View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

NBER WORKING PAPER SERIES

PHYSICIAN INCOME EXPECTATIONS AND SPECIALTY CHOICE

Sean Nicholson
Nicholas S. Souleles

Working Paper 8536
http://www.nber.org/papers/w8536

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2001

The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2001 by Sean Nicholson and Nicholas S. Souleles. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Physician Income Expectations and Specialty Choice
Sean Nicholson and Nicholas S. Souleles
NBER Working Paper No. 8536
October 2001

ABSTRACT

In spite of the important role of income expectations in economics, economists know little about how people actually form these expectations. We use a unique data set that contains the explicit income expectations of medical students over a 25-year time period to examine how students form income expectations. We examine whether students condition their expectations on their own ability, contemporaneous physician income, and the ex post income of physicians in their medical school cohort. We then test whether a model that uses the students' explicit income expectations to predict their specialty choices has a better fit than a model that assumes income expectations are formed statically, and a model that bases income expectations on ex post income.

Sean Nicholson
The Wharton School
University of Pennsylvania
nicholss@wharton.upenn.edu

Nicholas S. Souleles
The Wharton School
University of Pennsylvania
and NBER
souleles@wharton.upenn.edu

I. Introduction

Income expectations play a central role in many economic studies, particularly studies of schooling decisions and occupational choice. Economists rarely observe peoples' income expectations and therefore must make assumptions about how individuals form these expectations. There is little agreement among economists regarding whether income expectations are formed fully rationally or not, and what information people use when forming income expectations (Manski, 1993). In his "cobweb" model of occupational labor supply, Freeman (1971) assumes that income expectations are static, or adaptive. College students expect to earn the contemporaneous mean income of individuals who are already in the contemplated profession. Sloan (1970) also assumes expectations are static in his study of physician specialty choice by deriving medical students' expected income associated with each specialty based on contemporaneous physician income. By contrast, Willis and Rosen (1979) assume that income expectations are rational; high school graduates understand the process that will generate their future income. Students compare the expected earnings associated with attending college versus completing high school, and select the utility-maximizing schooling level. Hay (1991), who also assumes that medical students have rational income expectations, extends the Willis and Rosen (1979) model to allow medical students to choose from among three possible specialties (rather than two schooling levels). Willis and Rosen (1979) and Hay (1991) test their models by instrumenting for individuals' ex-post income with information known at the time of the schooling decision, such as gender and ability.

In all four of the studies cited above, the authors find that students are more likely to choose a particular occupation or a schooling level if they have a relatively large difference in expected income between that occupation/schooling level and its alternative. However, as Manski (1993) points out, without any evidence of how people form income expectations, "the most that one can do is infer the decision rule conditional on maintained assumptions on expectations." Manski (1993) has shown that misspecifying how income expectations are formed can lead to incorrect parameter estimates, such as the responsiveness of enrollment to the rate of return to schooling.

This paper directly examines how medical students form income expectations. Specifically, to what extent does a medical student condition his own expected income on the mean income of physicians who are currently practicing, the ex post income of physicians in the specialty in which he plans to practice, measures of his ability, and other individual characteristics such as age, gender, and race? Are students' expectations forward-looking, in that they correctly anticipate future trends in specialty income?[1] Medical students are some of the brightest people in the country, so if we find that their income expectations are not forward-looking, it would be hard to imagine that other students have forward-looking expectations. We then test whether students' explicit income expectations help predict their specialty choices, decisions for which expectations of future returns should be critical. We separately compare the fit of a specialty choice model that uses the students' explicit income expectations with a model that assumes students form income expectations statically, and a model that uses ex post income.

We use a unique data set that contains direct measures of medical students' subjective income expectations and a rich set of demographic and ability measures. The Jefferson Longitudinal Study contains information on all 7,433 medical students who graduated from Jefferson Medical College, a large medical school in Philadelphia, between 1968 and 1998. The students were surveyed in their fourth year of medical school and asked to predict the following: the specialty in which they will practice, their income five, 10, and 20 years after completing residency training (i.e., their income with five, 10, and 20 years of post-residency experience), peak career income, and characteristics of their medical practice.

There have been a few studies of subjective income expectations available in other surveys. Most

[1] Some papers indirectly test the rationality of occupational choice without observing peoples' subjective income expectations. Zarkin (1985) examines whether prospective teachers incorporate forecastable demand conditions into their decision to enter the occupation. He finds that future student enrollment rationally affects the occupational decisions of secondary school teachers, but not elementary school teachers. Siow (1984) assumes that prospective lawyers form income expectations rationally and expect future cohorts of students to arbitrage away any rents that would otherwise occur from a wage shock. He finds evidence consistent with his model.

of these studies focus on testing the rationality of the expectations (e.g., Das and van Soest, 1997; Dominitz, 1998; Nicholson and Souleles, 2000). There has been little analysis, however, of the determinants of expectations and of whether expectations help predict people's behavior. Souleles (1999) analyzes the time series and cross-sectional variation in expectations in the Michigan Surveys of Consumer Attitudes and Behavior, regarding variables like household income and financial position, inflation and aggregate economic activity. He shows that such questions help predict households' consumption and precautionary saving, as well as their portfolio allocations (Souleles, 2000).

Most surveys of expectations like the Michigan Surveys ask about expectations over a one-year horizon, so they cannot be used to analyze low frequency life-cycle behavior. Also, the answers to the expectational questions are often constrained to be discrete (e.g., Will your income increase, decrease, or stay the same over the next year?). The Jefferson Longitudinal Study is well suited to examine how students form income expectations and whether income expectations help predict specialty choice. The Jefferson survey solicits open-ended, quantitative income expectations. The database also includes information on student performance during medical school, including scores on national board exams. Furthermore, the sample period spans a 30-year time period during which there were tremendous changes in the health care market. This provides a large amount of variation in factors that should affect income expectations. With long-horizon income expectations we can analyze a person's occupational choice, a decision with substantial welfare implications.

We find that medical students do condition their expected income on their own ability and the contemporaneous income of physicians in the students' intended specialty. However, expectations are also forward-looking; students incorporate future trends in specialty income into their own expectations. We also find that the students' explicit income expectations are useful in predicting their behavior. The specialty choice model that uses students' explicit income expectations fits the data better than models that assume students form income expectations statically or that expectations match ex post income.

The paper is organized as follows. The two data sets used in this paper – the Jefferson

Longitudinal Study data set and the American Medical Association annual survey on physician income – are described in the following section. Section III presents the empirical model we use to analyze how medical students form income expectations and how those expectations affect their choice of specialty. Results are reported in Section IV and concluding comments in Section V.

II. Data

In 1968 Jefferson Medical College began surveying its medical students in their fourth year of school. Students were asked to predict various aspects of their future medical practice and career: the specialty in which they will practice, the number of hours per week they will work, the proportion of time they will devote to research and teaching, the percentage of their patients who will be low income, and the age at which they will retire. Between 1968 and 1979, students were also asked to state the income, after medical expenses and before taxes, that they expected to receive 5, 10, and 20 years after completing residency training, and the peak income they expected to receive during their career. Students who graduated after 1979 have been asked to predict their peak income only. Students were asked to report their income in real terms rather than trying to guess the inflation rate.[2] The Jefferson Longitudinal Study now contains information on 7,433 individuals, most of whom are now practicing physicians. Sample means are presented in Table 1.

