Volume Title: Frontiers in the Economics of Aging

Volume Author/Editor: David A. Wise, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-90304-4

Volume URL: http://www.nber.org/books/wise98-1

Publication Date: January 1998

Chapter Title: The Impact of Intrafamily Correlations on the Viability of Catastrop
Insurance

Chapter Author: Matthew Eichner

Chapter URL: http://www.nber.org/chapters/c7303

Chapter pages in book: (p. 275 - 300)

6 The Impact of Intrafamily Correlations on the Viability of Catastrophic Insurance

Matthew J. Eichner

This paper explores the relationship between health care expenditures of spouses. If expenditures are due largely to random shocks, the household's medical expenditures are smoothed when two or more family members each draw from the distribution. On the other hand, if shocks are positively correlated, the potential exists for negative wealth shocks that are greater than those that would be predicted based on studies of the persistence of individual medical expenditures over time. Under traditional systems of insurance, with relatively low coinsurance levels, the consequences of any putative positive correlation across family members in expenditures are relatively mild. But under the sort of high-deductible insurance that is currently attracting interest from the policy community, wealth and utility effects may be appreciable. And under systems that include both high-deductible insurance and medical savings accounts, intrafamily correlations might dramatically change the accumulation of wealth in such accounts over a working lifetime.

Intrafamily correlations might exist for medical, economic, or behavioral reasons. Certain conditions, such as those related to contagious diseases or automobile accidents, might affect multiple members of a family. In addition, and as described in Eichner (1997), spending by one family member under most traditional employer-provided insurance plans lowers the price of care faced by other family members and may thereby induce additional expenditure. Finally, there may be fundamental differences in behaviors across families. Some individuals may make choices about their lives that lead, in the short

Matthew J. Eichner is assistant professor of finance and economics at Columbia University Graduate School of Business and a faculty research fellow of the National Bureau of Economic Research.

The author thanks Jeffrey Brown, Jeffrey Kling, James Poterba, and David Wise for helpful comments, and the National Institute on Aging, the Sloan Foundation, and the National Bureau of Economic Research for financial support.

run or the long run, to lower health care expenditures. And such individuals may seek out, as partners, individuals with similar values. In extreme cases, behaviors exhibited by one family member may directly affect expenditures by a spouse. The recently documented consequences of secondhand smoke provide such an example. Equally important, individuals may display different levels of inclination to initiate contact with the health care system. And again, there is every reason to suppose that individuals with particular levels of "taste for medical care" will choose similar individuals as partners.

The empirical strategy used in this paper will seek to differentiate between these various explanations for intrafamily correlations. To do this, I will look at individuals covered under different sorts of insurance plans and their expenditures for different sorts of care. By comparing the correlation between spousal expenditures in plans with coinsurance to the correlation between spousal expenditures in plans without coinsurance, I can measure the extent of the price effect. And by comparing the correlation between spouses for different sorts of expenditures, I can at least begin to sort out how much of the observed correlation in expenditures is due to factors other than taste. For example, the correlation in total medical expenditures between spouses may well be driven by correlation in expenditures for routine care. But if I restrict attention to expenditures related to surgical interventions, the effects of contagious disease and certain taste parameters are likely eliminated.

I also seek to determine where in the distribution of expenditures the correlations appear to be strongest. For example, there may be large correlations at lower expenditure levels produced by taste parameters, price effects, or contagious diseases. But the correlation could diminish higher in the distribution. In other words, there may be a great deal of information about whether a husband spends nothing or $300 on medical care in whether his wife spends nothing or $300. But there is likely less information concerning the husband's spending patterns in whether the wife spends $300 or $30,000.

The remainder of this paper is divided into six sections. Section 6.1 uses a simple framework to demonstrate the effect of intrafamily correlations on the standard microeconomic insurance problem. The analysis suggests that substantial intrafamily correlations can dramatically alter the value to individuals of insurance against the risk of medical expenditures. Section 6.2 introduces the data consisting of claims records of employees of two Fortune 500 firms and their spouses. The critical distinction between these two firms is that the first provides medical coverage with standard coinsurance provisions consisting of copayments and deductibles, while the other provides first-dollar coverage of almost all expenditures. In order to measure correlations for specific sorts of expenditures, it is necessary to adopt a strategy for aggregating claims that are likely due to the same medical condition. Section 6.3 describes the strategy I adopt in this paper and provides some descriptive evidence that it accomplishes the stated goal. Section 6.4 introduces the basic econometric framework utilized in this work. Rather than trying to faithfully capture the extremely long right tail of the expenditure distribution, I seek to estimate the probability that an individual's expenditure falls above a particular threshold. And most important, I will explore whether the fact that an individual's expenditures exceed a threshold provides information about the spending of his or her partner. In section 6.5, I consider the possibility that correlations may change over a working lifetime as a function of the aging process. Finally, section 6.6 presents some illustrative utility calculations and discusses the implication of the estimation results for systems of medical savings accounts and catastrophic insurance.

6.1 Risk, Insurance, and Intrafamily Correlation

This section presents the traditional microeconomic insurance problem, but modified to allow for the possibility of losses that are correlated among family members covered under a single insurance policy. I will show that positive correlation in losses within families unambiguously increases the value to families of purchasing insurance. This is reflected in their willingness to pay larger spread, or risk premium, over the expected loss.

Consider first a single risk-averse individual with a concave utility function $U(\cdot)$ such that $U'(\cdot) > 0$ and $U''(\cdot) < 0$. The standard insurance problem considers the expected utility of this individual integrated over all possible losses, which are assumed to follow some density $f_L(l)$:

$$E\ U(M - L) = \int_L U(M - l) f_L(l)\, dl,$$

where M is income. Willingness to pay for insurance coverage is usually expressed in terms of a certainty equivalent, the amount an individual would pay with probability one to avoid drawing some realization l from the density $f_L(l)$. This certainty equivalent c is found by solving

$$U(M - c) = \int_L U(M - l) f_L(l)\, dl.$$

If the certainty equivalent c equals the expected value of the loss, then the individual is unwilling to pay for insurance. But if $U(\cdot)$ is concave as assumed above, c will always exceed the expected loss. As developed in Arrow (1971), the individual is then willing to pay a risk premium to avoid the uncertainty in drawing from the density $f_L(l)$.

Now I consider the same problem, but modified to consider, not an individual facing a loss with density function $f_L(l)$, but a married couple facing a realization of two losses l_1 and l_2 with the joint density function $f_{L_1,L_2}(l_1, l_2)$. The critical issue, of course, is how the joint density function affects the certainty equivalent and hence the risk premium that the family is willing to pay to avoid the uncertainty of drawing l_1 and l_2 from $f_{L_1,L_2}(l_1, l_2)$. In particular, the issue will be the correlation coefficient between l_1 and l_2 implied by the joint density function.

Suppose a second joint density $g_{L_1,L_2}(l_1, l_2)$ differs from $f_{L_1,L_2}(l_1, l_2)$ only in that the correlation between l_1 and l_2 is higher, so $\rho_g > \rho_f$. The expected loss L is the same under either of the functions, since $E(L) = E(l_1) + E(l_2)$ does not depend on the correlation between the two individual losses. But the variance of the family loss L, $V(L) = V(l_1) + V(l_2) + 2\rho\sigma_{l_1}\sigma_{l_2}$, is an increasing function of the intrafamily correlation ρ.

