On the Use of Expectations Data in Estimating Structural Dynamic Choice Models

Wilbert van der Klaauw, *Federal Reserve Bank of New York*

Despite the importance of expectations in models of decision behavior under uncertainty, few empirical economists have made use of subjective expectations data in estimating such models. Assuming that expectations about future behavior accurately portray optimal future behavior conditional on current information, it is shown that such data can provide similar information about the decision process as can data on current or retrospective behavior. The value of self-reported choice expectations is illustrated by using information on respondents' expected future occupation in the estimation of a structural dynamic model of teacher career decisions under uncertainty.

I. Introduction

Over the past 2 decades, economists have become increasingly involved in the collection, measurement, and analysis of subjective expectations. The interest in subjective expectations is not surprising given the importance of expectations in economic models of intertemporal decision making and in models of decision making under uncertainty more generally.

I have benefited from helpful comments and suggestions provided by Ken Wolpin, Chris Flinn, Basit Zafar, Jeff Dominitz, Peter Arcidiacono, Dana Goldman, Jim Walker, and seminar participants at Georgetown, Laval, New York, and Pennsylvania State Universities, University College London, and the University of Pennsylvania. The views and opinions offered in this article do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System as a whole. Contact the author at wilbert.vanderklaauw@ny.frb.org.

[*Journal of Labor Economics,* 2012, vol. 30, no. 3]

0734-306X/2012/3003-0002$10.00

These models typically relate the distribution of choices to the distribution of preferences and expectations in the population. The goal in standard revealed preference analysis is then to infer individuals' decision rules and preferences from observed choice behavior. However, as pointed out by Manski (1993, 2002), without placing much structure on the problem, preferences and expectations often cannot both be recovered from the choice distribution alone. The degree of underidentification is often severe, as shown, for example, by Magnac and Thesmar (2002) in the context of dynamic discrete decision models, with observed choices generally found to be consistent with several combinations of expectations and preferences.

Reflecting a relative scarcity of subjective expectations data, skepticism about their reliability, and an absence of an appropriate methodology for incorporating these data in the estimation of structural models, the approach prevalent in the economics literature has been to make strong untested assumptions on expectations and to infer preferences conditional on the maintained assumptions. Typically it is assumed that agents in the model use the same information as that available to the researcher, doing so in the same way, with expectations being rational with their subjective probability distributions coinciding with the true probability distributions.[1] Misspecification of the information set or of the expectations formation process generally will lead to biased preference parameter estimates (Manski 1993).

In fact, relatively little is known about how individual agents form expectations: about what is in their information set, how such information is used, and how expectations are updated over time. In recent years several economists have begun to collect and use subjective expectations data to investigate their validity, information content, and the way they are correlated with characteristics of individuals and their environments (Manski 2004). In addition, researchers have started to explore how such data can be combined with choice data to recover preferences under milder assumptions on how agents evaluate the likelihood of future events. While this promising line of research is relatively new and substantial hurdles remain, its potential for improving our limited understanding of a key element of many models of economic decision making has already become abundantly apparent.

[1] An alternative approach that also does not make use of expectations data but that does not impose explicit assumptions about the expectations formation process is that adopted by Houser, Keane, and McCabe (2004). In modeling and analyzing choice behavior in an experimental dynamic game, they specify expected total future payoffs (reflecting expected future choices and payoffs) as flexible functions of state variables. Using data on observed choices, they then use model estimates to categorize individuals' behavior into three different types, which they label "near-rational," "fatalistic," and "confused."

This article provides an illustration of the value of subjective expectations data in studying economic behavior. More specifically, it shows how frequently available expectations regarding future choice decisions can be incorporated into the estimation of structural dynamic choice models. Just as current choices are taken to portray optimal behavior given current information, expectations about future choices portray optimal future behavior conditional on current information. These data can therefore provide useful information about the decision process in the same way as do data on current or retrospective behavior. Like differences in actual choices, differences in reported expectations can therefore be explained using the same behavioral model.

The expectations data used in this study represent respondents' expectations about their personal occupation and employment status several years into the future. While showing the value of such data in estimating a structural dynamic model of teacher career choices, it is important to note that the methodology adopted here is applicable to the study of other choice decisions and to other types of expectations data as well. The recent study by van der Klaauw and Wolpin (2008) represents another example of the applicability of this approach. In estimating a structural model of retirement and saving decisions, it exploits subjective expectations data on a large set of future events, including individuals' expected date of retirement, expected social security benefits, and self-reported longevity and bequest expectations.

In this study I show how expectations data can be combined with data on actual choices to obtain more precise parameter estimates, while assuming that the two data sources used are consistent, that is, assuming that the expectations data were generated by the same model governing the actual choices.[2] In addition, along the lines proposed by Wolpin and Gonul (1985), I will use estimates of the model obtained from data on observed behavior alone to test whether the reported expectations, which must be a function of the same structural parameters, are consistent with this model.

This article is organized as follows. The next section provides a brief discussion of the recent literature on the use of expectations data in studying economic decision behavior. A simple dynamic model of occupational choice and career mobility is presented in Section III. Section IV describes

[2] While there are important differences, in some respects this approach of incorporating subjective data in estimating a structural model is similar to the use of subjective information on reservation wages in estimating job search models, as in Lancaster and Chesher (1983) and Flinn and Del Boca (1984). In that literature, reservation wage data are typically used to identify some of the model parameters, while in the current case the expectations data represent overidentifying information. See Wolpin (1999) for a related discussion of potential efficiency gains derived from using data on choice expectations.

the data set and the estimation of the model, and this is followed by a brief discussion of the parameter estimates. Section V describes the self-reported expectations data, presents validation tests of these data, and discusses the manner in which they can be incorporated in the estimation of the structural model. Estimates obtained after incorporating the expectations data are also presented. Finally, Section VI offers some concluding comments and areas for future research.

II. Earlier Studies Using Expectations Data

Over the past 2 decades there has been a marked increase in interest among economists in the measurement and analysis of individuals' subjective expectations (Manski 2004).[3] A number of large-scale consumer surveys, such as the National Longitudinal Surveys, the Panel Study of Income Dynamics, and especially the Health and Retirement Study, have elicited respondents' subjective expectations and intentions about various future life events or choices, such as mortality, fertility, retirement, income, schooling, and occupation. More specialized surveys, such as the Survey of Economic Expectations (Dominitz and Manski 1997a, 1997b) and surveys conducted as part of the Federal Reserve Bank of New York's Household Inflation Expectations Project (Bruine de Bruin et al. 2010), have elicited respondents' subjective probability distributions of various personal and macroeconomic events. In addition to asking for a respondent's point forecast or a "percent chance" assigned to binary outcomes, these surveys elicit probability distributions reflecting respondents' beliefs and uncertainty about future realizations of continuous variables.