Information on student performance has been added to the Jefferson Longitudinal Study. Medical students must pass three national exams before they can receive a license to practice medicine in the United States. Part 1 of the National Board of Medical Examiners (NBME) test is administered after the second year of medical school and covers the classroom material taught during the first two years (e.g., anatomy, physiology, pharmacology). Jefferson students received an average score of 204.4 on Part

[2] For example, the 1970 survey question was worded as follows: "In answering the following questions relating to income, please assume that dollars maintain their 1970 value. What do you think your own gross personal income (after professional expenses, but before income taxes) will be 5 years after

1 of the NBME, slightly below the national average of 210.[3]

Many economists who study occupational choice and schooling decisions assume that individuals base their income expectations on the contemporaneous mean income of earlier cohorts of students who are now employed in the occupation of interest. We use data from American Medical Association (AMA) surveys to characterize the information that was available to the Jefferson students.[4] The AMA reports the mean and median medical practice income (income after professional expenses but before taxes) by specialty for physicians in the following age groups: under 36 years old, between 36 and 45, between 46 and 55, between 56 and 65, and over 65 years old. We assume that the mean income for an age group corresponds to physicians at the midpoint of the age range (e.g., the mean income of a 40-year-old physician is assumed to be the mean income reported for the 36-45 year old group). We linearly interpolate between age-specific observations to estimate the national, cross-section age-income profile for each specialty in each year.[5] Most students complete medical school between the ages of 26 and 28, so the age-income profile closely approximates the experience-income profile. The income expectations of the Jefferson students are then matched with the corresponding national mean income of physicians who are currently practicing in the specialty in which each medical student intends to enter. For example, consider a student who graduates from medical school in 1977, plans to become a surgeon, and expects his income with five years of experience to be $185,000.[6] In 1976 the mean income of surgeons with five years of experience from the AMA survey was $172,000, measured in 1996 dollars. When analyzing the

---

completing residency training?"

[3] The second part of the NBME exam is administered in the fourth year of medical school and the third part is administered in the first year of the students' residency program. We focus on the Part 1 score as a measure of student ability and performance because for some students the second and third parts occurred after they stated their income expectations.

[4] Each year the AMA surveys a nationally representative, random sample of about 4,000 practicing physicians.

[5] Specifically, we assume that physicians complete medical school at age 26, spend the number of years as a resident as required by their specialty, and begin practicing medicine immediately after completing residency training.

determinants of income expectations, we compare this person's expected income with the mean contemporaneous income of surgeons in 1976 who had five years of experience (Y^N_5, where the superscript refers to "national" and the subscript refers to years of post-residency experience). We lag the national data by one year because the AMA survey is published one year after the survey is conducted. The same procedure is used to assign a corresponding national mean income for students' expected income with 10 and 20 years of experience. The peak contemporaneous income in a specialty is the maximum age-specific mean in the national cross-section experience-income profile.

By 1987 the student described above should have completed his five-year surgery residency program and have been practicing medicine for five years. In 1986 the mean income of surgeons with five years of experience from the AMA survey was \$205,000, also measured in 1996 dollars. The mean income of surgeons with five years of experience increased by \$33,000 (\$205,000 - \$172,000) from the time the student formed his expectation to the time the student had five years of experience. In some of the analysis that follows, we examine whether students correctly anticipated the future change in income and incorporated this change into their own expectations.

Betts (1996) and Nicholson (1999) have shown that there is substantial variation among college and medical students in how much they know about wages. Similarly, Souleles (1999) even found variation in peoples' forecasts of aggregate variables like inflation and economic activity. We refer to another series of questions on the Jefferson survey to incorporate heterogeneity of market information. Jefferson medical students were asked to estimate the practice income *currently* being earned by physicians in six different specialties: family practice, internal medicine, surgery, pediatrics, obstetrics/gynecology, and psychiatry. We calculate the accuracy of each student's market information by taking the difference between the student's estimate of the current income in the specialty in which they plan to enter and the median income of physicians already practicing in that specialty, as reported in

---

[6] All expected and realized income in this paper are converted to 1996 dollars using the urban CPI.

the AMA surveys.[7] Jefferson students under predict the prevailing median income in their preferred specialty by $17,400, on average, as reported in the last row of Table 1.

III. Empirical Method

a. Income Expectations

We begin by examining how students form income expectations. Let $EY_{i,j,t=0}$ represent the income that student i in his fourth-year of medical school (t=0) expects to receive when he has j years of post-residency experience. We allow a student's information set to consist of personal characteristics (X_i); the contemporaneous mean national income of physicians with j years of experience who are practicing in year 0 in the specialty the student intends to enter ($Y^N_{j,t=0}$, where "N" refers to "national"); an indicator variable for the specialty the student expects to enter (S); an indicator variable for the student's graduation year (T); a measure of how accurately the student estimates the prevailing income of physicians in the specialty he plans to enter ($Y^{N,est} - Y^N$); and the future growth of physician income in the student's specialty ($Y^N_{j,t=j} - Y^N_{j,t=0}$):

$$(1)\quad EY_{i,j,t=0} = \alpha_1 X_i + \alpha_2 Y^N_{j,t=0} + \alpha_3 S + \alpha_4 T + \alpha_5 (Y^{N,est}_{i,t=0} - Y^N_{t=0}) + \alpha_6 (Y^N_{j,t=j} - Y^N_{j,t=0}) + u_1$$

Through a series of ordinary least squares regressions of form (1), we test whether and to what extent students' condition their expected income on the above information.

Our measure of contemporaneous physician income ($Y^N_{j,t=0}$) is conditioned only on specialty and experience level. If medical students only observe the specialty and experience level of practicing physicians and expectations are static, α_2 will equal one and the other coefficients in equation (1) will be

[7] About three-quarters of the Jefferson students expected to enter one of these six specialties.

zero. Non-zero α_1 coefficients are consistent with either rational expectations or static expectations where students observe characteristics other than just specialty and experience level (e.g., ability, race, gender). When the specialty and year indicators are excluded from the regression, α_2 is identified by variation in physician income between specialties at a point in time, and variation over time in all specialty incomes. When the specialty and year indicators are included, α_2 is identified by within-specialty income variation over time. Therefore, a positive coefficient on α_2 in the latter specification indicates that students are quite knowledgeable about *relative* changes in specialty incomes.

The coefficient α_5 measures the extent to which students' misinformation about contemporaneous physicians' income is incorporated into their own income expectations. The final variable in equation (1) measures the future growth rate of physician income in the student's specialty. If students had perfect foresight, α_6 would be close to one.[8] More generally, α_6 measures the proportion of future income growth that is forecasted by students. For example, if students anticipate that the demand for surgical services will increase in the future, they might expect their future income to be higher than that of previous cohorts of surgeons. A positive coefficient on α_6 would be the strongest indication that students are forward-looking when forming their income expectations.

b. Specialty Choice

Some economists question the value of survey questions such as those asking students to state their expected income. It is worth noting, however, that most variables in household datasets are based on self-reported information. Instead of repeating the well-known advantages and disadvantages of survey questions, we formally test whether subjective income expectations help predict people's behavior.

[8] The physician's own income realization $Y_{i,j}$ is not in his information set at time t=0. By contrast, $Y^N_{j,t=j}$ is the average income in his specialty in year j. Because $Y^N_{j,t=j}$ can be interpreted as the projection of $Y_{i,j}$ on time and specialty dummies, it can be interpreted as a rational expectations forecast of $Y_{i,j}$, and so is a valid regressor in equation (1).

Specifically, we analyze the specialty choice decisions of medical students using three different assumptions about how students form income expectations. We compare the fit of a model that uses students' subjective income expectations to the fit of models where medical students are assumed to have static income expectations – they expect to receive the contemporaneous mean income of practicing physicians -- or forward-looking expectations – medical students expect their income with 10 years of experience, for example, to equal the mean income that was actually received by their cohort 10 years later. If students' specialty choices can be predicted more accurately in the former model relative to the latter two models, this provides evidence that subjective income expectations can help explain individuals' behavior.

Ninety-seven percent of U.S. medical school graduates enter a residency training program after completing school. Residency positions are available in 26 different specialties, which range in length from three years for primary care specialties (family practice, internal medicine, and pediatrics) to five years for surgical specialties (e.g., orthopedic surgery and general surgery). In 1997 the mean income of primary care physicians was $155,000, considerably lower than non-primary care physicians ($230,000). However, due to the apparently favorable non-monetary attributes of the primary care specialties and barriers to entry into some non-primary care specialties, each year a majority of the graduating medical students enter primary care residency programs (Nicholson, 1999).[9]

Our model of specialty choice and income expectations is based on those of Hay (1991) and Willis and Rosen (1979). Medical students consider the monetary and non-monetary attributes of each specialty and choose the specialty that maximizes their expected lifetime utility. For simplicity, we consider the choice between a primary care (S=0) and a non-primary care specialty (S=1).[10] The

[9] Many students who begin primary care residency programs sub-specialize after completing their initial program (e.g., internal medicine residents receiving further training in cardiology), so that the majority of physicians in the U.S. practice in a non-primary care specialty.