If both density functions $g_{L_1,L_2}(l_1, l_2)$ and $f_{L_1,L_2}(l_1, l_2)$ yield the same expected loss but the loss under $g_{L_1,L_2}(l_1, l_2)$ has higher variance than the loss under $f_{L_1,L_2}(l_1, l_2)$, a risk-averse family will always prefer to face density function $f_{L_1,L_2}(l_1, l_2)$ rather than $g_{L_1,L_2}(l_1, l_2)$. Both imply the same expected loss, but the variance under $g_{L_1,L_2}(l_1, l_2)$ is greater. This intuitive idea can be formalized using the idea of second-order stochastic dominance. The density $f_{L_1,L_2}(l_1, l_2)$ is said to second-order stochastically dominate $g_{L_1,L_2}(l_1, l_2)$ if both have the same expected value and

$$\int_0^{l_1}\int_0^{l_2} G_{L_1,L_2}(l_1, l_2)\,dl_2\,dl_1 \geq \int_0^{l_1}\int_0^{l_2} F_{L_1,L_2}(l_1, l_2)\,dl_2\,dl_1$$

for any l_1 and l_2, where G and F are the distribution functions corresponding to density functions g and f, respectively. This condition will always be satisfied if the variance of the total loss, the sum of l_1 and l_2, described by the density $g_{L_1,L_2}(l_1, l_2)$ exceeds the variance of the total loss described by $f_{L_1,L_2}(l_1, l_2)$. As shown in Rothschild and Stiglitz (1970), second-order stochastic dominance is sufficient to establish

$$\int_{L_1,L_2} f_{L_1,L_2}(l_1, l_2)U(Y - l_1 - l_2)\,dl_1\,dl_2 > \int_{L_1,L_2} g_{L_1,L_2}(l_1, l_2)U(Y - l_1 - l_2)\,dl_1\,dl_2$$

or, equivalently, that the expected utility is higher if the loss follows $f_{L_1,L_2}(l_1, l_2)$ rather than $g_{L_1,L_2}(l_1, l_2)$.

This result suggests that the utility consequences of insurance schemes, and particularly those that involve relatively high potential out-of-pocket payments, may look quite different when evaluated on a family rather than on an individual basis. In particular, schemes that may have acceptable consequences for the utility of a single individual under all states of nature might have the potential to impose ruinous losses on families if substantial intrafamily correlations in medical expenditures exist. The empirical analysis described in this paper is an effort to measure the extent of these potentially important intrafamily (and particularly spousal) correlations in medical expenditures.

6.2 Firm Claims Data

The empirical analysis described in this paper uses claims data from two Fortune 500 firms. The firms differ along several dimensions. Most important for this analysis, firm 1 offers its employees a plan that requires them to bear none of the cost of their medical care, while firm 2 offers its employees a

Table 6.1 **Summary Data Description**

	Firm 1	Firm 2
Mean employee spending in 1992 ($)	1,355	1,832
Standard deviation of employee spending in 1992 ($)	5,177	7,968
Mean spousal spending in 1992 ($)	1,176	2,084
Standard deviation of spousal spending in 1992 ($)	4,253	8,872
Mean employee age	43.7	41.0
Mean spouse age	41.2	40.1
Mean difference in ages (employee age minus spouse age)	2.58	0.94
Percentage of male employees	95.04	74.69
Number of married couples	13,273	6,313

choice of plans all of which incorporate more traditional coinsurance provisions. In addition, firm 1 employees are located at a single site, while firm 2 employees are spread over six locations. All of the insurance plans at both firms incorporate limited case management for certain high-cost medical conditions and concurrent reviews of hospital stays. None cover pharmacy charges, mental health treatment, substance abuse treatment, or dental care, which are all covered separately under "carve-out" arrangements.

Married couples were formed from all individuals filing claims between 1990 and 1993 who could be unambiguously matched with a spouse. The analysis was then restricted to those married couples formed by employees between ages 25 and 55 whose spouses were aged 20 to 60. In firm 1, 13,273 such couples were identified, while the corresponding number from firm 2 was 6,313. Table 6.1 shows some basic descriptive information about the data. Thus, on average, the employee at firm 2 is younger, is more likely to be female, and spends more in 1992.

Claims data provide excellent information on the timing and nature of contacts with the health care system, as well as who in the family received treatment. Claims data are less useful in determining who was covered by the insurance plan during a particular period. Only by observing an individual filing a claim do I know that an individual is present. And when an individual does not file a claim during a particular period, it is presently impossible to discern whether the individual has received no treatment or has separated from the firm. I hope that the impact of this issue is minimized by studying firms that have stable workforces and individuals who have not yet reached ages at which retirement is likely. To address this issue more directly, additional enrollment data is required that will become available in the near future.

6.3 Classification of Claims

Most contacts with the health care system generate multiple claims. For example, the individual who visits the doctor complaining of stomach pains might receive an examination, blood tests, and an X-ray. Each of these may generate separate and, in some cases, multiple claims. Each claim may carry a diagnosis code, a procedure code, both, or neither. Continuing with the above example, the examination may be coded for a digestive disorder (diagnosis) and a medical procedure (examination), while the blood test may be coded as a laboratory procedure without a diagnosis. The empirical work described in this paper seeks to demonstrate that, for certain types of care, intrafamily correlations are greater than for other types of care. To do this requires grouping claims together by procedure and diagnosis in a sensible way, so that all claims stemming from the stomach pain scenario sketched above can be classified as a single treatment episode.

The idea of grouping claims to form treatment episodes dates to the RAND experiment in the 1970s. As described in Newhouse (1993), clinicians participating in the study made notes that explicitly indicated which claims were related. This information, along with the chronological spacing of treatments, allowed all charges to be grouped that "reflect[ed] decisions about the same medical problem." These episodes were then assigned to one of four categories: hospital, physician and supplies, dental, and pharmacy.

Because claims data lack the explicit links relied upon by the RAND investigators, I have taken a different approach. Furthermore, since the claims data explicitly exclude pharmacy and dental charges, their classification of episodes into one of four categories does not lend itself to my purpose. Instead, I have adopted an algorithm using exclusively chronological criteria to group episodes. The algorithm is fairly simple, yet as I will argue, it produces episodes that are internally consistent, as well as consistent with reasonable priors about health care delivery.

Under this typology, an outpatient episode begins on the date an outpatient treatment is received, and a window of 14 days is opened. Any further claims during that period are grouped with the initial treatment. When each additional treatment is received, the window is extended for an additional two weeks from the date of service. When two weeks pass during which no additional claims are filed, the outpatient episode is deemed complete. The procedure with respect to inpatient episodes is similar, although the window is 28 rather than 14 days in length. If an episode begins with an outpatient claim, a subsequent inpatient claim for treatment lengthens the window measured from all further claims in the episode from 14 to 28 days.

Once episodes are defined, I then group them using a hierarchy of the ICD-9 and CPT codings, which designate particular diagnoses and procedures, respectively. At the top of the hierarchy are episodes related to injuries and poisonings, as reflected by the presence in the grouping of at least one claim with an ICD-9 code corresponding to such a diagnosis. The next level consists of those episodes containing at least one claim referencing an inpatient or outpatient surgical procedure. Lower still are episodes with an inpatient or outpatient medical procedure, again as indicated by the presence of an appropriately coded claim. Another level consists of episodes that consist of at least one claim indicative of a diagnostic test. And, finally, there is the residual.