Increased data collection and improved measurement have led to a rapidly growing number of studies involving subjective expectations data. Studies in which such data have been used can be broadly divided into two groups. The first group has been mainly concerned with testing the properties and validity of the reported expectations and analyzing their determinants and covariates. Many of these studies aim to test for "rationality," that is, whether expectations are unbiased and use all available information, by comparing expectations to actual realizations. Some studies of this type include Griliches (1980), Hamermesh (1985), Bernheim (1988, 1989, 1990), Hurd and McGarry (1995, 2002), Honig (1996), and Hurd, Smith, and Zissimopoulos (2004). More recently, a number of studies have begun to focus on how individuals form and update their expectations and to document and analyze the substantial heterogeneity in beliefs across respondents. Surprisingly, given the role of forward-looking

[3] There has been a long history of collecting expectations data, such as those used to generate the University of Michigan Consumer Sentiment Index and the Conference Board's Consumer Confidence Index, but these data have been used mainly for descriptive and prediction purposes.

behavior in economic theories, relatively little is known about what information individuals possess and use in forecasting future outcomes and about the way forecasts are formed. Some studies that have begun to investigate these issues include Dominitz (1998), Benitez-Silva and Dwyer (2005), Dominitz and Manski (2005), Lochner (2007), Delavande (2008a), Stinebrickner and Stinebrickner (2008), Bruine de Bruin et al. (2010), Galati, Heemeijer, and Moessner (2010), and Zafar (2011b).

In the second group of studies, self-reported expectations about future events or decisions have been used to help explain observed choice decisions. In earlier studies of this type, most analyses were reduced form in nature. For example, Bernheim and Levin (1989) used subjective expectations about future social security benefits to help explain current savings behavior. Sandell and Shapiro (1980), Shaw and Shapiro (1987), Gronau (1988), and Blau and Ferber (1991) used reported plans of future labor market separations and subjective preferences for future labor force participation in testing the human capital theory of job and occupational sex segregation. The use of expectations data in estimating such reduced-form model specifications raises a number of important concerns.

In a dynamic framework, expectations of future decisions and outcomes are functions of current information sets and thus will generally depend on the same observables and unobservables that affect current decisions. For example, expected future social security benefits will depend on the planned date of retirement and expected future earnings up to that date, which will generally have the same determinants as current work and savings decisions. Thus preferences and skills are likely to determine both current saving behavior and future social security benefits. As a result, treating subjective expectations as exogenous explanatory variables is likely to lead to endogeneity biases.[4] In some cases it may also be difficult, if not impossible, to disentangle the causal effects of expectations on current actions and vice versa. For example, do planned labor force separations lead to lower human capital investment on the job and a choice of jobs with lower wages and flatter wage-earnings profiles, or do lower wages and flatter wage profiles lead to higher quit rates, or are both the case?

More recently, a number of studies have endeavored to use subjective expectations data in the structural estimation of simple choice models. These studies have used such data to help overcome the identification problem of inferring preferences from observed choice behavior under

[4] Lochner (2007) explicitly deals with the potential endogeneity of the perceived probability of arrest in evaluating its effect on criminal behavior by using an instrumental variables approach. After first differencing out unobserved fixed effects, (twice) lagged criminal and arrest histories of the individual and the individual's siblings serve as instruments for beliefs about the probability of arrest.

uncertainty. For example, Delavande (2008b) used expectations data on, among other things, perceived risks of pregnancy to study college students' choice of birth control method, while Zafar (2009) used expectations on future earnings and several other outcomes to analyze students' choice of college major. Other examples include Bellemare, Kröger, and van Soest (2008), who analyzed choices in ultimatum games; Arcidiacono, Hotz, and Kang (2010), who studied college major choices; Blass, Lach, and Manski (2010), who estimated preferences for electricity reliability; Armantier et al. (2011), who assessed whether individuals make investment choices based on their inflation expectations; and Delavande and Kohler (2011), who investigated how risky sexual behavior in a high-HIV-prevalence environment is influenced by individuals' survival expectations. In these studies, it is assumed that individuals maximize the expected returns or benefits associated with a set of alternatives, where the net return or utility level associated with a given choice is specified explicitly as a function of preferences and expected outcomes. By directly using survey data on individuals' subjective expectations of outcomes, preferences can be recovered without requiring assumptions on how agents evaluate the likelihood of future events, therefore reducing the risk of misspecification biases.

While an important step forward, there are several important limitations to the approach adopted in these studies. First, its applicability is restricted to choice models in which relevant expected future returns can be fully captured by a finite set of measurable summary statistics. For example, if realizations of outcomes are correlated across choice alternatives, one generally would need to measure the entire joint subjective distribution of future choice-specific outcomes.[5] Furthermore, while the use of subjective expectations data in the estimation of structural choice models has been limited to models that are essentially static, most dynamic decision problems are sequential in nature. For example, in the case of college major choice or occupational choice models, in making current choice decisions individuals may consider the option to switch in the future.[6] Generally, in making current choices, individuals may consider future benefits associated with sequences of future choice decisions. In that case, a comparison of returns or utility levels associated with choice alternatives would involve consideration of the expected outcomes conditional on any possible sequence of choices up to that future period, as well as the probabilities of making these sequential choices. Clearly, in general the

[5] Note that unless one relies on risk neutrality of preferences, a comparison of expected returns would generally involve measuring not just the means but the whole outcome distributions conditional on choosing each alternative.

[6] Zafar (2009) and Arcidiacono et al. (2010) need to rule out such switching in their empirical models.

data requirements for fully measuring all relevant expectations would be a daunting, if not impossible, task.

A second limitation concerns the likely endogeneity of reported expectations. As discussed earlier, reported expectations are likely to reflect unobserved preference heterogeneity. Therefore, even if one actually could measure all relevant subjective expectations, they could not simply be treated as exogenous explanatory variables. For example, reported expectations about the likelihood of getting pregnant when using a particular contraceptive would likely reflect expected efforts to reduce the risk of pregnancy, which in turn would capture preferences for becoming (or not becoming) pregnant.[7] Solutions to the endogeneity problem generally require additional knowledge or assumptions regarding the expectations formation process.[8]

Third, the approach does not require one to understand how individuals form expectations. However, without an explicit model describing how expectations are formed, knowing preferences by themselves would not be sufficient for addressing many interesting policy questions. Generally, to conduct counterfactual policy analyses with the goal of predicting behavior under a variety of conditions, one would need to understand and take into account how a new set of conditions will affect individual expectations.

These different limitations imply that, in modeling most intertemporal choice decisions or decision making under uncertainty, it will be difficult to circumvent altogether the need to impose some structure on the expectations formation process and on the way in which expectations may affect behavior. To fruitfully use subjective expectations data to explain observed choice behavior, one generally will need to explicitly model the expectations formation process jointly with a model of how expectations affect current choice behavior.

