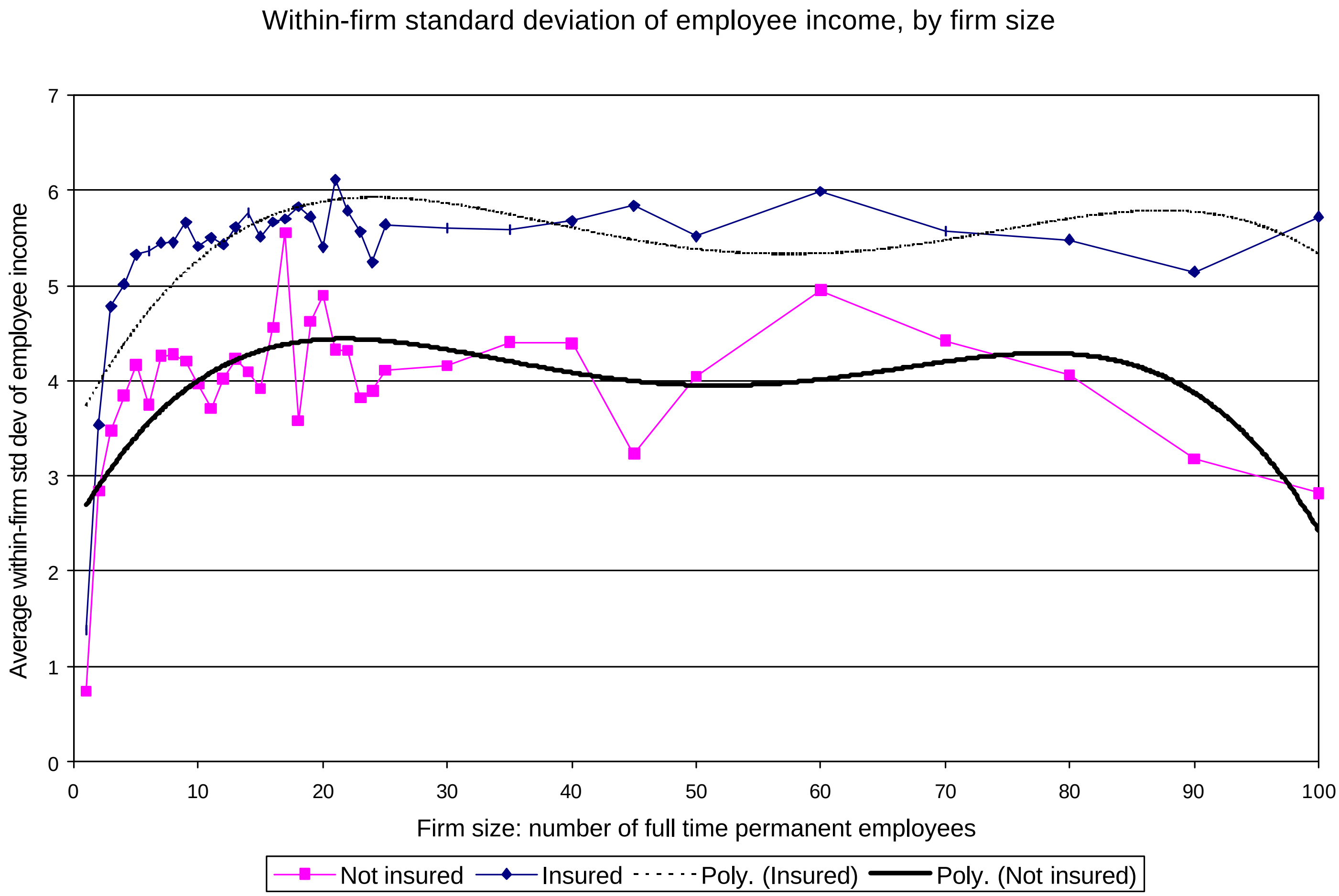

Figure 11. Within-firm standard deviations in average employee income, by collapsed firm size. RWJ 1997 EHIS data, N=18,712. Firm average income was calculated using income proportions for each income interval (<10k, 10-14k, 14-20k, 20-30k, 30k+) multiplied by midpoints or assumed values for endpoints (8, 12, 17, 25, 35). Collapsed firm sizes were generated as previously discussed. Standard deviations within each firm were calculated for using the weighted deviations of income from this firm specific average income. These standard deviations were then averaged for each collapsed firm size separately for firms offering and not offering insurance. Polynomial trend lines are bold for uninsured firms, dashed for insured firms.