Table 6.2 **Treatment Episodes by Type**

	Firm 1			Firm 2		
Episode Type	Percentage of Episodes	Percentage of Cost	Mean Cost ($)	Percentage of Episodes	Percentage of Cost	Mean Cost ($)
Injury/poisoning	12.47	21.46	1,526	10.93	24.81	1,863
Surgical	18.78	43.40	2,068	18.15	48.46	2,191
Medical	35.41	26.46	669	49.35	19.56	325
Diagnostic	18.62	4.54	218	8.80	2.95	276
Residual	14.71	4.34	264	12.77	4.22	271

Table 6.3 **Treatment Episodes by Length**

	Firm 1			Firm 2		
Episode Length	Percentage of Episodes	Percentage of Cost	Mean Cost ($)	Percentage of Episodes	Percentage of Cost	Mean Cost ($)
1 Day	66.21	14.20	192	61.69	11.96	159
< 1 Week	10.79	14.93	1,238	10.79	11.21	852
< 2 Weeks	10.28	14.54	1,266	11.40	11.23	807
< 3 Weeks	5.17	9.34	1,619	6.25	8.70	1,141
< 4 Weeks	2.19	7.20	2,943	2.97	6.80	1,878
< 6 Weeks	2.39	10.53	3,937	2.97	9.31	2,575
< 8 Weeks	1.01	5.63	4,988	1.34	6.24	3,835
< 3 Months	1.09	8.93	7,348	1.45	10.67	6,049
> 3 Months	0.88	14.70	15,027	1.15	23.89	17,109

Note that in this typology each level subsumes the levels below. An individual receiving only a cardiogram would produce an episode coded as a diagnostic procedure. If he or she had a cardiogram and also a medical examination, the episode would be categorized as medical. A cardiogram and medical examination followed by open heart surgery would be classified as a surgical episode.

Table 6.2 shows the breakdown of episodes constructed from claims filed by employees and spouses at the two firms during the 1990–92 period. In both firms, injury/poisoning and surgical episodes account for a disproportionate share of expenditures. For example, only 18 percent of the episodes of treatment for firm 2 employees are classified as surgical, but these account for almost one-half of expenditures. On the other hand, the more numerous medical episodes account for less of the total cost. Again referring to firm 2, the 49 percent of episodes classified as medical account for only 20 percent of expenditures. In both firms, the importance of the so-called residual episodes, those that cannot be classified elsewhere in the hierarchy, is relatively small as the cost share of these episodes is below 5 percent.

Table 6.3 describes the treatment episodes, defined using the algorithm and classifications outlined above, in terms of their length, that is, the number of days from the treatment that initiates the episode to the final treatment included in the grouping. The distribution of episode lengths is quite similar across the two firms. And, as in table 6.2, the category that contains most of the episodes accounts for a relatively small percentage of the cost.

6.4 Estimation of Intrafamily Correlations

A primary goal of this work is to assess empirically the magnitude of intrafamily correlations in medical expenditures. In this section, I outline three possible approaches to this analysis and describe the bivariate probit model that I choose to apply. I then show how the results of the bivariate probit analysis can be easily interpreted using the concept of conditional probability.

One option in measuring the correlation is to estimate a system of two regressions, one for the employee and one for the spouse, and to allow some covariance structure between the equations. This sort of estimation is generally referred to as seemingly unrelated least squares (SUR). The problem is fundamentally complicated by the fact that both equations, employee expenditures and spousal expenditures, are censored, with a substantial proportion of individuals spending nothing. Estimation of an SUR system with censored dependent variables requires restrictive assumptions concerning the distribution of both expenditures and the disturbance terms and would still prove computationally intractable because construction of the likelihood function would require evaluating "hybrid" density and distribution functions for the assumed bivariate distribution.

Another approach estimates the joint distribution function at particular points. For example, controlling for demographic factors, I might estimate the probability that the spouse spends more than $1,000 while the employee spends more than zero. This is done using the bivariate probit specification. The assumption of joint normality is relatively harmless since I do not need to claim that I accurately capture the shape of the probability mass between any two points of support. The probability in which I am interested is the only parameter of the relevant Bernoulli distribution. In other words, I care only about whether expenditures are above or below some threshold. The exact nature of the distribution above and below this threshold is unimportant; only how much of the density is above and below the threshold is critical.

A third possibility involves a bivariate ordered probit model. Instead of estimating the joint probability that the spouse's expenditures exceed some level and the employee's expenditures exceed some level, I would estimate the joint probability that the spouse's expenditures fall in a particular range and that the employee's expenditures fall in a particular range. Here the assumption of joint normality is more restrictive, since the outcome of interest is not simply whether a realization falls above or below some threshold.

Another troublesome implication of the bivariate ordered probit is that the correlation is constrained to be equal across the entire expenditure distribution.

One might imagine that positive expenditures by a spouse are correlated with positive expenditures by an employee. But it is unlikely that the correlation persists higher in the distribution and that expenditures above \$1,000 by a spouse are as correlated with expenditures above \$1,000 by an employee. The results are therefore very sensitive to the choice of expenditure ranges. I will opt to avoid this issue by using the simple bivariate probit framework to estimate the probabilities of exceeding certain points in the joint distribution.

A simple probit model of expenditures is written

$$P(Y > y) = P(Z\beta + \varepsilon > 0),$$

where y is a particular threshold of interest (often zero) and Z is a vector of independent variables including a constant. Making the distributional assumption that ε follows a normal distribution, the relation can be rewritten

$$P(Y > y) = \Phi(Z\beta),$$

where Φ is the distribution function that gives the probability in the lower tail of the normal distribution.

This framework is easily modified to deal with a system of two equations. Suppose

$$P(Y_1 > y_1 \text{ and } Y_2 > y_2) = P(\beta_1 Z_1 + \varepsilon_1 > 0 \text{ and } \beta_2 Z_2 + \varepsilon_2 > 0),$$

where the subscripts 1 and 2 refer to the equations. The distributional assumption is then

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right),$$

and the relevant probability can be written

$$P(Y_1 > y_1 \text{ and } Y_2 > y_2) = \Phi_2(\beta_1 Z_1, \beta_2 Z_2, \rho),$$

where Φ_2 is the distribution function for the bivariate normal distribution written above.

The parameters β_1, β_2, and ρ can be estimated using standard maximum likelihood methods. The exact specifications for firm 1 are

$$Z_i\beta_i = \alpha_i + \beta_{i1}\text{Age}_i + \beta_{i2}\text{Male}_i + \beta_{i3}\text{Age}_i * \text{Male}_i.$$

For firm 2, five location and two plan indicators are added for each member of the couple:

$$Z_i\beta_i = \alpha_i + \beta_{i1}\text{Age}_i + \beta_{i2}\text{Male}_i + \beta_{i3}\text{Age}_i * \text{Male}_i$$

$$+ \beta_{i4}1(\text{Plan} = 1) + \beta_{i5}1(\text{Plan} = 3) + \sum_{k=1}^{5} \beta_{i(k+5)}1(\text{Loc} = k).$$

Table 6.4 Bivariate Probit Estimates

	Firm 1		Firm 2	
Variable	$P(Y_i > 0)$	$P(Y_i > 500)$	$P(Y_i > 0)$	$P(Y_i > 500)$
Employee				
Age	0.0044	0.0087	0.0198**	0.0213**
	(0.0070)	(0.0068)	(0.0040)	(0.0039)
Male	−0.8722**	−0.7100*	−0.3841*	−0.6720**
	(0.3035)	(0.2998)	(0.1815)	(0.1845)
Age*Male	0.0113	0.0063	0.0043	0.0062
	(0.0072)	(0.0071)	(0.0045)	(0.0045)
Constant	0.3362	−0.5680	−0.3036	−0.8075
Spouse				
Age	0.0227**	0.0167**	0.0211**	0.0153**
	(0.0015)	(0.0016)	(0.0022)	(0.0022)
Male	−0.2051	−0.5640	−0.5999**	−0.8340**
	(0.2780)	(0.3325)	(0.1697)	(0.1877)
Age*Male	−0.0058	0.0056	−0.0017	−0.0056
	(0.0062)	(0.0073)	(0.0040)	(0.0044)
Constant	−0.7805	−1.3413	0.6905	−0.7984
Rho	0.4097**	0.2183**	0.6905**	0.3995**
	(0.0122)	(0.0151)	(0.0133)	(0.0189)
Other covariates	None	None	Five location and two plan indicators	Five location and two plan indicators

Note: Numbers in parentheses are standard errors.
*Statistically distinct from zero ($p < 0.05$).
**Statistically distinct from zero ($p < 0.01$).