Instead of specifying a model that, with the available expectations data, can be estimated without having to make any assumptions regarding the expectations formation process, in this article I take a more conventional approach in specifying a dynamic model of teacher career decisions where individuals are assumed to have rational expectations and to maximize expected lifetime utility. After estimating the model using observed choice data, I then evaluate whether reported subjective expectations about future

[7] There is also potential for endogeneity caused by cognitive dissonance, where individuals report beliefs that are consistent with their behavior, as well as estimation biases due to reporting errors such as those associated with rounding (Zafar 2011a).

[8] For example, to address this issue Bellemare et al. (2008) model the way beliefs and unobserved preferences are correlated. Similarly, Delavande and Kohler (2011) specify a recursive joint model of choice behavior and beliefs and impose exclusion restrictions.

occupation and employment status are consistent with the expectations implied by the model. Finally, I explore how such data can be integrated into the estimation of the model. In the next section I begin by presenting a simple model of teacher career choices.

III. A Dynamic Model of Teacher Career Decisions

The model presented below characterizes each individual's initial occupational choice decision of whether or not to become a teacher, as well as subsequent occupational mobility decisions (i.e., exit out of and reentry into teaching) in each year since graduating from a teacher training program. These career choices are constrained by the arrival of teaching job offers. The model also incorporates the labor force participation decision itself to explain temporary exits (particularly of women) from the labor market. Each occupational choice and work decision involves a trade-off between pecuniary and nonpecuniary rewards in the teaching and nonteaching sectors, as well as the utility derived when not working in the labor market. Because individuals face uncertainty about current and future economic conditions, these career decisions involve a formation of expectations about future earnings, nonpecuniary benefits, and employment opportunities in each occupation. In this sense the model is similar to those of Gotz and McCall (1985) and Keane and Wolpin (1997).

Upon graduating from a teacher training program, each graduate is assumed to maximize the present value of utility over a known finite horizon (T) by choosing whether to work as a teacher (if such a job is available), work in the nonteaching sector, or choose not to work in the labor market. The objective of the individual is to maximize

$$E\sum_{t=1}^{T}\delta^{t-1}U(P_t, C_t), \tag{1}$$

where the utility function is specified as

$$U(P_t, C_t) = \alpha C_t - b_{1t}I(P_t = 1) - b_{2t}I(P_t = 2), \tag{2}$$

by choosing a path $\{(P_t \in I_t, C_t \in \Re); t = 1, \dots, T\}$, where the choice decision P_t equals $P_t = 0$ if the individual opts for the nonmarket alternative, $P_t = 1$ when choosing to work as teacher, and $P_t = 2$ if deciding to work in the nonteaching sector; and C_t represents consumption in period t of a composite good; $I(\ )$ is the indicator function, with $I = 1$ if the argument is true and $I = 0$ if not; I_t represents the set of choice options for P_t in period t; δ is the subjective discount factor; and E is the expectations operator.

In the specification of the utility function, α represents the marginal utility of consumption and b_{1t} and b_{2t} represent the disutility of working in the two sectors of the labor market relative to the utility of staying at home. The disutility of working in each occupation (which could be negative) will depend on the individual's preferences for each different type of work and on the nonpecuniary benefits provided by the occupation. To model this, we specify

$$b_{kt} = X'\beta_{k1} + S'_{kt}\beta_{k2} + u_{kt}, \qquad k = 1, 2, \tag{3}$$

where X is a vector of individual characteristics, including the individual's race, sex, type of degree obtained, and a constant term. The vector S_{kt} includes the time-varying variables age and the individual's total number of years of work experience $\exp_{kt}$ in occupation k since graduation from a teacher training program. Occupation-specific work experience evolves over time according to the law of motion:

$$\exp_{kt} = \exp_{kt-1} + I(P_{t-1} = k),\ \exp_{k0} = 0,\ k = 1, 2. \tag{4}$$

The disutility and nonpecuniary benefits associated with working in occupation or sector k is thus allowed to depend on the individual's work experience, age, and characteristics X. This dependence reflects both differences across individuals in tastes for working in occupation k and the varying degree of access within each occupational sector to jobs with higher nonpecuniary benefits. The stochastic components u_{kt} in equation (3) represent unobserved individual differences in preferences and nonpecuniary returns in period t, which, as discussed below, can be serially correlated.

The period-specific budget constraint is given by

$$C_t = N_t + W_{1t}\mathcal{I}(P_t = 1) + W_{2t}\mathcal{I}(P_t = 2), \tag{5}$$

where N_t represents nonlabor income in period t and W_{kt} are the wage earnings an individual receives in period t when choosing occupation k. Wage earnings in each employment sector depend on total work experience in that occupation, a vector Z of individual characteristics affecting the earnings in occupation k, and a quadratic trend in calendar time (with yr_t representing the calendar year corresponding to period t), to capture a trend in average salary levels over time:

$$\begin{aligned} W_{1t} = {} & Z'\gamma_{11} + \gamma_{12}\exp_{1t} + \gamma_{13}\exp_{1t}^2 \\ & + \gamma_{14}\exp_{1t}^3 + \gamma_{15}\text{yr}_t + \gamma_{16}\text{yr}_t^2 + \nu_{1t}. \end{aligned} \tag{6}$$

$$\begin{aligned} W_{2t} = {} & Z'\gamma_{21} + \gamma_{22}\exp_{2t} + \gamma_{23}\exp_{2t}^2 + \gamma_{24}\exp_{2t}^3 \\ & + \gamma_{25}\exp_{1t} + \gamma_{26}\text{yr}_t + \gamma_{27}\text{yr}_t^2 + \nu_{2t}. \end{aligned} \tag{7}$$

The vector Z includes a constant, the individual's race, sex, types of degrees obtained, and SAT score. It further includes the state's average manufacturing wage earnings over the sample period, as an indicator of the average strength of regional demand for labor. Teacher salary schedules differ from school district to school district but within a school district depend solely on educational background and teaching experience. The vector of individual characteristics Z was included in the teacher wage equation to allow for the possibility that teachers with desirable characteristics may be able to obtain jobs in better-paying school districts. The average state's manufacturing wages were included in the teacher wage equation as a (crude) proxy for variations in the average teaching salary across states and school districts. Note that while nonteaching wages may depend on teaching experience, teacher salaries do not depend on $\exp_{2t}$ as actual teacher salary schedules do not depend on nonteaching work experience.

Earnings in each occupation are further stochastic, depending on a random component ν_{kt} with mean zero, representing stochastic fluctuations in earnings over time. At the time of each period's choice decision, each individual knows both the current value of W_{kt} in each sector k and the wage structure in (6) and (7) but does not know the future values of W_{kt}.