[10] We group psychiatry with the three traditional primary care specialties for this analysis because the

difference in the expected utility of entering a non-primary care versus a primary care specialty, I, is defined as a function of individual characteristics, Z:

$$(2) \quad I_i = \gamma Z_i + u_i$$
$$\Pr(\text{choose } S = 1) = \Pr(I \geq 0) = \Phi(\gamma Z),$$
$$\Pr(\text{choose } S = 0) = \Pr(I < 0) = 1 - \Phi(\gamma Z),$$

where u is has a standard normal distribution. One component of the net benefit of choosing a non-primary care specialty is the income a student expects to receive in a non-primary care specialty relative to the income they expect to receive in a primary care specialty. Students who actually enter a non-primary care specialty expect to receive income Y_1, and students who actually enter a primary care specialty expect to receive income Y_0:

$$(3) \quad Y_{i1} = \beta_1 X_i + \varepsilon_{i1}$$
$$Y_{i0} = \beta_0 X_i + \varepsilon_{i0},$$

where **X** is a subset of **Z**. We assume that a student's debt, race, and age affect their choice of specialty but have no effect on the income they expect to receive in each specialty. If students expect to retire at the same age, relatively young medical students will tend to prefer non-primary care specialties because they have more working years over which to recoup their investment in the relatively long non-primary care residency programs. With liquidity constraints, students with substantial debt might prefer a primary care specialty because this allows them to begin paying off their loans more quickly. Finally, it is conceivable that the distribution of equalizing differences between primary and non-primary care differs between whites and non-whites. Bhattacharya (2000) also assumes that debt influences specialty choice

mean income of psychiatrists and the length of psychiatric residency training programs are more similar to the primary than non-primary care specialties.

but not income.

We assume that u and ε_1 have a bivariate normal distribution with means of zero, standard deviations of one and σ_1, respectively, and correlation ρ_1; u and ε_0 are assumed to have a bivariate normal distribution with means of zero, standard deviations of one and σ_0, respectively, and correlation ρ_0. In the empirical application of this model, Y will either refer to a medical student's expected peak income over the course of their career, or their income with 10 years of experience.

The Jefferson survey records a student's expected income in their chosen specialty only. That is, we observe Y_1 when I > 0 and Y_0 when I < 0. The expected income of students who enter a non-primary care specialty is truncated on a positive expected net benefit of choosing such a specialty, and the expected income of students who enter a primary care specialty is truncated on a negative expected net

$$
\begin{aligned}
(4)\quad E[Y_1 | Y_1 \text{ observed}] &= E[Y_1 | I \geq 0] \\
&= \beta_{1'} X + E[\varepsilon_1 | I \geq 0] \\
&= \beta_{1'} X + \rho_1 \sigma_1 \lambda_1 \\
&= \beta_{1'} X + \beta_{\lambda 1} \lambda_1, \\
\text{where } \lambda_1 &= \left(\frac{\phi(\gamma Z)}{\Phi(\gamma Z)}\right)
\end{aligned}
$$

$$
\begin{aligned}
(5)\quad E[Y_0 | Y_0 \text{ observed}] &= E[Y_0 | I < 0] \\
&= \beta_{0'} X + E[\varepsilon_0 | I < 0] \\
&= \beta_{0'} X + \rho_0 \sigma_0 \lambda_0 \\
&= \beta_{0'} X + \beta_{\lambda 0} \lambda_0, \\
\text{where } \lambda_0 &= -\left(\frac{\phi(\gamma Z)}{1 - \Phi(\gamma Z)}\right)
\end{aligned}
$$

benefit of choosing a non-primary care specialty:

If selection into specialties is non-random, an ordinary least squares regression of students' expected income on student characteristics would yield inconsistent estimates of β_1 and β_0. Consistent estimates can be obtained if one includes the Mills ratios, or selection-correction terms, λ_1 and λ_0.

λ_1 is non-negative, so a positive coefficient $\beta_{\lambda 1}$ would indicate that the observed expected income of students who enter a non-primary care specialty is greater than the income that students who chose a primary care specialty would expect to earn if they instead chose a non-primary care specialty. Since λ_0 is non-positive, a negative coefficient $\beta_{\lambda 0}$ would indicate that the expected income of students who chose a primary care specialty is also biased upward relative to the population expected income, where the population includes students who enter both specialties. If, on the other hand, $\beta_{\lambda 1}$ and $\beta_{\lambda 0}$ were both positive, students who enter non-primary care specialties would have an absolute advantage, in terms of expected income, relative to students who enter primary care specialties.

Following Hay (1991) and Willis and Rosen (1979), our estimation strategy consists of three steps. First, we estimate equation (2) with a probit model to obtain the reduced-form estimates of γ. We then compute λ_1 for students who actually enter a non-primary care specialty and λ_0 for students who actually enter a primary care specialty. In the second step we estimate β_1 and $\beta_{\lambda 1}$ by regressing the expected income of students who chose a non-primary care specialty on observed characteristics and λ_1. Likewise, we estimate β_0 and $\beta_{\lambda 0}$ by regressing the expected income of students who chose a primary care specialty on observed characteristics and λ_0.

In order to estimate the responsiveness of specialty choice to expected income, we need to estimate the income a student would expect to receive in the specialty they decided *not* to enter. According to our model, a student who chose a primary care specialty would have the following expected

$$(6)\ \ E\,[Y_1 \,|\, Y_0\ observed\,] = Y_1^* = E\,[Y_1 \,|\, I < 0] \\ = \beta_1 X + E\,[\varepsilon_1 \,|\, I < 0] \\ = \beta_1 X + \beta_{\lambda 1}\lambda_0$$

income in a non-primary care specialty:

A similar approach is used to predict the counterfactual expected income (Y_0^*) for students who actually entered a non-primary care specialty. For each student we calculate the difference in expected income between a non-primary and primary care specialty as Y_1 - Y_0^* for students who chose a non-primary care specialty, and Y_1^* - Y_0 for students who chose a primary care specialty.

In the third step of the empirical model we re-estimate the probit equation (2) after including for each student the difference in expected income between non-primary and primary care specialties. Variables that are assumed to influence expected income but not directly affect the non-monetary value of a specialty, such as a student's board score, are excluded from the final specialty choice regression. We estimate the standard errors of this final probit by jointly bootstrapping all three steps.

We assess the usefulness of subjective income expectations in three ways. First, we compare the log likelihood of the model where students are assumed to make specialty choices based on their unique subjective expected income to the log likelihood of the models where students' income expectations are assumed to be static or to match ex post income. Second, in each model we predict the specialty that each student will choose, and then compare the percentage of choices correctly predicted by each of the three models. The coefficients from the probit specialty choice regression yield a predicted latent utility for each student, which can be translated into a predicted probability of choosing non-primary care. If the predicted probability for a particular student is greater than the proportion of the sample that actually chose a non-primary care specialty, we predict that this particular student will choose a non-primary care specialty. Third, we estimate a specialty choice probit that includes both the subjective income expectation and a new variable defined as the difference between the student's static income expectation and their subjective income expectation. If the coefficient on latter coefficient is insignificant, then information on contemporaneous physicians' income does not provide *incremental* predictive power for specialty choice, beyond the information available in the expectations variables. We perform the same procedure for ex post income as well.

IV. Results

a. Income Expectations

We begin with an analysis of how medical students form expectations of their peak income. Every fourth-year medical student at Jefferson Medical College since 1968 has been asked to predict the peak income they will receive during their career, so our sample is large and covers a time period when the health care market has undergone profound change. The mean expected peak income for each cohort of Jefferson students between 1974 and 1998, measured in 1996 dollars, is depicted in Figure 1 for students entering family practice, and in Figure 2 for students entering surgery. These are two of the most popular specialties among the Jefferson students and therefore provide large sample sizes. The corresponding contemporaneous national peak incomes of practicing physicians, from the AMA cross-section surveys, are also depicted in these two figures (the AMA data are also reported in 1996 dollars).[11] For instance, for fourth-year medical students graduating in 1987 and entering family practice, the contemporaneous national peak income is the peak income of family practice physicians from the 1986 AMA survey.