The aim is to ensure that ρ, the parameter of interest, does not capture correlation related to both employee and spouse living in the same location, receiving care under the same plan, or being roughly the same age.

I begin by considering whether the sum of expenditures for all treatment episodes exceeds zero. I then continue by increasing the threshold from zero to $500 and then to $1,000. Results from this estimation are shown in table 6.4. The correlations between employees and spouses are large and quite precisely measured. In addition, the correlation decreases at higher levels of expenditure. For firm 1 couples, the correlation drops from 0.41 to 0.22 when the threshold is raised from zero to $500. A similar pattern is evident for firm 2 couples. This can be explained by the decreasing importance of taste parameters higher in the distribution. Whether a spouse chooses to visit the doctor with a common cold is likely correlated with whether an employee would make a similar choice. But whether a spouse would spend larger sums is likely less a function of the taste parameters that account for at least some of measured intrafamily correlations.

Finally, the estimated correlations for firm 1 are appreciably smaller than for firm 2. An obvious explanation for this regularity involves price effects that are present in firm 2 but not in firm 1. Recall that firm 2, unlike firm 1, requires coinsurance payments for standard medical care. These coinsurance payments take the form of a deductible that must be satisfied before insurance payments begin as well a copayment that continues from the point at which the deductible is met to some stop loss limit. As a result of the deductible and copayment requirements, each family faces a nonlinear price schedule for medical care, with the price falling once the deductible is satisfied and again once the stop loss limit is reached. Since the coinsurance provisions apply to the family as a whole, spending by one spouse can reduce, under plans like those in firm 2, the price of care faced by the other spouse.

Table 6.5 shows the correlation coefficients from additional estimation of joint probability distributions. Specifications are similar to those in table 6.4, but with additional expenditure thresholds and varying subsets of claims. As in table 6.4, the correlations are uniformly higher for couples covered by firm 2, suggesting the existence of substantial price effects. Furthermore, the correlations are largest when all episodes are included in the estimation. As episodes are restricted to those of the medical type, the correlations fall. Correlations between employee and spousal spending for surgical episodes, which are closer to the top of the hierarchy introduced in section 6.3, fall further. This suggests that moving up the hierarchy reduces the influence of taste parameters relative to external factors such as health shocks and provider decisions. Estimation results related to injury/poisoning episodes are an exception to this pattern, largely because a substantial fraction of injury/poisonings affect multiple family members.[1]

Increasing the threshold for each type of episode decreases the estimated correlation. Again, this is related to the relative importance of taste parameters as determinants of spending for different levels of expenditure. Both spouses may well opt to have the dermatologist remove benign growths within the same time frame. Expenditures for surgical removal of malignant tumors are much less likely to be correlated. Thus the correlation of 0.1628 for spending on surgical episodes above zero falls to essentially zero for spending on surgical episodes above $4,000.

To facilitate interpretation of these estimated correlations, it is useful to express them as conditional probabilities. This is done using the basic definition of conditional probability:

1. These results may seem to contradict the arguments presented in Eichner (1997) concerning the exogeneity of injuries and poisonings to family members. But those results were predicated on excluding all injury and poisoning claims that affected multiple family members. In addition, much of the exogenous variation was produced by injuries and poisonings to children, who are not included in the present analysis.

Table 6.5 **Estimated Correlations**

Spending	Firm 1	Firm 2
	All Episodes	
Above zero	0.4097	0.6905
	(0.0121)	(0.0133)
Above $500	0.2183	0.3995
	(0.0151)	(0.0189)
Above $1,000	0.2036	0.3150
	(0.0173)	(0.0215)
Above $2,000	0.1434	0.2200
	(0.0214)	(0.0261)
Above $4,000	0.1273	0.1396
	(0.0272)	(0.0351)
	Medical Episodes	
Above zero	0.2724	0.5730
	(0.0145)	(0.0159)
Above $500	0.1997	0.3022
	(0.0231)	(0.0256)
Above $1,000	0.1705	0.2635
	(0.0318)	(0.0333)
Above $2,000	0.1180	0.2221
	(0.0462)	(0.0523)
Above $4,000	0.1578	0.1111
	(0.0583)	(0.1168)
	Surgical Episodes	
Above zero	0.1628	0.3182
	(0.0177)	(0.0216)
Above $500	0.1283	0.2328
	(0.0241)	(0.0274)
Above $1,000	0.0914	0.1400
	(0.0278)	(0.0333)
Above $2,000	0.0688	0.0873
	(0.0336)	(0.0415)
Above $4,000	−0.0082	0.0447
	(0.0482)	(0.0552)
	Injury/Poisoning Episodes	
Above zero	0.2515	0.2732
	(0.0188)	(0.0271)
Above $500	0.2349	0.2312
	(0.0333)	(0.0373)
Above $1,000	0.1974	0.2039
	(0.0445)	(0.0455)
Above $2,000	0.0992	0.2747
	(0.0644)	(0.0540)
Above $4,000	0.1085	0.2399
	(0.0801)	(0.0727)

Note: Numbers in parentheses are standard errors.

$$P(Y_e > y \mid Y_s > y) = \frac{P(Y_e > y \text{ and } Y_s > y)}{P(Y_s > y)} = \frac{\Phi_2(Z_e\beta_e,\ Z_s\beta_s,\ \rho)}{\Phi(Z_s\beta_s)},$$

where the subscript e refers to the employee and the subscript s to the spouse. The conditional probability is evaluated for a couple with particular demographic characteristics. Here I will consider a 45-year-old male employee married to a woman of the same age. Table 6.6 shows the conditional probabilities for different thresholds and subsets of episodes and again makes clear that the correlations are strongest in the lower portions of the distributions. For example, a spouse with positive expenditures increases the probability of the employee's incurring positive expenditures by 0.2689. But a spouse with expenditures exceeding $2,000 only increases the probability of the employee's exceeding that level by 0.06.

6.5 Age-Dependent Correlations

In evaluating the effects of medical savings accounts in combination with catastrophic health insurance, a critical issue involves accumulation of balances in the accounts over a working lifetime. Intrafamily correlations, which potentially depend on the age of the couple, could have major effects on accumulation, particularly if the correlation increases with age so that the relation is strongest during the years in which the largest medical expenditures are likely to occur. This might occur if the effect of certain behavior decisions made over a lifetime produce increasingly important health consequences as a couple ages. For example, smoking (either first- or secondhand) is surely such a behavioral factor, and the deleterious consequences of smoking become manifest, not when individuals are in their 20s and 30s, but when they have reached more advanced stages of life.

I can easily reparametrize the bivariate probit model to allow for the dependence of the correlation coefficient on the age of the employee:

$$P(Y_e > y \text{ and } Y_s > y) = \Phi_2(\beta_e Z_e,\ \beta_s Z_s,\ \rho(\text{Age})),$$

where $\rho(\text{Age})$ is a simple linear function:

$$\rho(\text{Age}) = \gamma_0 + \gamma_{\text{Age}} \cdot$$

Estimates of β_e, β_s, and γ are obtained as before using maximum likelihood.

The results of this estimation, presented in table 6.7, show that correlation decreases with employee age, although in most cases the coefficient on age is not statistically distinct from zero. For example, each additional year is predicted to decrease the correlation in the bivariate normal distribution for positive expenditures in firm 1 by -0.0016. As shown in table 6.8, the implied correlations for couples with employees of ages 30, 40, and 50 are 0.4334, 0.4174, and 0.4014, respectively.