The correlation structure of the different error terms in the model is specified as follows:

$$u_{kt} = \mu_k + \omega_t, \qquad k = 1, 2, \tag{8}$$

$$\nu_{kt} = \kappa_k \mu_k + \xi_{kt}, \qquad k = 1, 2, \tag{9}$$

where μ_k denotes a person- and alternative-specific time-invariant heterogeneity component and κ_k are wage-specific factor loadings. The component ω_t represents transitory unobserved changes in the disutility of working across individuals and over time, and the ξ_{kt} are individual-specific transitory wage shocks. The three transitory random components, ξ_{1t}, ξ_{2t}, and ω_t, are assumed to be joint normally distributed with variance-covariance matrix Σ, to be independently distributed over time and individuals, and to be uncorrelated with μ_1 and μ_2.[9]

The distribution of the permanent unobserved heterogeneity components μ_1 and μ_2 is specified to be discrete joint multinomial. Accordingly, we distinguish between J different "types" of individuals, where each type j, $j = 1, \cdots, J$, is characterized by a different vector $\underline{\mu}_j = (\mu_{1j}, \mu_{2j})$. The population proportions of each type are given by $q_j = \Pr(\mu_1 = \mu_{1j}, \mu_2 = \mu_{2j})$, $j = 1, \cdots, J$. In the estimation of the model I allow for four types of individuals, who differ in the values of μ_1 and μ_2, each of which can take on two different values, representing a low or high preference or value of working in each occupation.[10] The population proportions are defined as

[9] Identification requires a normalization of one of the parameters. Var(ω_t) was therefore fixed to one.

[10] See van der Klaauw (1996) for a similar specification of the unobserved heterogeneity distribution.

$$\Pr(\mu_1 = 0, \mu_2 = 0) = q_1,$$

$$\Pr(\mu_1 = \rho_1, \mu_2 = 0) = q_2,$$

$$\Pr(\mu_1 = 0, \mu_2 = \rho_2) = q_3,$$

$$\Pr(\mu_1 = \rho_1, \mu_2 = \rho_2) = 1 - q_1 - q_2 - q_3.$$

Note that, due to the permanent heterogeneity components, the errors u_{kt} and v_{kt} may be correlated across time and, when μ_1 and μ_2 are correlated, also across choice alternatives, even if all off-diagonal elements of Σ are zero.

One aspect of each period's occupational choice decision that has not yet been discussed concerns the definition and evolution over time of the choice set I_t. During the seventies and eighties, the time period covered by our data, the number of individuals seeking and applying for teaching jobs greatly exceeded the number of vacancies in teaching. Rather than assume that each individual has the option to work as a teacher in each period, I will therefore incorporate the lack of vacancies and existence of search frictions by allowing for the possibility that the choice set I_t may not include the teaching option in some periods. In addition, I will allow the probability of such an event to vary across individuals by characterizing the realization of a teaching job offer in each period by an arrival rate that depends on a vector of individual characteristics Y_t, containing the individual's race, degree background, age, and teaching experience. It is further assumed that all individuals currently teaching (with $P_{t-1} = 1$) will always have the option to remain in teaching. Given that during the sample period of our data few teachers were laid off, I do not believe this to be a very restrictive assumption. Accordingly, the arrival rate is specified as

$$\Pr(I_t = J_0 | P_{t-1} = 1) = 1,$$

$$\Pr(I_t = J_0 | P_{t-1} = k) = \Phi(Y_t'\omega), \qquad k = 0, 2,$$

$$\Pr(I_t = J_1 | P_{t-1} = k) = 1 - \Pr(I_t = J_0 | P_{t-1} = k), \qquad k = 0, 1, 2,$$

where $J_0 = \{P_t \in (0, 1, 2)\}$, $J_1 = \{P_t \in (0, 2)\}$, and $\Phi(\cdot)$ is the standard normal distribution function. The probability of receiving a teaching job offer in each period is assumed to be known to the individual.

In deciding each period whether to work in the teaching, nonteaching, or household sector, the individual compares the sum of current and expected discounted future utility associated with each option. Expected future utility in turn depends on the expected future growth in wage earnings and in nonpecuniary benefits, that is, on the rate of return to total and occupation-specific work experience, in each sector. The dependence of wage earnings, the disutility of working (and nonpecuniary benefits of working), and future teaching job offer arrival rates on the individual's employment history therefore causes an individual to consider in the current decision its effects on future utility levels and choices through a change in work experience. If work experience accumulated in one occupational sector has a lower wage return in the other, we can expect occupational mobility to decline with the number of years in the labor market. A high return to work experience will also lead to an increase in the opportunity cost of leaving the labor force.

The dynamic programming solution is then as follows. Substituting the budget constraint into the utility function, utility equals

$$\bar{U}_t(P_t) = \begin{cases} \alpha N_t & \text{when } P_t = 0, \\ \alpha(N_t + W_{1t}) - b_{1t} & \text{when } P_t = 1, \\ \alpha(N_t + W_{2t}) - b_{2t} & \text{when } P_t = 2. \end{cases}$$

The individual's maximization problem in each period t, $t = t_0, \dots, T$ can then be stated as follows:

$$\max_{\{d_{ks} \in I_s, s \geq t\}} E\left[\sum_{s=t}^{T} \delta^{s-t} \sum_{k=0}^{2} \bar{U}_s(k) d_{ks} | \Omega_t\right], \tag{10}$$

where Ω_t is the relevant information set or state space in period t, containing all factors known to the individual in that period that either affect current returns or the probability distribution of future returns and where $d_{ks} = 1$ if alternative k is chosen in period s and $d_{ks} = 0$ if not, and $\sum_{k=0}^{2} d_{ks} = 1$.

An alternative "reduced form" representation of the maximization problem can be obtained by substituting both earnings equations into the utility function in (9). The utility levels associated with each choice alternative can then be defined as

$$\bar{U}_t(k) = \begin{cases} \alpha N_t & \text{when } k = 0, \\ \alpha N_t + X_t'\lambda_1 + (\alpha\kappa_1 - 1)\mu_1 + \varepsilon_{1t} & \text{when } k = 1, \\ \alpha N_t + X_t'\lambda_2 + (\alpha\kappa_2 - 1)\mu_1 + \varepsilon_{t2} & \text{when } k = 2, \end{cases}$$

where the reduced-form coefficients λ_i are functions of the utility and the occupation-specific earnings equations parameters and the vector $\mathbf{X}_t$ consists of all explanatory variables in equations (3), (6), and (7) combined. The composite errors are defined as $\varepsilon_{kt} = \alpha\xi_{kt} - \omega_t$ and, given the distributional assumptions made earlier, are joint normally distributed.