Between 1974 and 1987 the expected peak income of students entering family practice corresponded very closely with the peak incomes of practicing physicians (Figure 1). At the start of this time period the income of family practitioners fell substantially and the Jefferson students adjusted their own expectations accordingly. The two lines in Figure 1 diverge in 1987; family practice income has increased in real terms while the students' expectations have remained fairly constant. Nonetheless, the two time series remain correlated. Medical students entering surgery after 1983 expected their peak income to be less than that of practicing surgeons (Figure 2). As surgeons' incomes increased in the late 1980s and early 1990s, the students' expectations likewise increased. In 1993 the students' mean

[11] The AMA did not conduct a physician survey in 1977, 1980, and 1981 so contemporaneous peak income by specialty is not available for those years.

expected peak income decreased by 25 percent from the previous year, probably in response to the Clinton administration's intent to reform the health care system. This drop occurred even though the income of practicing surgeons nationally changed very little in 1993. It appears that when the Clinton reform plan failed the medical students entering surgery in 1994 increased their expectations sharply.

One possible explanation for the divergence of expected from contemporaneous income in the mid 1980s is that Jefferson Medical College, like most medical schools, began accepting more female students in the 1980s. The population of practicing physicians, however, was still predominantly male until the 1990s. Since female medical students have lower income expectations than their male colleagues, as we demonstrate below, an increase in the proportion of medical students who are female will reduce the mean expected income, all else equal. An alternative explanation is that over time the ability of the Jefferson students might have fallen relative to the national average.[12] In both specialties the expected peak income decreased substantially between 1978 and 1981. This temporary decline might have occurred because it took students several years to incorporate the unusually high inflation rates into their own forecasts of real income.[13]

We perform a series of regressions of equation (1) in order to examine the determinants of income expectations. In the first column of Table 2, a student's expected peak income is regressed on personal characteristics only in order to examine the role of ability. Ability can affect expected income in two ways. A person of relatively high ability might be more likely to choose and be admitted into a high-paying specialty; and/or there might be returns to ability within a specialty. When indicator variables for a student's intended specialty are omitted, the coefficient on ability will capture both of these effects. Students who received a score on Part 1 of the NBME in the top quartile among Jefferson students are

[12] Nationally, the failure rate on Part 1 of the NBME has fallen from about 16 percent to 5 percent between 1990 and 1997, while the percentage of Jefferson students failing the exam has remained fairly constant at about three to five percent during this period.
[13] Souleles (1999) also finds in the Michigan surveys that inflation expectations were lower than actual inflation at this time.

assigned a value of one for the high board score variable; students who scored in the bottom quartile are assigned a value of one for the low board score variable. The coefficient on the high board score variable is significantly positive but small in magnitude; ability has a relatively minor effect on income expectations. Students who perform well on the board exam after their second year of medical school expect their peak career income to be about $10,000 higher than students with average performance. This represents a 5.7 percent premium for people with relatively high scores. Women expect their peak income to be $50,000 less than men. However, female physicians generally work fewer hours than male physicians. One advantage of the Jefferson data set is that students were asked to report the number of hours they expect to work. Controlling for expected hours reduces the coefficient on the female indicator by 50 percent, but it remains negative and statistically significant (Table 4 below).

We now address the question of whether students adjust their expected income by the contemporaneous difference in mean incomes between specialties. In the second regression of Table 2 we include the contemporaneous national peak income of practicing physicians ($Y^N_{peak,t=0}$) in the specialty the student intends to enter in the year in which the student states his expectation. The estimated coefficient on this variable is 0.67; a one-dollar increase in the peak income of physicians in a particular specialty is associated with a 0.67 dollar increase in the expected peak income of a medical student who plans to enter that specialty. The significance of this coefficient and the increase in R^2 from 0.07 to 0.25 suggests that students do condition their expectations on the contemporaneous income in the specialty they plan to enter. This result is not inconsistent with rational expectations, however. If income is serially correlated across cohorts, rational expectations should be partly correlated with contemporaneous income. Since the coefficient is less than one, expectations are not strictly adaptive; they do not depend only on contemporaneous income. The female coefficient decreases in absolute value because women are more likely to enter a low-paying specialty than men.

Indicator variables for the specialty a medical student intends to enter and their graduation year are added to the third regression of Table 2. We begin by assuming specialty choice is exogenous, or at least predetermined. The next section models the specialty choice directly. The specialty coefficients measure the average difference in expected income, relative to family practice, over the entire 1974-1997 time period. The coefficient on the national peak income ($Y^N_{peak,t=0}$) of 0.22 is still statistically different from zero although smaller than in the previous specification. This coefficient is now identified by income variation within a specialty over time. The peak incomes of physicians in the six different specialties between 1973 and 1996 are plotted in Figure 3. Specialty incomes did not always move together during this time period. For example, the incomes of surgeons and obstetricians increased in the late 1980s relative to the other four specialties, and decreased relative to the other specialties in 1993 and 1994. These results suggest that students do incorporate relative changes in specialty income into their own expectations, but not on a dollar for dollar basis.

Four of the five specialty coefficients are positive and statistically significant in the third regression, and the coefficients for surgery and obstetrics are large. Thus, students condition their expectations on more than just the current income in their intended specialty. Consider the coefficient of $73,000 on the surgery indicator variable. Conditioning on the contemporaneous income of physicians, students who intended to become surgeons expect their peak income to be $73,000 higher than students who intend to become family practitioners (the omitted specialty indicator), on average. This implies that prospective surgeons were forward-looking; they expected income to grow in the future relative to family practice. In fact, over the 1973 to 1996 period, the real peak income of surgeons and obstetricians increased by 32 percent and 18 percent, respectively, while the peak income of family practitioners decreased by 18 percent (see Figure 3).

In the fourth regression of Table 2 we add a variable that measures a student's knowledge of current physician income. The "information accuracy" variable is the difference between a student's assessment of the current income of physicians in the specialty he plans to enter and the contemporaneous national median income of physicians in that specialty, as measured by the AMA surveys. The estimated coefficient on this variable is 0.84; a student's market misinformation is incorporated almost dollar for dollar into their own income expectations. If, for example, a student who plans to become a pediatrician believes that pediatricians currently make $10,000 more than they actually do, his expected peak income will be $8,400 higher than a similar person who perfectly predicts a pediatrician's income. Adding the information accuracy variable increases the R^2 from 0.30 to 0.45. Hence, a considerable amount of the variation in income expectations is due to heterogeneity of information regarding the physician market. The coefficient on the contemporaneous national peak income is substantially higher in the fourth specification relative to the third specification, and the coefficients on the surgery and obstetrics indicator variables are substantially lower. This latter result suggests that part of the reason students entering surgery and obstetrics expected their income to grow in the future relative to other specialties was that they overestimated the current income of physicians in those specialties.

The disadvantage of using the expected peak income variable is that we do not know when in their career a student expects his peak income to occur. By contrast, expected income information for 5, 10, and 20 years of experience are available for about two-thirds of the 2,300 students who graduated from Jefferson Medical College before 1980.[14] This more detailed information allows us to analyze the lifetime profile of expected income, and to test whether students were able to anticipate the future growth rate of specialty income. In Table 3 we pool each student's expected income observations for 5 and 10 years of experience and include a variable that measures the future growth rate of specialty income (the variable ($Y^N_{j,t=j}$ - $Y^N_{j,t=0}$) in equation (1)).[15] Standard errors have been corrected to allow for correlation in the error terms between the multiple observations for an individual, and an indicator variable is included for observations with 10 years of experience (five years is omitted).

The coefficient of 0.27 on the income growth variable in Table 3 indicates that students do incorporate future changes in specialty income into their own expectations. This is strong evidence that

---

[14] Beginning in 1980, students were asked to state their expected peak income only.

[15] Expected income with 20 years of experience is omitted because the AMA data for physicians with 20

income expectations are forward-looking, and not entirely adaptive. The second regression in Table 3 omits the income growth variable for purposes of comparison with the regressions in Table 2. Other than the contemporaneous mean income, the coefficient estimates do not change substantially when the income growth variable is included.