Table 6.6 **Conditional Probabilities**

	Firm 1	Firm 2
All Episodes		
$P(Y_e > 0 \mid Y_s > 0)$	0.6768	0.8118
$P(Y_e > 0 \mid Y_s < 0)$	0.4079	0.3303
$P(Y_e > 500 \mid Y_s > 500)$	0.3646	0.5441
$P(Y_e > 500 \mid Y_s < 500)$	0.2380	0.2885
$P(Y_e > 1{,}000 \mid Y_s > 1{,}000)$	0.2745	0.4077
$P(Y_e > 1{,}000 \mid Y_s < 1{,}000)$	0.1699	0.2241
$P(Y_e > 2{,}000 \mid Y_s > 2{,}000)$	0.1761	0.2596
$P(Y_e > 2{,}000 \mid Y_s < 2{,}000)$	0.1161	0.1534
$P(Y_e > 4{,}000 \mid Y_s > 4{,}000)$	0.1193	0.1349
$P(Y_e > 4{,}000 \mid Y_s < 4{,}000)$	0.0765	0.0867
Medical Episodes		
$P(Y_e > 0 \mid Y_s > 0)$	0.4357	0.6938
$P(Y_e > 0 \mid Y_s < 0)$	0.2667	0.3044
$P(Y_e > 500 \mid Y_s > 500)$	0.1924	0.2955
$P(Y_e > 500 \mid Y_s < 500)$	0.1051	0.7045
$P(Y_e > 1{,}000 \mid Y_s > 1{,}000)$	0.1198	0.2060
$P(Y_e > 1{,}000 \mid Y_s < 1{,}000)$	0.0629	0.0949
$P(Y_e > 2{,}000 \mid Y_s > 2{,}000)$	0.0684	0.1122
$P(Y_e > 2{,}000 \mid Y_s < 2{,}000)$	0.0394	0.0463
$P(Y_e > 4{,}000 \mid Y_s > 4{,}000)$	0.0610	0.0339
$P(Y_e > 4{,}000 \mid Y_s < 4{,}000)$	0.0269	0.0179
Surgical Episodes		
$P(Y_e > 0 \mid Y_s > 0)$	0.2354	0.4376
$P(Y_e > 0 \mid Y_s < 0)$	0.1584	0.2457
$P(Y_e > 500 \mid Y_s > 500)$	0.1323	0.2516
$P(Y_e > 500 \mid Y_s < 500)$	0.0875	0.1420
$P(Y_e > 1{,}000 \mid Y_s > 1{,}000)$	0.0967	0.1562
$P(Y_e > 1{,}000 \mid Y_s < 1{,}000)$	0.0697	0.1035
$P(Y_e > 2{,}000 \mid Y_s > 2{,}000)$	0.0649	0.0909
$P(Y_e > 2{,}000 \mid Y_s < 2{,}000)$	0.0490	0.0665
$P(Y_e > 4{,}000 \mid Y_s > 4{,}000)$	0.0287	0.0488
$P(Y_e > 4{,}000 \mid Y_s < 4{,}000)$	0.0299	0.0402
Injury/Poisoning Episodes		
$P(Y_e > 0 \mid Y_s > 0)$	0.2811	0.2886
$P(Y_e > 0 \mid Y_s < 0)$	0.1480	0.1459
$P(Y_e > 500 \mid Y_s > 500)$	0.1226	0.1684
$P(Y_e > 500 \mid Y_s < 500)$	0.0485	0.0803
$P(Y_e > 1{,}000 \mid Y_s > 1{,}000)$	0.0806	0.1169
$P(Y_e > 1{,}000 \mid Y_s < 1{,}000)$	0.0324	0.0551
$P(Y_e > 2{,}000 \mid Y_s > 2{,}000)$	0.0375	0.1093
$P(Y_e > 2{,}000 \mid Y_s < 2{,}000)$	0.0216	0.0362
$P(Y_e > 4{,}000 \mid Y_s > 4{,}000)$	0.0304	0.0748
$P(Y_e > 4{,}000 \mid Y_s < 4{,}000)$	0.0159	0.0242

Table 6.7 **Bivariate Probit Estimates: Correlation Parametrized as a Function of Employee Age**

	Firm 1		Firm 2	
Variable	$P(Y_i > 0)$	$P(Y_i > 500)$	$P(Y_i > 0)$	$P(Y_i > 500)$
Employee				
Age	0.0053	0.0087	0.0196**	0.0211**
	(0.0070)	(0.0069)	(0.0040)	(0.0039)
Male	−0.8365**	−0.7134*	−0.3896*	−0.6748**
	(0.3045)	(0.3008)	(0.1816)	(0.1844)
Age*Male	0.0105	0.0063	0.0045	0.0063
	(0.0072)	(0.0071)	(0.0045)	(0.0045)
Constant	0.2992	−0.5653	−0.2966	−0.7988
Spouse				
Age	0.0227**	0.0166**	0.0211**	0.0152**
	(0.0015)	(0.0016)	(0.0022)	(0.0022)
Male	−0.2051	−0.5505	−0.5960**	−0.8299**
	(0.2781)	(0.3322)	(0.1698)	(0.1874)
Age*Male	−0.0059	0.0053	−0.0017	−0.0056
	(0.0062)	(0.0073)	(0.0040)	(0.0044)
Constant	−0.7814	−1.3381	0.5027	−0.7975
Rho				
Age	−0.0016	−0.0050*	−0.0007	−0.0037
	(0.0018)	(0.0023)	(0.0018)	(0.0022)
Constant	0.4814**	0.4393**	0.7177**	0.5547**
	(0.0814)	(0.1015)	(0.0814)	(0.0950)
Other covariates	None	None	Five location and two plain indicators	Five location and two plain indicators

Note: Numbers in parentheses are standard errors.
*Statistically distinct from zero ($p < 0.05$).
**Statistically distinct from zero ($p < 0.01$).

Thus correlations appear likely to be lower in the later years of life when big expenditures typically occur. Such a pattern is at least partially explained by the fact that when individuals are relatively young and healthy the taste parameters that are presumably at least partially jointly determined between partners are relatively more important than in later life, when health shocks and provider decisions play a larger role in determining health outlays. And, based on this empirical estimation, the cumulative effect of behavior over the life cycle appears to be swamped by the random shocks in the later stages of life.

Table 6.8 **Estimated Correlations Parametrized as a Function of Age: All Expenditures**

Spending	Age	Constant	ρ(30)	ρ(40)	ρ(50)
		Firm 1			
Above zero	−0.0016	0.4814**	0.4334	0.4174	0.4014
	(0.0018)	(0.0814)			
Above $500	−0.0050	0.4393**	0.2893	0.2393	0.1893
	(0.0023)	(0.1015)			
Above $1,000	−0.0006	0.2308*	0.2128	0.2068	0.2008
	(0.0026)	(0.1175)			
Above $2,000	−0.0010	0.1896	0.1596	0.1496	0.1396
	(0.0032)	(0.1444)			
Above $4,000	0.0046	−0.0798	0.0582	0.1042	0.1502
	(0.0041)	(0.1882)			
		Firm 2			
Above zero	−0.0007	0.7177**	0.6967	0.6897	0.6827
	(0.0017)	(0.0687)			
Above $500	−0.0037	0.5547**	0.4437	0.4067	0.3697
	(0.0022)	(0.0950)			
Above $1,000	−0.0074**	0.6274**	0.4054	0.3314	0.2574
	(0.0025)	(0.1080)			
Above $2,000	−0.0078*	0.5492**	0.3142	0.2372	0.1592
	(0.0031)	(0.1328)			
Above $4,000	−0.0076	0.4665**	0.2385	0.1625	0.0865
	(0.0042)	(0.1826)			

Note: Numbers in parentheses are standard errors.