Given the utility function specification above, the maximum expected present discounted value of lifetime utility at time t, $t < T$, equals

$$V_t(\Omega_t) = \max_{i \in I_t} [\bar{U}_t(i) + \delta E[V_{t+1}(\Omega_{t+1}) | d_{it} = 1, \Omega_t]], \tag{11}$$

where the information set Ω_t at time t contains the current realizations of the error terms ε_{it}, the vector $\mathbf{X}_t$ (which includes measures of the decision history until t), the values of μ_1 and μ_2, and the choice set I_t. The expectation in (11) is taken with respect to all stochastic components in

Ω_{t+1}, including the realization of next period's choice set (i.e., the arrival of teaching job offers), and the realization of the stochastic earnings and utility components, conditional on Ω_t and $d_{it} = 1$.

It is possible to derive all $V_t(\Omega_t)$ functions $t = 1, \ldots, T$ and to solve for the optimal policy at each t by exploiting the finite horizon nature of the dynamic programming problem. In period T, we have $V_T(\Omega_T) = \max_{j \in I_T}[\bar{U}_T(j)]$. Further, for each period $t < T$ and for each state vector X_t and error vector $\underline{\mu}$, we can define two values $\varepsilon^*_{kt}(X_t, \underline{\mu})$, $k = 1, 2$, such that

$$\begin{aligned}\delta E[V_{t+1}(\Omega_{t+1})|d_{0t} = 1, X_t, \underline{\mu}] - \varepsilon^*_{kt} &= X'_t\lambda_k + (\alpha\kappa_k - 1)\mu_k \\ &+ \delta E[V_{t+1}(\Omega_{t+1})|d_{kt} = 1, X_t, \underline{\mu}].\end{aligned} \tag{12}$$

Then the optimal policy for each information vector X_t and heterogeneity vector $\underline{\mu}$ when the choice set $I_t = J_0$ equals

$$\begin{cases} d_{1t} = 1, d_{0t} = 0, d_{2t} = 0 & \text{iff } \varepsilon_{1t} \geq \varepsilon^*_{1t}(X_t, \underline{\mu}) \text{ and} \\ & \varepsilon_{1t} - \varepsilon_{2t} \geq \varepsilon^*_{1t}(X_t, \underline{\mu}) - \varepsilon^*_{2t}(X_t, \underline{\mu}), \\ d_{2t} = 1, d_{0t} = 0, d_{1t} = 0 & \text{iff } \varepsilon_{2t} \geq \varepsilon^*_{2t}(X_t, \underline{\mu}) \text{ and} \\ & \varepsilon_{1t} - \varepsilon_{2t} < \varepsilon^*_{1t}(X_t, \underline{\mu}) - \varepsilon^*_{2t}(X_t, \underline{\mu}), \\ d_{0t} = 1, d_{1t} = 0, d_{2t} = 0 & \text{iff otherwise.} \end{cases} \tag{13$'$}$$

And when $I_t = J_1$,

$$\begin{cases} d_{2t} = 1, d_{0t} = 0, d_{1t} = 0 & \text{iff } \varepsilon_{2t} \geq \varepsilon^*_{2t}(X_t, \underline{\mu}), \\ d_{0t} = 1, d_{1t} = 0, d_{2t} = 0 & \text{iff } \varepsilon_{2t} < \varepsilon^*_{2t}(X_t, \underline{\mu}). \end{cases} \tag{13$''$}$$

The two values ε^*_{1t} and ε^*_{2t} divide the two-dimensional space up into three regions, in each of which one (assuming no ties) of the alternatives is optimal. Given the specified normal distribution for the ε_{kt}'s, the decision rule in each period, the terminal value function V_T, and the Bellman equation (11), it is possible to solve, by backward recursion, for all $V_t(\Omega_t)$ functions and all ε^*_{kt} values. Note that this involves the calculation of the expectations $E[V_{t+1}(\Omega_{t+1})|d_{it} = 1, X_t, \underline{\mu}]$, each of which involves the evaluation of a bivariate normal integral.

IV. Data and Estimation

To estimate the model, I will use data from the National Longitudinal Study of the High School Class of 1972 (NLS-72). This study surveyed over 22,000 high school seniors in 1972 and includes five additional follow-up surveys through 1986, at which point most members were in their early thirties. While the final 1986 follow-up survey was limited to only a subset of the original NLS-72 participants, the sample design oversampled teachers and potential teachers by including all those who had previously reported having completed teacher training. Therefore, the NLS-72 surveys combined

Table 1
Descriptive Statistics

Variable	Mean	SD/ (Frequency)	Number of Observations
Sample of 817 individuals:			
Years in sample	9.093	1.514	817
Age in first period	22.717	1.165	817
$\exp_{11}$	.039	.272	817
$\exp_{21}$	.078	.375	817
Nonwhite	.075	(61)	817
Female	.736	(602)	817
BEd	.811	(662)	817
MEd	.052	(43)	817
MA	.028	(23)	817
Science	.021	(17)	817
SAT	926.7	184.0	817
Manufacturing wage	17.740	2.583	817
Sample of 7,428 person-year observations:			
age_t	26.754	2.793	7,428
$\exp_{1t}$	2.019	2.360	7,428
$\exp_{2t}$	1.469	2.080	7,428
W_{1t}	15.806	4.744	2,207
W_{2t}	16.883	7.465	1,738
$P_t = 1$	.450	(3,342)	7,428
$P_t = 2$	.363	(2,693)	7,428

NOTE.—Teacher earnings, W_{1t}, are calculated for the sample of teachers with nonmissing wage information. Earnings in the nonteaching sector, W_{2t}, are calculated for workers with nonmissing earnings information in the nonteaching sector only. Both earnings are in thousands of 1982 dollars. Frequencies are in parentheses. All entries are weighted using the sample weights. See the data appendix for definitions of the other acronyms.

provide a valuable source for the study of the early career decisions and mobility patterns of a cohort of teachers. The analysis will be restricted to the subsample of individuals who were part of the final 1986 follow-up survey and who became eligible or qualified to teach, that is, graduated from a teacher training program, during the 1976–79 period. The latter group is defined to include all individuals who received at least one of the following (*a*) a bachelor's degree in education, (*b*) a master's degree in education, or (*c*) a teaching certificate. The first observation year for each individual in the sample is then the year in which the individual became qualified to teach and left full-time education. While the final observation year for most individuals is the final survey year 1986, for a small number it is instead the year after which information about their career history was missing or incomplete. For the resulting unbalanced panel of 817 individuals, the average number of years available is about 9 years per individual. Summary statistics of the variables used in the study are given in table 1. Definitions of these variables are provided in the data appendix.