In Table 4 we pool the four expected income observations for each student (income with 5, 10, and 20 years of experience, and peak income). We also include expected characteristics of the students' medical practices. Students who expect to spend more time on teaching and research expect to make less. Controlling for the number of hours per week a student expects to work, women still expect an annual income $25,000 lower than men. Students expect their income to increase by $35,000 between the 5th and 10th years of experience, but only by $24,000 between the 10th and 20th years of experience. The peak income is expected to be $12,000 higher than their income with 20 years of experience, on average. The students' responses imply a compound, annual real growth rate in income of 6.1 percent between years 5 and 10, but only 1.6 percent between years 10 and 20.

b. Specialty Choice

In order to examine whether subjective income expectations data help predict behavior, we analyze the decision by medical students to enter a primary or non-primary care specialty after graduating from medical school. We first analyze the specialty choice decision under the assumption that income expectations are static or adaptive; i.e., that students expect their peak lifetime income in primary and non-primary care to be equal to the contemporaneous peak incomes of practicing physicians in these two specialty groups. The peak contemporaneous incomes are taken from the American Medical Association annual surveys.[16] Since we do not have the individual-level AMA data, in this first model we do not

---

years of experience is not yet available for most of the Jefferson graduates.

[16] We weight the peak incomes in surgery and obstetrics to derive a contemporaneous non-primary care

control for non-random selection into the various specialties.

The first column of Table 5 presents coefficient estimates from a probit model of the specialty choices of 2,458 medical students who graduated between 1971 and 1997 under the assumption that students have static income expectations. The dependent variable is one if a student chooses a non-primary care specialty, and a zero otherwise. Of greatest interest is the variable measuring the difference between the contemporaneous peak income of non-primary care and primary care physicians. Its estimated coefficient is 0.00389, and significant. A $10,000 increase in the contemporaneous income of non-primary relative to primary care physicians is associated with an increase of 0.014 (from 0.36 to 0.374) in the probability that a medical student will choose a non-primary care specialty. Men and non-white students are more likely to choose a non-primary care specialty relative to their peers. We correctly predict the specialty choices of 57.3 percent of the students in this first model.

We compare the fit of this model where students are assumed to have static income expectations to a model that uses the students' explicit income expectations. The latter model is estimated in three steps to control for the possibility of non-random selection into the specialties. We first estimate a reduced-form probit model as specified by equation (2). Coefficient estimates from this model are reported in the second column of Table 5. The coefficients on gender, age, and race have the same sign and a similar magnitude as in the specification where students are assumed to have static income expectations. Board scores and the students' knowledge of contemporaneous physician income (income information accuracy) have been shown in the previous section to affect income expectations.[17] We

---

peak income. The weights are based on the number of practicing physicians in each of these two non-primary care specialties. Likewise, we weight the peak incomes in family practice, internal medicine, pediatrics, and psychiatry to derive a contemporaneous primary care peak income.

[17] The difference in a student's income information accuracy is the difference between their perception of the mean contemporaneous non-primary care income and the actual contemporaneous mean non-primary care income, minus the difference between their perception of the mean contemporaneous primary care income and the actual contemporaneous mean primary care income.

assume that these two variables affect specialty choice through income expectations only, and they are therefore included in the reduced-form probit but not the probit regressions that include income expectations. The percentage of choices correctly predicted (58.3) is slightly higher than in the model with static income expectations. The coefficients from the reduced-form probit regression are used to derive a Mills ratio for each student, as specified in equations (4) and (5).

For students entering non-primary care and primary care specialties, we separately regress the students' subjective expected income on personal characteristics and the Mills ratios. The coefficient on the Mills ratio (8) is positive and significant in the non-primary care regression, and negative and insignificant in the primary care regressions (Table 6). This indicates that the students who actually chose non-primary care expect to earn more in that specialty than would students who actually chose primary care. Women expect to earn substantially less than men in both specialties, but particularly in the non-primary care specialties.

The coefficients from the first two columns of Table 6 are used to estimate each student's counterfactual expected income, as outlined in equation (6). We then re-estimate the probit specialty choice equation after including the difference in the expected peak income between non-primary and primary care for each student. Coefficient estimates for the selection-corrected probit model are reported in the fifth column of Table 5, and the standard errors in the sixth column are estimated by jointly bootstrapping the three equations. The coefficient on the difference in students' income expectations is positive, significant, and four-times larger than the income coefficient from the model with static income expectations. A $10,000 increase in a student's expected income in non-primary relative to primary care is associated with an increase of 0.057 in the probability of entering non-primary care (from 0.360 to 0.417).

One way to measure the usefulness of information on students' subjective income expectations is to compare the log likelihood of the specification with the subjective expectations to the model with static income expectations. The log likelihood of the model with subjective income expectations (-819) is considerably larger than the log likelihood of the model where students expect to earn the contemporaneous peak income (-1,561). Furthermore, we correctly predict the specialty choice for 85.6 percent of the sample in the model with subjective income expectations, versus 57.3 percent in the model with static expectations. Students' explicit income expectations appear to be much more useful for predicting specialty choice than static expectations.

The final column of Table 5 includes both the students' subjective income expectations and a variable defined as the difference between the static and subjective income expectations. The coefficient on the former variable is positive and significant as before. The coefficient on the former variable, which represents information in contemporaneous physician incomes that is not present in the students' subjective income expectations, is insignificant. Including information on contemporaneous income hardly changes the fit of the model as measured by the log likelihood or the percent of choices correctly predicted. Results from the previous section showed that subjective income expectations are based, in part, on the contemporaneous income of practicing physicians. These new results demonstrate that conditional on subjective income expectations, contemporaneous incomes are no longer informative about specialty choices.

We have also estimated a version of the specialty choice model (results not shown) that contains indicator variables for the year a student completed medical school in order to allow for changes over time in the non-monetary attributes of the specialties. The coefficient of the difference in subjective expected income, which is now identified by variations between students within a cohort, is slightly larger in magnitude (0.0189 versus 0.0165) than the model without the year indicators, and the other results are essentially unchanged.

As a further test of the usefulness of subjective income expectations data for predicting revealed preferences, we compare the results of the model that uses subjective income expectations with a model that assumes medical students' income expectations fully match ex post income. Specifically, we assume that when fourth-year medical students were choosing a specialty, they expected their income with 10 years of experience to equal the mean income that was actually received by their cohort 10 years later. We use 10 years of experience because it is sufficiently far into a physician's career that it should correlate closely with a person's peak income (which we use in the subjective income expectation model), and yet it still allows us to analyze a reasonably large sample of students. We omit from the analysis students who graduated after 1986 because we do not observe the actual income with 10 years of experience for their cohort.

Consider a student who is completing medical school in 1980 and forming expectations regarding their income in family practice. After completing a three-year family practice residency program, her 10^{th} year of experience would occur in 1993. In the model where expectations are based on ex post income, we assume that this student's expected income in family practice is equal to the mean income of family practitioners with 10 years of experience in 1993. Data on physician incomes are taken from the American Medical Association (AMA) annual surveys. As before, we weight expected incomes in the various specialties to derive an overall expected income for primary and non-primary care. Since we do not have the individual-level AMA data, in the ex post income expectations model we do not control for non-random selection into the various specialties. The sample for this specification of the model consists of 1,149 students who graduated from Jefferson Medical College between 1971 and 1986.

In the first column of Table 7 we present the coefficient estimates from the specialty choice probit when expectations are based on ex post income. The variable of greatest interest is the difference between the actual mean income of non-primary and primary care physicians with 10 years of experience for physicians in a student's cohort. This coefficient is negative and insignificant. The model correctly predicts the specialty choice for 50.6 percent of the students.

For purposes of comparison, we re-estimate the subjective income expectation model with this smaller sample of students. The results are reported in the remaining columns of Table 7. In the fifth column, the coefficient on the difference in students' subjective income expectations (0.0112) is slightly smaller than in the previous specification (0.0165 in Table 5) and is not significantly different from zero. Our estimation method is identical to that of Table 5, so the lack of significance on the subjective income expectations variable is due to the smaller sample size. The model correctly predicts the specialty choices for 72.8 percent of the students.