Figure 12. Between firm standard deviations in average employee income, by collapsed firm size. RWJ 1997 EHIS data, N=18,712. Firm average income was calculated using income proportions for each income interval (<10k, 10-14k, 14-20k, 20-30k, 30k+) multiplied by midpoints or assumed values for endpoints (8, 12, 17, 25, 35). Collapsed firm sizes were generated as previously discussed. Standard deviations were calculated for each collapsed firm size separately for firms offering and not offering insurance. Polynomial trend lines are bold for uninsured firms, dashed for insured firms.

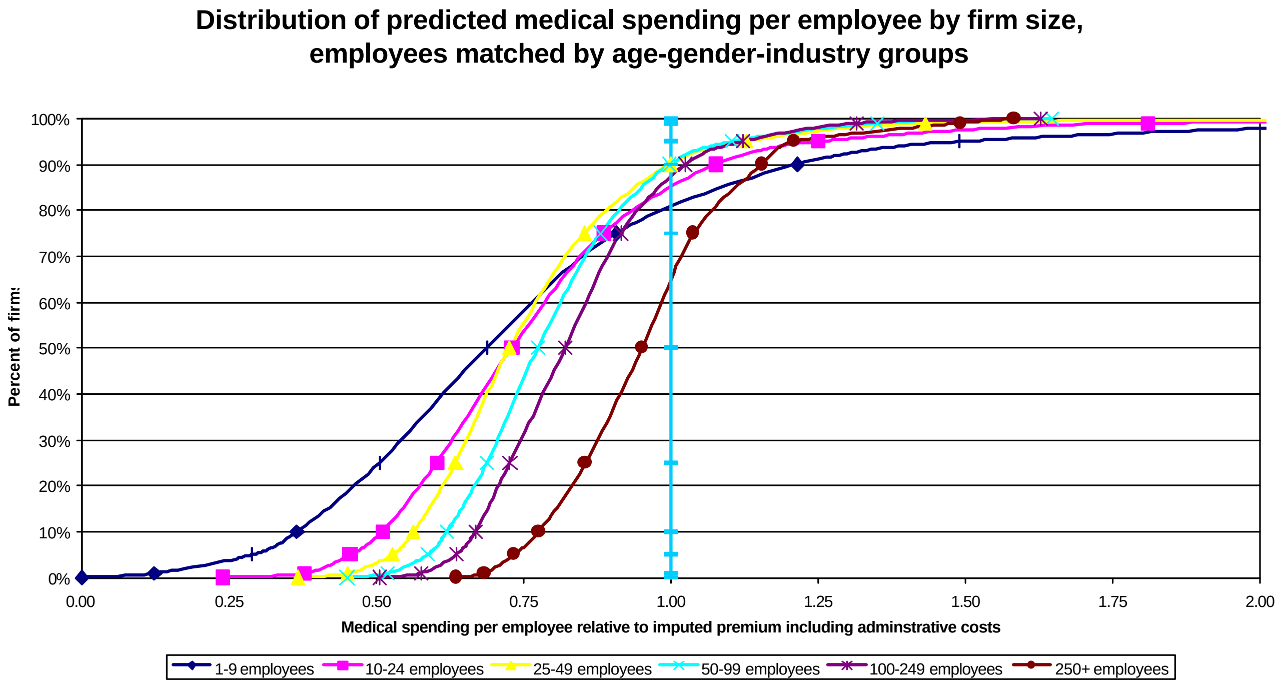

Figure 13. Distribution of predicted medical spending per employee by firm size categories, employees matched by age-gender-industry groups. These distributions were generated by combining RWJ 1997 EHIS and MEDSTAT 1998-99 data. 250 random samples of employees from the MEDSTAT data were drawn matching the age-gender-industry intervals of each firm appearing in the EHIS data. Predicted spending for each individual was generated using demographic and lagged spending splines as in Table 5. Imputed premiums were calculated assuming fully community rating. Imputed premiums are the grand mean average cost (1619) increased by average administrative cost percentages based on Chu and Trapnell (2002). The average predicted spending per employee for each draw was deflated by the average imputed premium appropriate for that firm size. The percentiles of this distribution for each firm size category were used to generate curves shown above.