For each individual k the choice of each alternative i is observed for T_k periods. In those periods in which the individual works, wage earnings are also observed. Let the decision set for individual k be $\underline{d}_t^k = [d_{0t}^k, d_{1t}^k, d_{2t}^k]$ and $\mathbf{d}^k = [\underline{d}_1^k, \ldots, \underline{d}_{T_k}^k]$, where d_{it}^k specifies the actual choice of alternative i for individual k at time t. Thus, $\underline{d}_t^k$ is the vector defining the alternative chosen at time t by individual k, and d^k is the vector describing the choice sequence over the individual's observed sample period. Further, let $\mathbf{w}_1^k = [W_{11}^k, \ldots, W_{1T_k}^k]$ and $\mathbf{w}_2^k = [W_{21}^k, \ldots, W_{2T_k}^k]$ be the sequences of the teacher and nonteacher earnings observed for individual k, elements of which will be zero (missing) if in that period the individual did not work in that sector or if earnings data are missing.

The objective is to estimate the structural parameters, θ, given the observed data on the individuals' choices and occupation-specific earnings, where θ includes the utility function parameters (α and the β_{kj} parameters), the parameters in the two earnings equations (γ_1 and γ_2), the teaching job offer probability parameters (ω), the discount factor (δ), and the error distribution parameters, ρ_1, ρ_2, $\{q_j, j = 1, \cdots, J\}$, κ_1, κ_2, and Σ.

Estimates of the structural parameters of the model can be obtained using relatively standard maximum likelihood methods.[11] Given the optimal policy in (13′) and (13″), it is possible to calculate for each pair of vectors $(X_t, \underline{\mu})$ the probability that alternative i is chosen in period t as

$$\begin{aligned}\Pr(d_{it} = 1|\mathrm{X}_t, \underline{\mu}) = \Upsilon \Pr(d_{it} = 1|\mathrm{X}_t, \underline{\mu}, J_0) \\ + (1 - \Upsilon)\Pr(d_{it} = 1|\mathrm{X}_t, \underline{\mu}, J_1),\end{aligned} \tag{14}$$

where $\Upsilon = \Pr(I_t = J_0|P_{t-1} = k)$ is the arrival rate of teaching job offers defined earlier. The choice probabilities $\Pr(d_{it} = 1|\mathrm{X}_t, \underline{\mu}, J_k)$ for each choice set J_k and alternative i are equal to the probability that the values of the two normally distributed error terms ε_{1t} and ε_{2t} satisfy the conditions described in (13′) and (13″). The calculation of these choice probabilities therefore requires the evaluation of a bivariate normal integral. For example,

$$\begin{aligned}\Pr(d_{1t} = 1|\mathrm{X}_t, \underline{\mu}, J_0) &= \Pr[\varepsilon_{1t} \geq \varepsilon_{1t}^*(\mathrm{X}_t, \underline{\mu}), \\ &\quad \varepsilon_{1t} - \varepsilon_{2t} \geq \varepsilon_{1t}^*(\mathrm{X}_t, \underline{\mu}) - \varepsilon_{2t}^*(\mathrm{X}_t, \underline{\mu})] \\ &= \int_{\varepsilon_{1t}^*}^{\infty} \int_{-\infty}^{\varepsilon_{1t}+\varepsilon_{2t}^*-\varepsilon_{1t}^*} \phi(\varepsilon_{1t}, \varepsilon_{2t}) d\varepsilon_{2t} d\varepsilon_{1t}.\end{aligned} \tag{15}$$

where $\phi(\cdot, \cdot)$ represents the joint normal density function of ε_{1t} and ε_{2t}.

The likelihood function for our sample of K individuals is then defined as

[11] For reviews of solution and estimation methods for similar dynamic programming models, see Eckstein and Wolpin (1989) and Rust (1991, 1994, 1996).

$$L(\theta) = \prod_{k=1}^{K} L_k = \prod_{k=1}^{K}\sum_{j=1}^{J} L_{kj}q_j = \prod_{k=1}^{K}\sum_{j=1}^{J} \Pr(\mathbf{d}^k, \mathbf{w}_1^k, \mathbf{w}_2^k|\theta, \underline{\mu}_j)q_j,$$

$$= \prod_{k=1}^{K}\sum_{j=1}^{J} (\Pr[\underline{d}_{T_k}^k, W_{1T_k}^k, W_{2T_k}^k|\underline{d}_{T_k-1}^k, \dots, \underline{d}_2^k, \underline{d}_1^k]$$

$$\times \cdots \times \Pr[\underline{d}_2^k, W_{12}^k, W_{22}^k|\underline{d}_1^k] \Pr[\underline{d}_1^k, W_{11}^k, W_{21}^k])q_j,$$

where the conditioning on θ and $\underline{\mu}_j$ in the second equation has been omitted to simplify notation. The joint probability terms can further be written as the product of a conditional and marginal probability as follows:

$$\Pr[\underline{d}_t^k, W_{1t}^k, W_{2t}^k| \cdot] = \Pr[\underline{d}_t^k| \cdot , W_{1t}^k, W_{2t}^k] \Pr(W_{1t}^k, W_{2t}^k| \cdot).$$

Each of the choice probabilities $\Pr[\underline{d}_t^k| \cdot , W_{1t}^k, W_{2t}^k]$ is equal to the probability that the chosen alternative is the optimal one (given the employment history and the values of the current period's wage offers), which is equal to the probability, for each possible choice set I_t, that the draw of the $(\varepsilon_{it})_{i\in I_t}$ vector falls in the region of the (ε_t) space where the chosen alternative is optimal. With normally distributed ε's, the likelihood function equals the product of weighted averages of multinomial probit probabilities such as the one in (15).[12] Thus, estimating the model involves calculating these probabilities for each individual and time period. As we saw earlier, the backward recursive solution to the dynamic programming problem will provide us with these probabilities.

The estimates of the structural parameters and their standard errors are shown in table 2.[13] Considering first the earnings equation estimates, the most interesting results are the much higher returns in the nonteaching sector for having a master's degree, a science degree, and a higher SAT score, as well as the relatively large gender wage gap in the nonteaching sector relative to the teaching sector. The estimate of α, the marginal utility of consumption, is large and positive significant, which implies that wage considerations are important in decisions to enter and remain in teaching. The positive coefficients of $\exp_{1t}$ and $\exp_{2t}$ indicate that the disutility of working in either sector increases with previous work experience, and the positive coefficient on age implies that utility associated with working in the teaching sector declines with age.

The arrival rate parameter estimates show that the probability of receiving a teaching job offer was greater for those with a bachelor's degree

[12] In computing choice probabilities that condition on an observed wage, the relevant normal joint density in (15) is conditional on the implied value of one of the wage error terms ν_{kt}.