In seventh column of Table 7 we estimate a probit model that includes students' subjective income expectations as well as the difference between the ex post income of the students' cohorts and the students' explicit income expectations. The coefficients on both expectation variables are positive, although neither is significant. Therefore, we reject the hypothesis that ex post income provides additional predictive power for specialty choice. Including ex post income increases the log likelihood slightly and improves the percentage of correct specialty choice predictions by about one percentage point. When entered individually, however, the ex post income variable has a negative coefficient whereas the coefficient on the subjective income expectation variable is positive.

V. Conclusion

This paper analyzed the determinants of medical school student's income expectations. While medical students condition their income expectations on the contemporaneous income of physicians in the specialty they plan to enter, their expectations are not strictly static. Students who enter specialties that subsequently experience income growth relative to other specialties report income expectations that are relatively high. Female medical students expect to earn substantially less than male medical students, even after controlling for the number of hours per week they expect to work. Students who perform relatively well on a national board exam taken during medical school expect income that is slightly higher than their colleagues. Students' misinformation about physicians' contemporaneous income affects their own expectations almost dollar for dollar.

We also find that subjective income expectations help explain medical students' specialty choice decisions. A $10,000 increase in a student's expected income in non-primary relative to primary care specialties is associated with an increase of 0.057 in the probability of entering a non-primary care specialty. Subjective income expectations are more useful in predicting specialty choice than the adaptive expectations often used in the literature, and subjective income expectations summarize most of the explanatory power of the ex-post income of a student's cohort. More generally, these results suggest that subjective expectational questions can help predict people's behavior, including their investment in human capital.

Acknowledgments

Funding for this study was provided by the University of Pennsylvania Research Foundation, Jefferson Medical College, and the Rodney L. White Center for Financial Research. Dean Joseph S. Gonnella, Jon Veloski, Mary Robeson and other members of the Center for Research in Medical Education and Health Care at Jefferson Medical College provided access to and assistance with the Jefferson Longitudinal Study. Patricia Danzon, Mark Pauly, Rob Lemke, Dan Polsky, and members of the Applied Economics workshop at The Wharton School have provided helpful suggestions. Valuable contributions were provided by our research assistants: Josh Beer, Matt Cinque, Hanli Mangun, and Peter Margetis.

References

Betts, Julian R., 1996, "What Do Students Know About Wages?," *Journal of Human Resources* 31(1): 27-56.

Bhattacharya, Jayanta, 2000, "Lifetime Returns to Specialization in Medicine," mimeo.

Das, Marcel and Arthur van Soest, 1997, "Expected and Realized Income Changes: Evidence From the Dutch Socio-economic Panel," *Journal of Public Economics* 32: 137-154.

Dominitz, Jeff, 1998, "Earnings Expectations, Revisions, and Realizations," *The Review of Economics and Statistics*: 374-388.

Freeman, Richard, 1971. The Market for College-Trained Manpower. Cambridge, MA: Harvard University Press.

Manski, Charles F., 1993, "Adolescent Econometricians: How do Youth Infer the Returns to Schooling?," in Studies of Supply and Demand in Higher Education, eds. C. Clotfelter and M. Rothschild, Chicago: University of Chicago Press: 43-57.

Nicholson, Sean and Nicholas S. Souleles, 2000, "Physician Income Expectations, Income Realizations, and Practice Behavior," mimeo.

Nicholson, Sean, 1999, "Physician Specialty Choice Under Uncertainty," mimeo.

Nicholson, Sean, 1999, "Medical Students Perceptions of Physician Income," mimeo.

Siow, Aloysius, 1984, "Occupational Choice Under Uncertainty," *Econometrica* 52(3): 631-645.

Sloan, Frank A., 1970, "Lifetime Earnings and Physicians' Choice of Specialty," *Industrial and Labor Relations Review* 24: 47-56.

Souleles, Nicholas S., 1999, "Consumer Sentiment: Its Rationality and Usefulness in Forecasting Expenditure," mimeo.

Souleles, Nicholas S., 2000, "Household Portfolio Choice, Hedging Motives, and Transactions Costs," mimeo.

Willis, Robert J. and Sherwin Rosen, 1979, "Education and Self-Selection," *Journal of Political Economy* 87(5): S7-S36.

Zarkin, Gary A., "Occupational Choice: An Application to the Market for Public School Teachers," *Quarterly Journal of Economics* 100(2): 409-446.

Table 1: Sample Means and Standard Deviations[18]

Variable	Mean	Standard Deviation
Age at graduation	26.9	2.93
Female	0.245	0.430
White	0.862	0.345
Part 1 NBME board exam score	204.4	17.9
Debt ($000)		
Expected income, in 1996 dollars ($000)		
5 years of experience: EY_5	101.9	61.5
10 years of experience: EY_{10}	153.8	77.0
20 years of experience: EY_{20}	189.4	92.4
Peak income: EY_{peak}	185.5	109.9
Expected hours worked/week	61.5	12.0
Expected percentage of time will devote to treating poor patients	22.6	16.0
Expected percentage of time will devote to teaching and research	22.3	18.3
Chosen specialty		
- internal medicine	0.199	0.399
- family practice	0.118	0.323
- pediatrics	0.053	0.225
- surgery	0.191	0.393
- ob/gyn	0.046	0.210
- radiology	0.031	0.174
- anesthesiology	0.023	0.151
- psychiatry	0.026	0.158
- pathology	0.016	0.127
- other	0.297	0.457

[18] Students who graduated after 1979 were not asked to predict their income 5, 10, and 20 years after completing residency training, nor asked to predict the number of hours they would work and how they would allocate those hours. Sample means for these variables reflect the responses from students who graduated before 1980.

Table 2: Determinants of Expected Peak Income ($000)

	Coefficient	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Female	-49.7*	3.17	- 37.6*	2.83	-30.2*	3.01	-27.3*	2.74
White	-9.73	5.16	10.0*	4.80	1.77	4.82	1.55	4.34
High board score	10.5*	4.38	5.02	3.93	7.96*	3.85	7.77*	3.47
Low board score	-1.76	3.91	4.22	3.53	3.60	3.42	2.98	3.07
National peak income, Y^N_{peak} ($000)			0.670*	0.0302	0.218*	0.0861	0.587*	0.077
Accuracy of income information ($000) $(Y^{N,est} - Y^N)$							0.837*	0.0467
Expected specialty (family practice is omitted)								
- internal medicine					16.2*	5.28	6.00	4.69
- pediatrics					- 4.71	3.18	5.17	2.96
- surgery					72.9*	11.4	33.6*	9.94
- ob/gyn					63.1*	9.57	29.5*	8.22
- psychiatry					14.1*	7.19	17.1*	6.47
Constant	191*	5.38	38.0*	8.04	100*	15.4	57.4*	14.0
Indicator variables for year of graduation						Included		Included
Observations	2,716		2,716		2,716		2,716	
R^2	0.07		0.25		0.30		0.45	

* = significantly different from zero at the five percent level

Table 3: Determinants of Income Expectations From Pooled Regressions ($000)

Dependent variable: expected income with 5 and 10 years of experience

	5 and 10 years		5 and 10 years	
	Coefficient	S.E.	Coefficient	S.E.
Female	-19.0*	3.48	-18.1*	3.40
High board score	2.48	3.57	3.66	3.50
Low board score	0.319	3.88	0.310	3.75
National income at t=0 ($000): $Y^N_{j,t=0}$	0.509*	0.0944	0.342*	0.0869
Accuracy of income information ($000)	0.402*	0.0339	0.393*	0.0336
Growth in national income between year 0 and year j ($000): $Y^N_{j,t=j} - Y^N_{j,t=0}$	0.270*	0.0482		
Year of income expectation (5 yrs omitted):				
- 10 years experience	23.7*	3.80	35.0*	3.29
Expected specialty indicator				
- internal medicine	-10.5*	3.65	-2.74	3.38
- pediatrics	7.07	4.76	4.25	4.77
- surgery	-34.2*	8.15	-0.509	5.88
- ob/gyn	-21.6*	6.84	-0.563	5.66
- psychiatry	-6.66	5.38	-1.31	5.27
Constant	37.8*	11.9	47.2*	11.2
Observations	1,440		1,525	
R^2	0.41		0.39	

Note: indicator variables for a student's graduation year are included in each regression.
* = significantly different from zero at the five percent level.