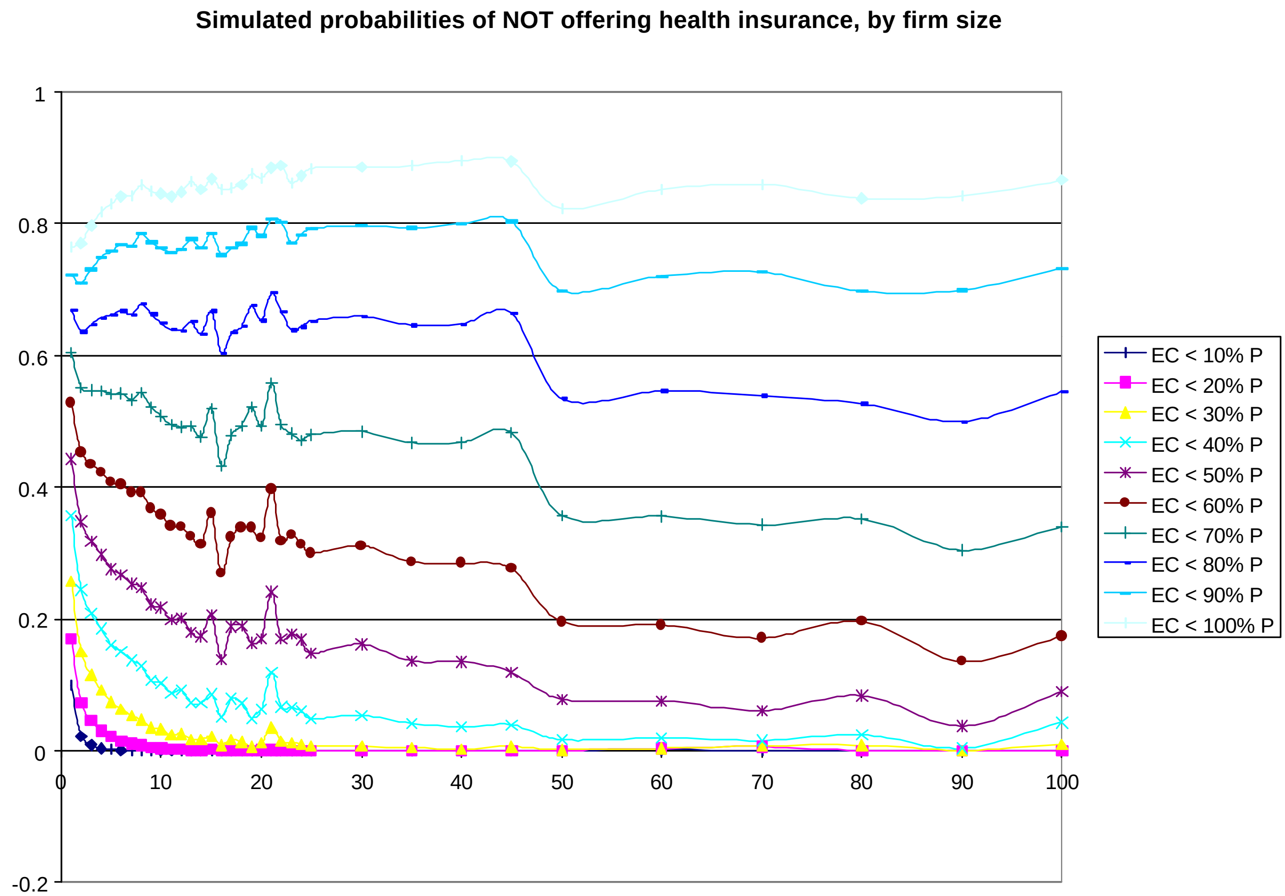

Figure 14. Simulated probabilities that firms do not offer insurance, by collapsed firm size. Generated by combining RWJ 1997 EHIS and MEDSTAT 1998-99 data. Collapsed firm sizes were generated as previously discussed. 250 random samples of employees from the MEDSTAT data matching the age-gender-industry intervals were drawn for each firm appearing in the EHIS data. For each random draw, the average predicted medical cost (the expected cost or EC) was calculated using age and prior year spending splines. Imputed premiums (P) were calculated assuming fully community rating. Imputed premiums are the grand mean average cost (1619) increased by average administrative cost percentages from Chu and Trapnell (2002). Each curve plots proportions of the random draws for firms in that size interval that would NOT choose to offer insurance if the decision rule is to not buy if EC < á P where á takes on values from 10% to 100%.