[13] The discount factor was fixed at 0.9. The finite horizon T corresponds to age 45. Note that the maximum observed age in our panel is 33, which, given a discount rate of 0.9, suggests that the results are unlikely to be very sensitive to an increase in T.

in education and for individuals who were somewhat older. Those with more teaching experience, on the other hand, were less likely to receive a teaching job offer than those with less teaching experience, possibly reflecting the trade-off between hiring better and more experienced teachers and hiring less costly inexperienced teachers. The error covariance estimates reveal a positive correlation between the two wage errors of about 0.6 and negative correlations between the disutility of working error u_t and the two wage errors. The estimates of the heterogeneity distribution parameters reveal the presence of significant permanent unobserved heterogeneity.

V. Self-Reported Expectations Data

Like many other surveys of individuals, the NLS-72 includes several questions regarding the respondent's expectations or intentions about future events or decisions. To illustrate the value and use of such data, we focus here only on one question in which individuals were asked about their career expectations. More specifically, the expectations data to be used in this study are the responses of the panel members to a question posed in the survey year 1979. In that year all individuals who participated in the NLS-72 were asked about their expected occupation and labor force status at the age of 30. The exact question asked was: "What kind of work will you be doing when you are 30 years old? (circle *one* that comes closest to what you expect to be doing)." Given an average age in 1979 of 25, the expectation therefore refers on average to 5 years in the future. In addition to the homemaker/not working and school teacher options, individuals could choose from a list of 15 additional occupations, including clerical work, craftsman, farmer, manager, services, sales, and others. For the purposes of this study, the answer to this question asked in period t will be represented by the variable ES_t defined as

$$\mathrm{ES}_t = \begin{cases} 0 & \text{if not working,} \\ 1 & \text{if school teacher,} \\ 2 & \text{if in an nonteaching occupation.} \end{cases}$$

Table 3 provides cross tabulations of the responses with both the individual employment status in the survey year 1979 and with the actual labor force status at age 30. The fact that the diagonal elements in the bottom part of the table are generally much larger than the off-diagonal elements clearly indicate that the expectations data contain information about actual future behavior.[14] The top part of the table also indicates that

[14] Note that even in the absence of aggregate shocks, differences between the mean expected and the actual proportions choosing each state do not imply that the expectations are not rational (see Manski 1990).

Table 2
Estimates of Life Cycle Model

Symbol	Variable	Estimate	SE
Utility function parameters:			
α	C_t	.426*	.070
β_{10}	Constant	−.297	.548
β_{111}	Nonwhite	−.275	.230
β_{112}	BEd	−.372*	.148
β_{113}	MEd	.981*	.445
β_{114}	MA	.996*	.378
β_{115}	Science	−.878	.649
β_{116}	Female	.257	.153
β_{12}	age_t	.164*	.031
β_{13}	$\exp_{1t}$	.370*	.058
β_{20}	Constant	3.922*	.839
β_{211}	Nonwhite	−.258	.188
β_{212}	BEd	−.238*	.122
β_{213}	MEd	.456	.351
β_{214}	MA	.376	.316
β_{215}	Science	.411	.467
β_{216}	Female	−.195	.135
β_{22}	age_t	.033	.026
β_{23}	$\exp_{2t}$	.089	.055
Arrival rate teaching jobs:			
ω_1	Constant	−.985	1.041
ω_2	$\exp_{1t}$	−.095*	.020
ω_3	yr_t	−.246*	.053
ω_4	age_t	.083	.050
ω_5	Nonwhite	.102	.108
ω_6	BEd	.170*	.089
ω_7	MEd	−.150	.276
ω_8	MA	−.471	.331
ω_9	Female	.083	.085
Error covariance matrix:			
$\text{Cov}(\omega_t, \xi_{1t})$		−3.346*	.365
$\text{Var}(\xi_{1t})$		25.786*	2.195
$\text{Cov}(\omega_t, \xi_{2t})$		−6.664*	.451
$\text{Cov}(\xi_{1t}, \xi_{2t})$		19.691*	2.340
$\text{Var}(\xi_{2t})$		53.513*	3.844
Teacher earnings equation:			
γ_{111}	Constant	8.561*	.862
γ_{112}	Nonwhite	.711*	.272
γ_{113}	BEd	1.260*	.236
γ_{114}	MEd	2.400*	.554
γ_{115}	MA	.469	1.946
γ_{116}	Science	3.295*	.768
γ_{117}	SAT	−1.317*	.383
γ_{118}	Female	−1.009*	.199
γ_{12}	$\exp_{1t}$	.612*	.240
γ_{13}	$\exp^2_{1t}$	.280	.572
γ_{14}	$\exp^3_{1t}$	−.690	.411
γ_{15}	yr_t	−1.175*	.158
γ_{16}	yr^2_t	1.097*	.126
γ_{17}	Manufacturing wage	1.562*	.270
Nonteacher earnings equation:			
γ_{211}	Constant	2.173	1.234
γ_{212}	Nonwhite	.842	.554
γ_{213}	BEd	.244	.360
γ_{214}	MEd	3.810*	.610
γ_{215}	MA	4.763*	.535
γ_{216}	Science	3.227*	.935
γ_{217}	SAT	2.787*	.531

Table 2 (*Continued*)

Symbol	Variable	Estimate	SE
γ_{218}	Female	−4.651*	.267
γ_{22}	$\exp_{2t}$	1.516*	.229
γ_{23}	$\exp_{2t}^2$	−.706*	.355
γ_{24}	$\exp_{2t}^3$	.343*	.142
γ_{25}	$\exp_{1t}$	−.101	.076
γ_{26}	yr_t	−1.623*	.191
γ_{27}	yr_t^2	1.297*	.157
γ_{28}	Manufacturing wage	2.841*	.520
Heterogeneity distribution:			
ρ_1		2.273*	.245
ρ_2		.127	.168
κ_1		2.314*	.324
κ_2		66.059*	88.016
q_2		.407*	.021
q_3		.176*	.019
q_4		.242*	.022
δ	Discount factor	.90	
L	Log likelihood	16,761.2	

NOTE.—For definitions of the acronyms, see the data appendix.
* Significant at the 5% level.

the expectations data provide information beyond that contained in the individual's current labor force status.

I will interpret the answer to the posed question on the expected occupation and labor force status at the age of 30 to represent the choice alternative that at the current date has the greatest probability of maximizing the individual's utility at age 30, that is, as the alternative with the greatest probability of being chosen at age 30 (i.e., the mode).[15] With this interpretation, it is clear that these expectations or intentions data contain information about individual choice behavior. Future behavior will depend in part on conditions known to the individual at the time of the survey and in part on events that have not yet occurred and are not perfectly foreseeable. In our model the actual stochastic process generating these subsequent events (the random preference shock u_t, the arrival of teaching job offers, and future wage shocks ν_{1t} and ν_{2t}) has been specified up to a vector of unknown parameters. Given these specifications and the associated optimal decision rules (13′) and (13″), each future period's choice probabilities can be calculated for each possible work history in that future period. Consequently, it is possible to calculate the age 30 choice probabilities conditional on the current period's work history. These future choice probabilities will be a function of the same parameters that determine the current choice probabilities and work decisions.