*Statistically distinct from zero ($p < 0.05$).

**Statistically distinct from zero ($p < 0.01$).

6.6 Utility Consequences of Intrafamily Correlation

In this section, I will provide some illustrative calculations detailing the cost of correlation between partners under a prototypical catastrophic insurance plan. The key provision of the plan I consider is a $2,000 annual individual deductible so that a married couple can, at most, suffer a $4,000 out-of-pocket loss. And once again, I will focus on the case of a 45-year-old male employee and his wife of the same age. My empirical approach has thus far abstracted from the continuous nature of the expenditure distribution, and I will continue to do so in this context by assigning individuals to be either at or below the $2,000 individual annual deductible.

I will first consider the joint distribution defined by the following four probabilities: $P(Y_e \geq \$2{,}000$ and $Y_s \geq \$2{,}000)$, $P(Y_e \geq \$2{,}000$ and $Y_s < \$2{,}000)$, $P(Y_e < \$2{,}000$ and $Y_s < \$2{,}000)$, and $P(Y_e < \$2{,}000$ and $Y_s \geq \$2{,}000)$. This distribution is obtained from the previously estimated probit specification. For firm 1, this distribution can be written

	$Y_e < \$2,000$	$Y_e \geq \$2,000$
$Y_s < \$2,000$	0.7671 (0.0698)	0.1008 (0.0151)
$Y_s \geq \$2,000$	0.1089 (0.0670)	0.0233 (0.0193)

The standard errors appearing in parentheses below each estimated probability are obtained from the bivariate probit estimation results using the Taylor series approximation referred to as the delta method.

The analogous bivariate distribution for firm 2 is

	$Y_e < \$2,000$	$Y_e \geq \$2,000$
$Y_s < \$2,000$	0.6568 (0.0564)	0.1191 (0.0189)
$Y_s \geq \$2,000$	0.1660 (0.0527)	0.0582 (0.0174)

For each firm, I also construct a second bivariate distribution that has the same expected value but correlation equal to zero. This is done by applying a basic definition of independence and taking the product of the appropriate marginal distributions obtained by summing over the joint distributions shown above. For firm 1, this distribution is

	$Y_e < \$2,000$	$Y_e \geq \$2,000$
$Y_s < \$2,000$	0.7602 (0.0712)	0.1077 (0.0110)
$Y_s \geq \$2,000$	0.1157 (0.0712)	0.0164 (0.0159)

while for firm 2 the tabular representation is

	$Y_e < \$2,000$	$Y_e \geq \$2,000$
$Y_s < \$2,000$	0.6383 (0.0577)	0.1375 (0.0148)
$Y_s \geq \$2,000$	0.1844 (0.0573)	0.0397 (0.0151)

For each firm, I also calculate the mean expenditure for the appropriate ages and genders conditional on expenditures below \$2,000. For firm 1, this conditional expectation is \$340 for the employee and \$374 for the spouse. The corresponding numbers for firm 2 are \$234 and \$240. I will take these figures to represent the loss when expenditures do not exceed \$2,000 per person. Using these figures, the expected out-of-pocket loss for a firm 1 couple is \$926, while the figure for firm 2 is \$1,375. Note that these expected losses are identical under both the actual ($\rho > 0$) and constructed ($\rho = 0$) distributions.

My approach will compare the amount individuals are willing to pay to avoid facing the actual distribution with the amount they will pay to eliminate the uncertainty embodied in the constructed distribution. The difference I will interpret as some measure in dollars of the utility cost of intrafamily corre-

Table 6.9 **Estimated Cost of Correlations**

Income		Risk Premium ($)	Risk Premium if ρ Equals Zero ($)	Percentage Change
		Firm 1		
$30,000	$U(M) = \ln M$	13.12	12.32	6.10
	$U(M) = M^{0.25}$	9.80	9.20	6.12
	$U(M) = M^{0.5}$	6.50	6.11	6.00
	$U(M) = M^{0.75}$	3.24	3.04	6.17
$50,000	$U(M) = \ln M$	7.66	7.20	6.00
	$U(M) = M^{0.25}$	5.73	5.39	5.93
	$U(M) = M^{0.5}$	3.81	3.58	6.04
	$U(M) = M^{0.75}$	1.90	1.79	5.79
$70,000	$U(M) = \ln M$	5.41	5.09	5.91
	$U(M) = M^{0.25}$	4.05	3.81	5.93
	$U(M) = M^{0.5}$	2.70	2.54	5.93
	$U(M) = M^{0.75}$	1.35	1.27	5.93
		Firm 2		
$30,000	$U(M) = \ln M$	17.33	15.47	10.73
	$U(M) = M^{0.25}$	12.95	11.57	10.66
	$U(M) = M^{0.5}$	8.60	7.69	10.58
	$U(M) = M^{0.75}$	4.28	3.83	10.51
$50,000	$U(M) = \ln M$	10.08	9.01	10.62
	$U(M) = M^{0.25}$	7.54	6.75	10.48
	$U(M) = M^{0.5}$	5.02	4.49	10.56
	$U(M) = M^{0.75}$	2.50	2.24	10.40
$70,000	$U(M) = \ln M$	7.11	6.36	10.56
	$U(M) = M^{0.25}$	5.32	4.76	10.53
	$U(M) = M^{0.5}$	3.54	3.17	10.45
	$U(M) = M^{0.75}$	1.77	1.58	10.73

lation. The first step is to calculate the certainty equivalent, the amount the couple would be willing to pay to avoid facing the distribution. This is done by solving for c_a and c_c so as to equate utility with and without uncertainty:

$$U(M - c_a) = \sum_1^4 U(M - l_e - l_s) f_{L_e,L_s}(l_e, l_s, \rho),$$

$$U(M - c_c) = \sum_1^4 U(M - l_e - l_s) f_{L_e,L_s}(l_e, l_s, 0).$$

Here M is again income, L_e the loss incurred by the employee, L_s the loss incurred by the spouse, $f_{L_e,L_s}(l_e, l_s, \rho)$ denotes the actual distribution, and $f_{L_e,L_s}(l_e, l_s, 0)$ denotes the constructed distribution. Risk premiums can then be calculated assuming $f_{L_e,L_s}(l_e, l_s, \rho)$ and $f_{L_e,L_s}(l_e, l_s, 0)$ by subtracting the certainty equivalent under each distribution from the expected loss.

Table 6.9 shows these risk premiums for a variety of income levels under several different utility functions. These are surprisingly small. Under log utility and income of $30,000, for example, the risk premium for the actual firm 1 distribution is $13.12 while the risk premium for the constructed distribution is $12.32. The difference of $0.80 represents the cost of correlation under this set of assumptions. In this case, the cost of correlation represents 6.1 percent of the risk premium under the actual ($\rho > 0$) distribution. While the risk premium decreases with increasing income as expected, the fraction related to intrafamily correlation remains steady at about 6 percent for firm 1. A similar pattern is evident for firm 2, although the higher correlations due to price effects boost the cost of intrafamily correlation to about 10.5 percent of the risk premium.

I can also use the technique described above to investigate the counterfactual that the level of intrafamily correlation around the threshold of zero persists into the higher ranges of the distribution. I do this by looking at the probabilities that expenditures exceed zero and assuming that these instead reflect, as above, the probabilities that expenditures exceed $2,000. Obviously, the risk premiums will be larger. But my primary interest is, not the risk premiums themselves, but how much of the risk premium can be attributed to intrafamily correlation of expenditures.