Table 4: Determinants of Income Expectations From Pooled Regressions ($000)

Dependent variable: expected income with 5,10, and 20 years of experience, and expected peak income

	Coefficient	S.E.
Female	-24.5*	4.66
High board score	5.81	4.12
Low board score	1.59	4.11
Expected hours worked/week	-0.0354	0.134
Percentage of time expect to devote to low-income patients	-0.123	0.125
Percentage of time expect to spend performing teaching and research	-0.357*	0.122
National income at t=0 ($000)	0.343*	0.094
Accuracy of income information ($000)	0.537*	0.0408
Year of income expectation (5 yrs omitted):		
- 10 years experience	35.3*	3.54
- 20 years experience	59.2*	5.28
- peak income	71.6*	5.95
Expected specialty indicator		
- internal medicine	4.52	3.89
- pediatrics	13.8*	5.58
- surgery	22.7*	7.57
- ob/gyn	18.4*	6.08
- psychiatry	7.62	6.23
Constant	40.4*	14.4
Observations	2,144	
R^2	0.54	

Note : indicator variables are included for the student's graduation year
* = significantly different from zero at the five percent level.

Table 5: Specialty Selection Coefficients: Probit Analysis

Variable	Contemporaneous Cross-section income Coefficient	S.E.	Reduced form Coefficient	S.E.	Selection-corrected, subjective income expectations Coefficient	S.E.	Subjective expectations and contemporaneous Coefficient	S.E.
A: Diff. in student's expected income (non-primary - primary care), $000					0.0165*	0.00860	0.0134**	0.00533
B: Diff. in contemporaneous income of MDs	0.00389**	0.00103						
B – A							-0.00332	0.00389
Female	- 0.468**	0.0637	- 0.467**	0.0648	- 0.125	0.0888	-0.0941	0.0901
Age	- 0.0773	0.0847	- 0.0614	0.0855	0.0853	0.109	0.105	0.110
White	- 0.272**	0.0771	- 0.304**	0.0807	-0.430**	0.102	-0.464**	0.103
Debt ($00000)	- 0.278	0.203	-0.336	0.209	-0.522*	0.271	- 0.451*	0.272
Board score	---	---	0.0434*	0.0262	---	---		
Student's relative error in perceived contemporaneous income	---	---	0.0833	0.425	---	---		
Constant	-1.09	1.07	-4.80*	2.87	-1.28	1.38	-1.23	1.39
Observations	2,458		2,458		2,458		2,458	
Log likelihood	-1,561		-1,542		-819		-816	
Pseudo R^2	0.027		0.038		0.49		0.49	
Percent predicted correctly	57.3		58.3		85.6		85.5	

Dependent variable is one if a fourth-year medical student chose a non-primary care specialty and zero if they chose a primary care specialty.
Quadratic terms for debt, age, and the student's board score are included. ** = significantly different from zero at the 5 percent level; * = significantly different from zero at the 10 percent level.

Table 6: Determinants of Peak Income Expectations ($000)

Variables	Non-primary Care	Primary Care
Female	-76.7**	-22.1**
	(14.8)	(4.83)
Board score	1.82	-0.916
	(4.15)	(1.28)
Student's error in perceived contemporaneous income of practicing physicians	0.655**	1.03**
	(0.123)	(0.122)
λ	72.3*	-20.0
	(37.0)	(18.1)
Constant	38.8	248*
	(431)	(127)
Observations	886	1,583
R^2	0.25	0.30

Notes:
(1) indicator variables are included for the year a student graduated from medical school.
(2) A quadratic term for the board score is included.
(3) ** = significantly different from zero at the 5 percent level;
(4) * = significantly different from zero at the 10 percent level.

Table 7: Specialty Selection – Subjective, Contemporaneous Income Expectations vs. Ex Post Income Expectations

Variable	Ex Post Income of Student's Cohort Coefficient	S.E.	Reduced Form Coefficient	S.E.	Selection-Corrected, Subjective Income Expectations Coefficient	S.E.	Subjective Expectations and Ex Post Income Coefficient	S.E.
A: Diff. in student's expected income (non-primary - primary care), $000					0.0112	0.0118	0.0205	0.0190
B: Ex post income diff. for student's cohort of MDs	-0.000506	0.00338						
B – A							0.00902	0.00854
Female	- 0.407**	0.111	- 0.451**	0.113	- 0.183	0.128	-0.191	0.129
Age	- 0.278**	0.138	- 0.336**	0.140	-0.282*	0.152	-0.264*	0.152
White	- 0.380**	0.154	- 0.392**	0.158	-0.582**	0.177	-0.602*	0.177
Debt ($00000)	- 0.198	0.425	0.0635	0.453	-0.180	0.482	-0.0820	0.487
Board score	---	---	0.0460	0.0391	---	---		
Student's relative error in perceived contemporaneous income	---	---	0.166	0.537	---	---		
Constant	0.356**	0.174	-0.258	4.26	3.54*	1.87	2.39	1.93
Observations	1,149		1,149		1,149		1,149	
Log likelihood	- 719		- 707		- 549		- 546	
Pseudo R^2	0.015		0.032		0.25		0.25	
Percent predicted correctly	50.6		58.1		72.8		74.0	

Dependent variable is one if a fourth-year medical student chose a non-primary care specialty and zero if they chose a primary care specialty.
Quadratic terms for debt, age, and the student's board score are included. ** = significantly different from zero at the 5 percent level; * = significantly different from zero at the 10 percent level.