Actual versus simulated probabilities of NOT offering health insurance, by firm size

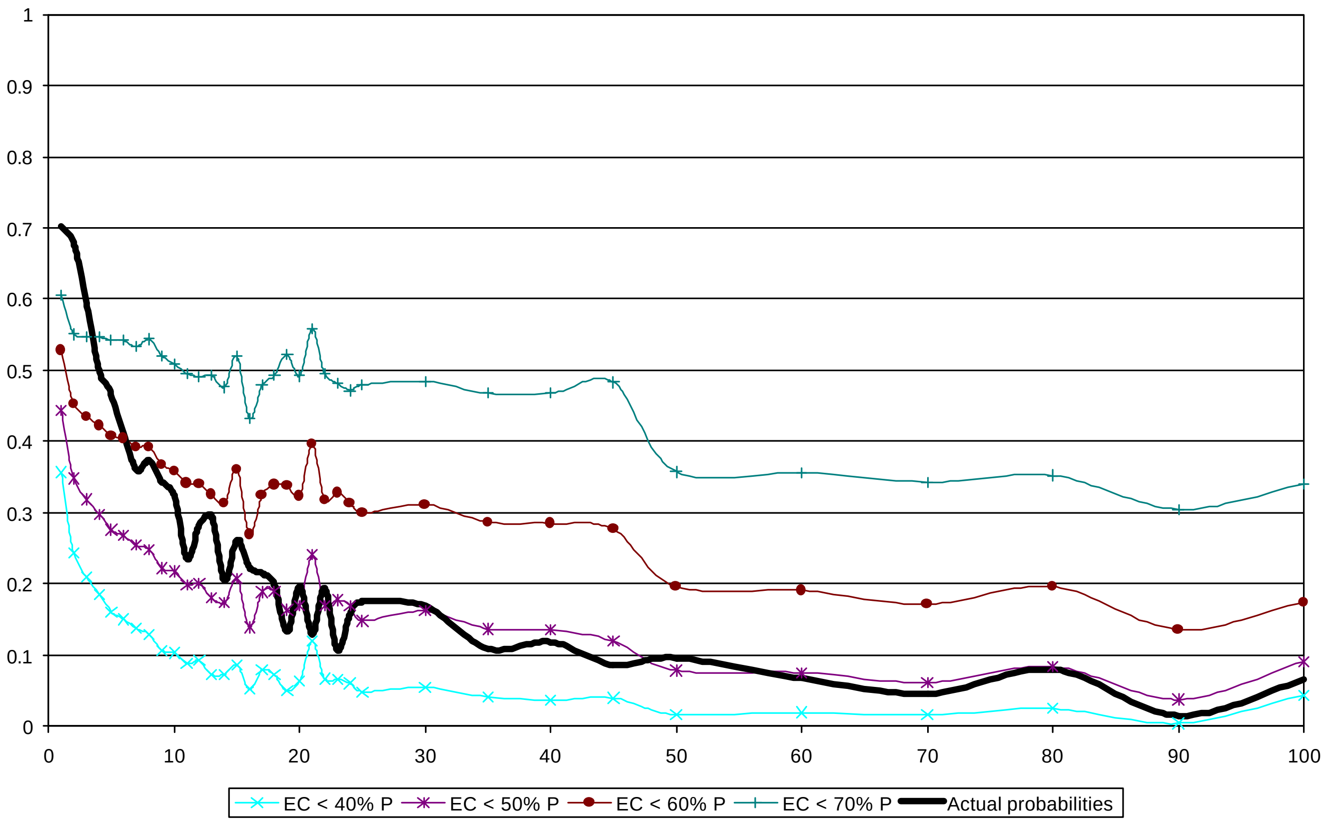

Figure 15. Actual versus simulated probabilities that firms do not offer insurance, by collapsed firm size. Collapsed firm sizes were generated as previously discussed. Actual curve is the same as in Figure 14, based on RWJ 1997 EHIS data. Simulated probabilities were generated by combining RWJ 1997 EHIS and MEDSTAT 1998-99 data. 250 random samples of employees from the MEDSTAT data matching the age-gender-industry intervals were drawn for each firm appearing in the EHIS data. For each random draw, the average predicted medical cost (the expected cost or EC) was calculated using age and prior year spending splines. Imputed premiums (P) were calculated assuming fully community rating. Imputed premiums are the grand mean average cost (1619) increased by average administrative cost percentages based on Chu and Trapnell (2002). Each curve plots proportions of the random draws for firms in that size interval that would not choose to offer insurance if the decision rule is to not buy if EC < á P where á takes on values from 40% to 70%.