The actual distribution relevant for these calculations and firm 1 is

	$Y_e <$ \$0	$Y_e \geq$ \$0
$Y_s <$ \$0	0.2194	0.1008
	(0.0287)	(0.0151)
$Y_s \geq$ \$0	0.1264	0.5461
	(0.0309)	(0.0367)

The distribution constructed from the marginal distribution is

	$Y_e <$ \$0	$Y_e \geq$ \$0
$Y_s <$ \$0	0.1132	0.2143
	(0.0225)	(0.0420)
$Y_s \geq$ \$0	0.2325	0.4400
	(0.0234)	(0.0425)

The corresponding distributions for firm 2 are

	$Y_e <$ \$0	$Y_e \geq$ \$0
$Y_s <$ \$0	0.2394	0.1650
	(0.0502)	(0.0568)
$Y_s \geq$ \$0	0.1927	0.4030
	(0.0516)	(0.0583)

and

	$Y_e <$ \$0	$Y_e \geq$ \$0
$Y_s <$ \$0	0.1747	0.2297
	(0.0469)	(0.0615)
$Y_s \geq$ \$0	0.2573	0.3383
	(0.0470)	(0.0616)

Table 6.10 Estimated Cost of Correlations: Counterfactual Distribution

Income		Risk Premium ($)	Risk Premium if ρ Equals Zero ($)	Percentage Change
		Firm 1		
$30,000	$U(M) = \ln M$	34.52	27.35	20.78
	$U(M) = M^{0.25}$	25.92	20.53	20.81
	$U(M) = M^{0.5}$	17.30	13.69	20.84
	$U(M) = M^{0.75}$	8.66	6.85	20.87
$50,000	$U(M) = \ln M$	20.04	15.86	20.90
	$U(M) = M^{0.25}$	15.04	11.90	20.90
	$U(M) = M^{0.5}$	10.04	7.94	20.92
	$U(M) = M^{0.75}$	5.02	3.97	20.93
$70,000	$U(M) = \ln M$	14.12	11.17	20.92
	$U(M) = M^{0.25}$	10.60	8.38	20.93
	$U(M) = M^{0.5}$	7.07	5.59	20.95
	$U(M) = M^{0.75}$	3.54	2.79	20.96
		Firm 2		
$30,000	$U(M) = \ln M$	32.13	21.96	31.64
	$U(M) = M^{0.25}$	24.16	16.49	31.72
	$U(M) = M^{0.5}$	16.15	11.01	31.79
	$U(M) = M^{0.75}$	8.09	5.51	31.87
$50,000	$U(M) = \ln M$	18.64	12.70	31.89
	$U(M) = M^{0.25}$	14.00	5.93	31.93
	$U(M) = M^{0.5}$	9.35	6.36	31.97
	$U(M) = M^{0.75}$	4.68	3.18	32.06
$70,000	$U(M) = \ln M$	13.13	8.93	31.99
	$U(M) = M^{0.25}$	9.86	6.70	32.02
	$U(M) = M^{0.5}$	6.58	4.47	32.05
	$U(M) = M^{0.75}$	3.29	2.24	32.08

With these distributions in hand, I can repeat the calculation of risk premiums, which is shown in table 6.10. Moving to a counterfactually more correlated distribution increases the fraction of the risk premium attributable to intrafamily correlation. For firm 1, the fraction rises to 21 percent, while for firm 2, with the presence of price effects, the fraction reaches 32 percent. Note, however, that the magnitudes of the risk premiums remain, as before, very small.

6.7 Summary and Conclusion

The empirical work in this paper suggests that, while correlation of expenditures among married partners is large at the low end of the expenditure distribution, the relation diminishes appreciably in the upper ranges. For example, I estimate correlations for positive expenditures of 0.41 for firm 1 and 0.69 for firm 2. But the correlations in expenditures above $4,000 are only 0.13 and

0.14, respectively. Thus, while these correlations may be important in the lower ranges of the expenditure distribution, they decrease appreciably in the ranges relevant for a discussion of catastrophic health insurance schemes. Furthermore, the utility cost of this correlation is quite small as a percentage of the risk premium required to induce an individual to face the uncertainty in a prototypical major risk policy. For the firm 1 plan, which like a prototypical catastrophic plan is structured so that spending by one family member does not reduce the price for care paid by other family members, only about 6 percent of the risk premium is due to intrafamily correlation. Thus the existence of intrafamily correlation is unlikely to appreciably complicate the analysis or implementation of catastrophic insurance coverage.

More generally, the calculations in this paper reveal that the entire risk premium, and not just the portion related to intrafamily correlation, is quite small under a set of standard assumptions about the form of the utility function. The risk premiums for the prototypical catastrophic plan, even abstracting from any behavior response to the greater cost sharing, are no more than $20 in the most extreme case. Such small numbers, of course, raise more general questions about why individuals are consistently observed paying a great deal in premiums to avoid relatively small amounts of additional risk. Additional insight into this behavior will be a major goal of future work involving the broader issue of plan choice.

References

Arrow, Kenneth A. 1971. *Essays in the theory of risk bearing.* Chicago: Markham.

Eichner, Matthew J. 1997. Incentives, price expectations, and medical expenditures: An analysis of claims under employer-provided health insurance. New York: Columbia University, Graduate School of Business. Mimeograph.

Newhouse, Joseph P. 1993. *Free for all: Lessons from the RAND Health Insurance Experiment.* Cambridge, Mass.: Harvard University Press.

Rothschild, M., and J. Stiglitz. 1970. Increasing risk I: A definition. *Journal of Economic Theory* 2:225–43.

Comment Thomas J. Kane

Pooling risk within groups, such as firms or families, yields greater opportunity for utility gains when those risks are not positively correlated. In this chapter,

Thomas J. Kane is associate professor of public policy at the John F. Kennedy School of Government, Harvard University, and a faculty research fellow of the National Bureau of Economic Research.

Matt Eichner constructs a unique data set on family medical expenditures to study the implications of such correlation. He asks four questions: First, how correlated are medical expenditures within couples? Second, is that correlation different for large and small expenditures? Third, how is the degree of correlation related to deductibles and coinsurance? Fourth, what are the utility consequences of such correlation for the value of medical insurance?

The Data

The author has assembled an impressive set of claims data for two Fortune 500 firms, matching spouses' expenditures and aggregating the claims data by health episodes. As the author's recent work demonstrates, such data can prove extremely useful for studying a range of questions related to families' claims behavior and insurance purchasing behavior.

An interesting component of the data, which the author exploits in this paper as well as in other work, is the difference between the two firms' insurance plans. One firm requires various types of coinsurance (copayments, deductibles); the other does not. As long as the incentive structure is the only difference between the two firms that would be relevant for the intrafamily correlation in insurance behavior (one firm is not a major manufacturer of "bungee" jumping equipment, for instance), such variation can shed light on how couples behave when price incentives differ.

The Structure of the Problem

Insurance companies are worried not about the total variance in health expenditures but primarily about the variance that is not related to easily observable characteristics. If everyone in the sample had the same expected health expenditures, Eichner's problem would be the fairly straightforward exercise of describing the joint distribution of actual expenditures by couples. However, his work is complicated by the fact that easily observed demographic characteristics, such as age and gender, are clearly related to expenditures. Therefore, we are less interested in the distribution of total expenditures by each person (y_i and y_j) than in the distributions of the orthogonal components of health expenditures (e_i and e_j) below:

$$y_i = X_i\beta + e_i,$$
$$y_j = X_j\beta + e_j.$$

Faced with the task of having to estimate β in order to study e_i and e_j, Eichner chooses not to use a highly restrictive parametric form, such as the joint normal distribution with a single correlation parameter ρ, to describe the data. Rather, he specifies a series of thresholds ($500, $1,000, $2,000, and $4,000) and uses a bivariate probit estimator to model the likelihood that the expenditures of either member of the couple are above or below each threshold. Since joint normality of e_i and e_j would imply the same β and ρ regardless of the threshold used, Eichner's framework allows him to study correlation at various points in the joint distribution of expenditures, while providing a natural test for joint normality.