Figure 1

Expected Peak Income in Family Practice, 1974-1998

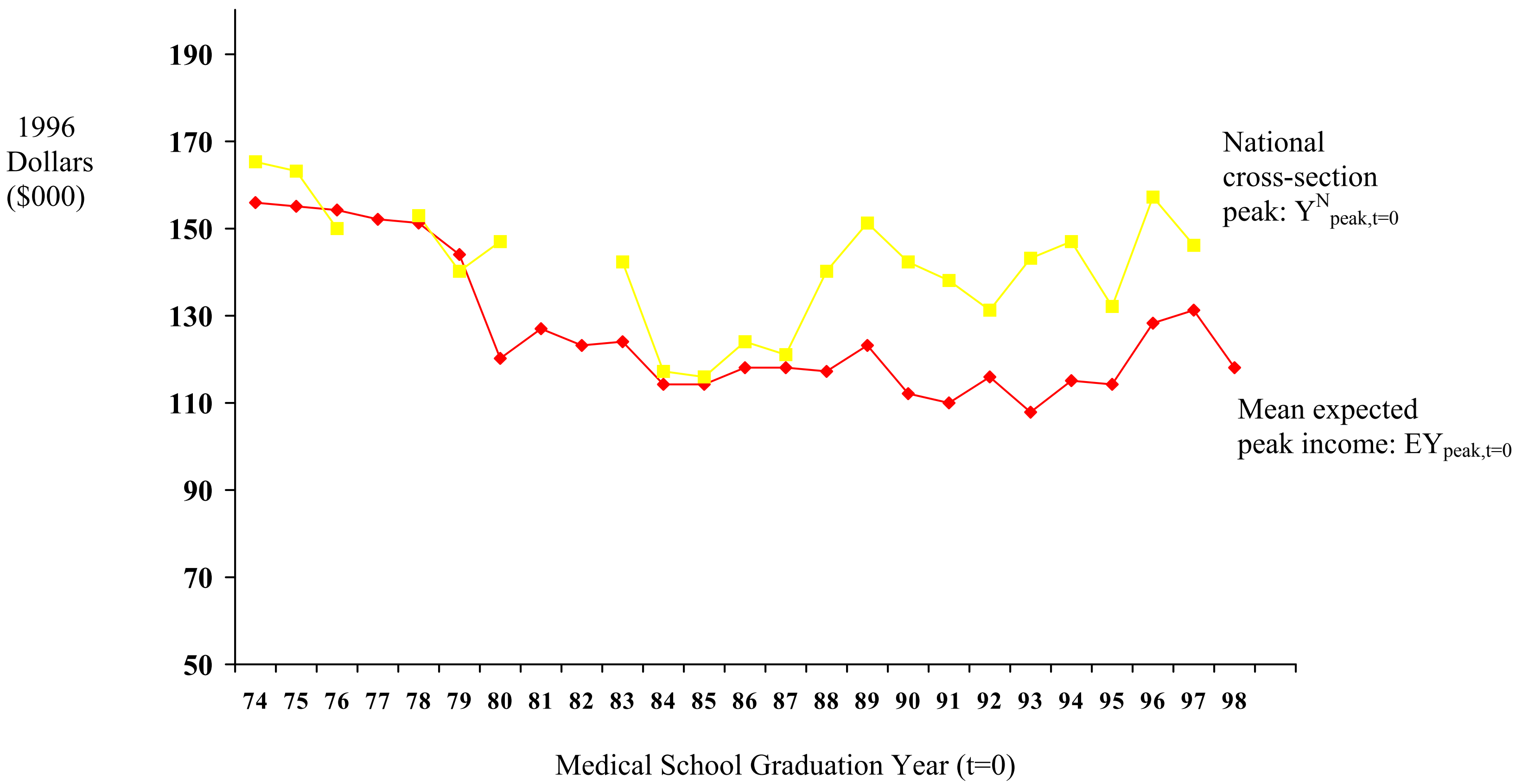


Figure 2

Expected Peak Income in Surgery, 1974-1998

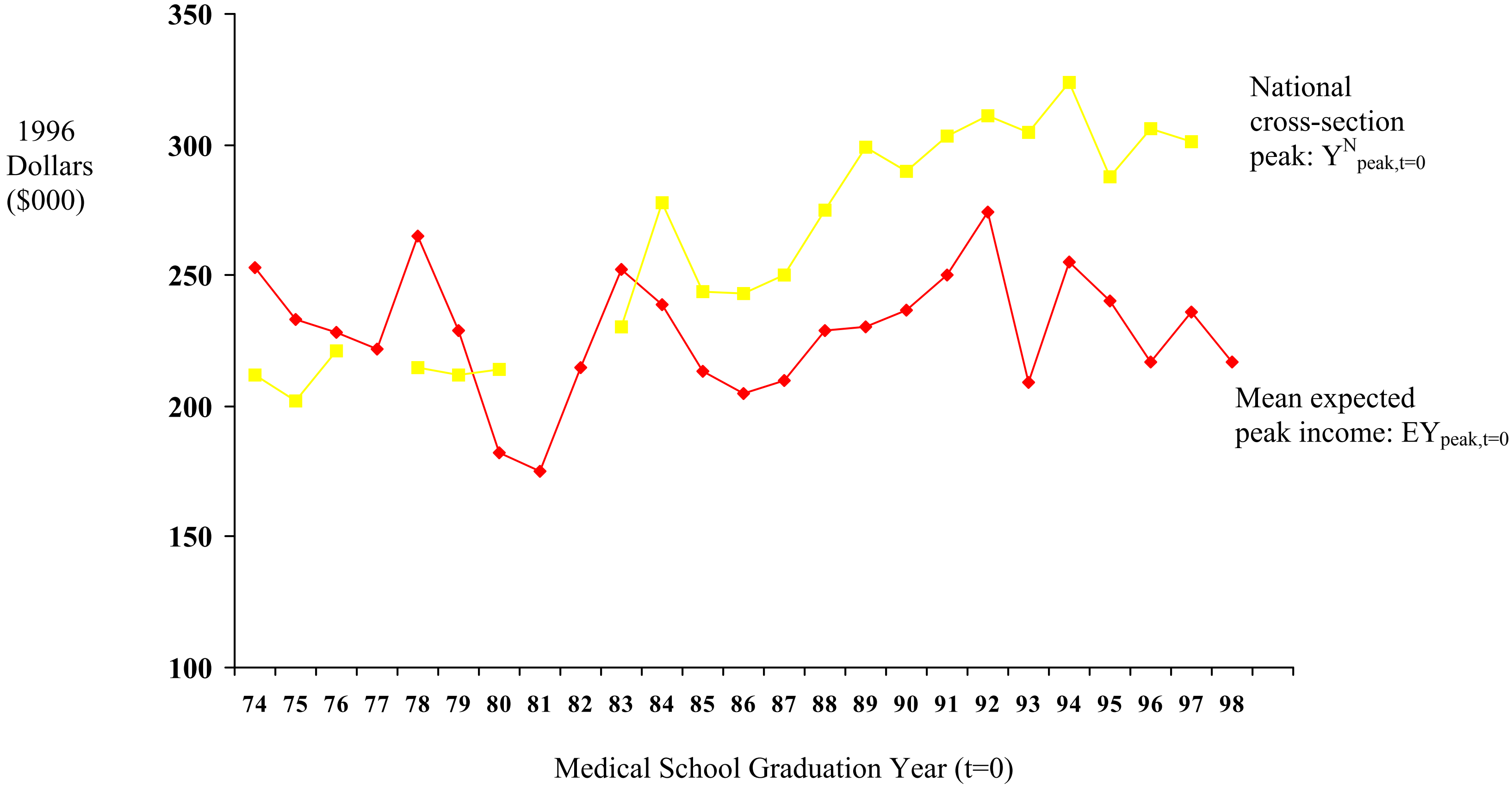
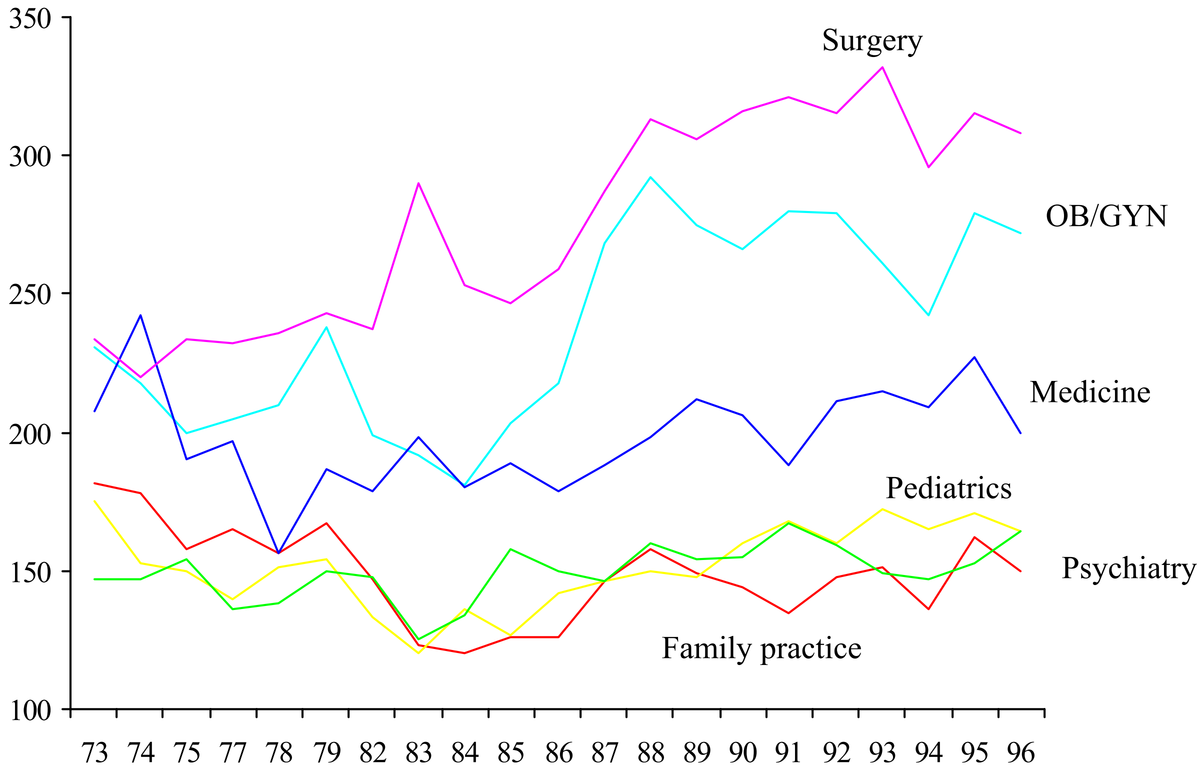


Figure 3

National Peak Physician Income by Specialty, 1973 - 1996
(1996 $000)



Source: AMA Socioeconomic Monitoring Study