These results of the first part of the paper are summarized below:

1. For each type of expenditure, the estimated correlations were lower at higher expenditure thresholds, although the estimated correlations generally remained positive.

2. The intracouple correlations in expenditures were generally lower for surgical episodes than for other types of medical episodes.

3. The intracouple correlations at levels of expenditure below $4,000 tended to be greater in firm 2, which had fewer coinsurance features.

4. There was little evidence to suggest that intracouple correlations varied with the couple's age.

The author draws the plausible conclusion that shared tastes for medical care are more likely to produce covariance in small expenditures than in larger expenditures such as major surgeries. This conclusion is bolstered by the finding that the magnitude of intracouple correlation was higher at the firm with less coinsurance, primarily at lower spending thresholds.

Although the results are both plausible and interesting, I have two concerns with the first section of the paper. First, I could imagine that different forms of measurement error in health care expenditures would affect different parts of the joint distribution differently. For example, suppose that a couple's expenditures were truncated—such as when the employee left the firm, died, or started using his or her spouse's insurance. Any error that resulted in the simultaneous loss of both spouses' expenditure data would lead to greater correlation in expenditure at lower levels. Moreover, any errors in the assignment of claims to different types of episodes—surgical, medical, or injury/poisoning—could produce more "random" variation in various parts of the joint distribution, depending on the nature of the error.

Whether such errors in categorization are important would depend on the question being asked. For instance, an insurance company setting annual deductibles would not necessarily care about the impact of using an arbitrary accounting unit such as a year, which would produce both left- and right-censored expenditure episodes. For setting annual deductibles, it is the distribution of truncated or uncompleted spells that matters. However, for drawing conclusions about the influence of "tastes" on joint expenditures, sorting out the implications of various types of measurement error could be important.

Second, finding different correlation coefficients at different cut points may be a cause for concern—not just for the assumption that the distribution is joint normal throughout the range of expenditures, but even for whether the bivariate probit is the right specification. For instance, such findings may suggest that the joint distribution of e_i and e_j is also nonsymmetric. In future work, the author might experiment with other specifications of the distribution of e_i and e_j.

Utility Implications of Intracouple Correlation

In section 6.6 of the paper, the author explores the implications of intracouple correlation for the value of an insurance policy with a $2,000 per person deductible. He calculates the certainty equivalent "cost" of that correlation in the following way. First, he uses his estimates to calculate the probability that a couple with a given set of characteristics would find themselves in each of four quadrants, with the employee spending above or below $2,000 and the spouse spending above or below $2,000. (Call these probabilities P^{00}, P^{01}, P^{10}, and P^{11}.) He then calculates the mean expenditures in each of the four quadrants. (Call these e^{00}, e^{01}, e^{10}, and e^{11}.) Using a log utility function with income M, the author calculates the risk premium under correlated expenditures by solving for c_a:

$$U(M - c_a) = P^{00}U(M - e^{00}) + P^{01}U(M - e^{01}) + P^{10}U(M - e^{10}) + P^{11}U(M - e^{11}).$$

Having correlated medical expenditures means that the likelihood of falling into each of the quadrants is not equal to the product of the individual likelihoods that the employee or his spouse has expenditures above or below the relevant threshold. For example, the likelihood of both persons spending above $2,000 ($P^{11}$) is greater than the product of the probability that each individually has expenditures above $2,000. Such correlation can be costly for a couple, and Eichner seeks to estimate this cost. To do so, he then calculates a second certainty equivalent, c_c, to approximate the certainty equivalent with independent expenditures by solving the following equation:

$$\begin{aligned} U(M - c_c) = {} & (P^{00} + P^{01}) * (P^{00} + P^{10})U(M - e^{00}) + \\ & (P^{00} + P^{01}) * (P^{01} + P^{11})U(M - e^{01}) + \\ & (P^{00} + P^{10}) * (P^{10} + P^{11})U(M - e^{10}) + \\ & (P^{01} + P^{11}) * (P^{10} + P^{11})U(M - e^{11}). \end{aligned}$$

The difference in risk premiums, $c_c - c_a$, provides an estimate of the "cost" to families of having correlated health expenditures. At a family income of $30,000, Eichner estimates that intrafamily correlation raises the risk premium by roughly 10 percent over what two individuals with uncorrelated expenditures would have been willing to pay.

However, the above calculation probably misstates the cost of intracouple correlation for at least two reasons: First, the method calculates only the utility cost of movements of probability mass *between* the four quadrants; it does not take into account the utility cost of movements of probability mass *within* each of the four quadrants. Particularly because expenditures are correlated below $2,000, much of the cost of the correlation presumably results from the in-

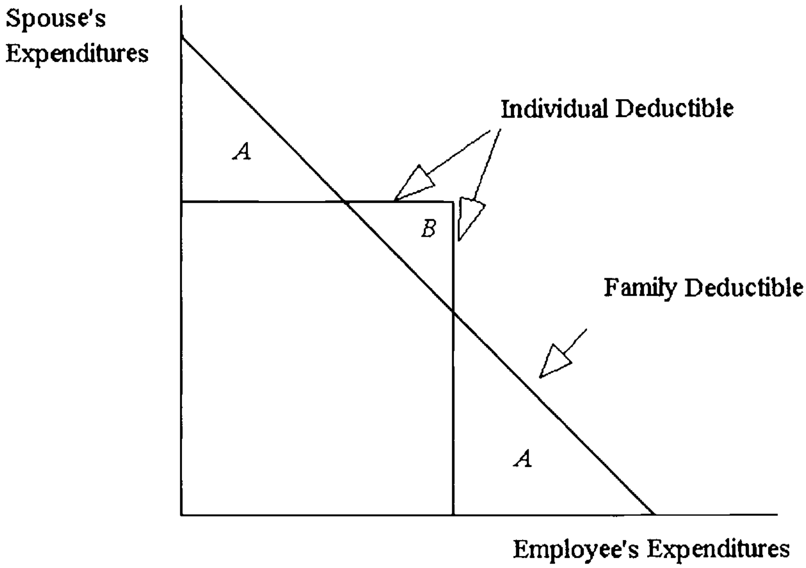

Fig. 6C.1 Families can trade individual risk (*A*) for less combined risk (*B*)

creased likelihood that the couple will face payments of $1,999 for each. Although Eichner's choice of the bivariate probit allows him to ignore the shape of the density surface within each of the quadrants, it also means that he is forced to focus only on between-quadrant shifts in density rather than the within-quadrant shifts.

Second, because the bonds of matrimony allow couples to enforce cost-sharing agreements better than two strangers with less correlated expenditures could do, they can avoid some of the costs of having correlated expenditures by purchasing a policy with a joint deductible. As illustrated in figure 6C.1, by buying a policy with a family deductible rather than an individual deductible, a family can trade off less insurance against individual risk (A) by buying more insurance against pooled risk (B). Here, the couple's deductible would be less than two times the individual deductible. Yet the insurance company's expected cost would be unchanged as long as the expected cost in the area B is just equal to the expected cost in the areas labeled A. Even though any two strangers on the street might also be able to benefit from pooling their resources and sharing the cost of health care, they lack the means that a family would have for enforcing it.

Conclusion

This paper is an early contribution to what promises to be an important line of research on families' claims behavior and insurance purchasing decisions. Expenditures among couples and for a given person over time are likely to be correlated. As Eichner points out, this may well have strong welfare implications for different types of insurance. Moreover, sorting out the source of the correlation is crucial. To the extent that expenditures are correlated *because* of price effects—for instance, having used up my family's deductible, I lower the cost for my wife's next procedure—rather than taste effects, such variation may also be a valuable source of exogeneity for studying the price elasticity of demand for medical care.