[15] Juster (1966), Manski (1990), and Blass et al. (2010) similarly interpret stated-choice responses as representing choice alternatives with the highest probability. Delavande and Rohwedder (2011) test this interpretation for reported expected social security claiming ages and find the reported point forecasts to be consistent with the mode of individual forecast distributions for 90% of respondents.

Table 3
Current, Expected, and Actual Future Occupation at Age 30

	Expected Status at Age 30			
	Homemaker/ Not Working	School Teacher	Other Specified Occupation	Total
Status in 1979:				
Not working	18 (.22, .26)	31 (.38, .09)	33 (.40, .10)	82 (.11)
Teaching job	32 (.08, .46)	281 (.68, .77)	100 (.24, .29)	413 (.53)
Nonteaching job	20 (.07, .29)	52 (.19, .14)	208 (.74, .61)	280 (.36)
Status at age 30:				
Not working	37 (.22,. 53)	71 (.41, .20)	64 (.37, .19)	172 (.22)
Teaching job	12 (.04, .17)	209 (.70, .57)	78 (.26,. 23)	299 (.39)
Nonteaching job	21 (.07, .30)	84 (.28, .23)	199 (.65, .58)	304 (.39)
Total	70 (.09)	364 (.47)	341 (.44)	775

NOTE.—Row and column percentages are given in parentheses. Each individual was asked the following question in October 1979: "What kind of work will you be doing when you are 30 years old? (circle *one* that comes closest to what you expect to be doing)." In addition to the homemaker/not working and school teacher option, a list of 15 additional occupations was given, including clerical work, craftsman, farmer, manager, services, sales, and so forth.

More formally, given the specified structure of the individual's maximization problem and given values of the parameters, the expected probability of choosing a particular alternative at age 30 corresponds to the probability that the error terms ε_1 and ε_2 in the corresponding period take values such that inequalities (13′) or (13″) hold, where this probability is calculated conditional on the current information set. This structure therefore allows us to calculate these future choice probabilities for each individual (and each type). Under the assumption that the behavioral model is correct (and ignoring sampling variation that causes the estimated parameters to differ from the true parameters), the alternative with the largest choice probability, that is, the most likely choice at age 30 at the current date, should then equal each individual's self-reported most likely choice at age 30.

Let us define the calculated or implied expected choice probabilities at age 30, given current information, as P_0^*, P_1^*, and P_2^*, where $P_j^* = \Pr(d_{jt+m} = 1 | X_t, \underline{d}_t, \underline{\mu}_l)$ for $j = 0, 1, 2$, where $t + m$ represents the year in which the individual is 30 years old. Then, with ES_t representing the expected (or most likely) choice in period $t + m$ reported in year t, as defined above, we have (for each type $\underline{\mu}_l$)

$$\Pr(ES_t = i | X_t, \underline{d}_t, \underline{\mu}_l) =$$

$$\Pr(ES_t = i | P_0^*, P_1^*, P_2^*) = \begin{cases} 1 & \text{iff } i = \arg\max \{P_j^*\} \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

for all $i = 0, 1, 2$, where the $\Pr(d_{jt+m} = 1|X_t, \underline{d}_t, \mu_l)$, $j = 0, 1, 2$ can be calculated as described earlier.[16]

Incorporation of these probabilities in the likelihood function will make the likelihood function discontinuous and nondifferentiable.[17] This problem is resolved once we allow for the possibility that individuals make errors in reporting their expectations.[18] It is likely that respondents may not take sufficient time to give a precise answer when responding to survey questions about expectations but use more precise forecasts when making actual career choices and in reporting choices made. While individuals are assumed to calculate future choice probabilities (the $\Pr(d_{jt+m} = 1|X_t, \underline{d}_t, \mu_l)$) correctly, instead of reporting the maximum of these probabilities, we assume that they report each alternative with probability

$$\Pr(ES_t = i|X_t, \underline{d}_t, \mu_l) = \frac{e^{\Pr(d_{it+m}=1|X_t,\underline{d}_t,\mu_l)/r}}{\sum_{j=0}^{2} e^{\Pr(d_{jt+m}=1|X_t,\underline{d}_t,\mu_l)/r}}$$

$$= \frac{e^{P^*_i/r}}{\sum_{j=0}^{2} e^{P^*_j/r}}, \quad i = 0, 1, 2. \tag{17}$$

Note that as $r \to 0$ these probabilities will approximate those in (16), that is, if $r = 0$, individuals would in fact report the alternative with the greatest expected future probability.[19] Thus r provides a measure of the degree of misreporting.

Note that in comparison to (16), the degree by which the choice probabilities in (17) will differ from one and zero will depend on how similar to each other the future choice probabilities are. If one alternative clearly has the greatest future probability of being chosen, the probability that the individual will report that alternative will be close to one. On the other hand, when two choices are almost equally likely to be chosen in

[16] Note that

$$\Pr(d_{jt+m} = 1 | X_t, \underline{d}_t, \mu_l) = \sum_{X_{t+m}} \Pr(d_{jt+m} = 1 | X_{t+m}, \underline{d}_t, \mu_l) \Pr(X_{t+m} | X_t, \underline{d}_t, \mu_l).$$

[17] A similar problem arises in the case of the maximum score estimator of Manski (1975, 1985). There the goal is to choose parameter values that maximize the number of correct choice predictions, where a prediction is either correct or incorrect. The likelihood function becomes a step function, complicating the maximization routine as well the derivation of the asymptotic properties of the estimator.

[18] Bernheim (1988, 1990) finds indirect evidence of the existence of reporting errors in expectations. In the job search literature, subjective reservation wages are similarly assumed to be measured with error.

[19] Note that when r becomes small, our allowance for reporting errors has the same effect or plays the same role as the smoothing method proposed by Horowitz (1992) to overcome the discontinuous and nondifferentiable likelihood problem for the maximum score estimator.

future period $t + m$ (in which case it may be more difficult for the individual to determine the one with the maximum probability), the reported expected future state could be either with equal probability and the probability of a reporting error will be greatest.

The expectations data can now be incorporated into the likelihood function to obtain

$$L(\theta) = \prod_{k=1}^{K} \sum_{j=1}^{J} \Pr(d^k, w_1^k, w_2^k, \mathrm{ES}_t^k | \theta, \underline{\mu}_j) \times q_j, \tag{18}$$

where

$$\Pr(d^k, w_1^k, w_2^k, \mathrm{ES}_t^k | \;\cdot\;) = \Pr[\underline{d}_{T_k}^k, W_{1T_k}^k, W_{2T_k}^k | \underline{d}_{T_k-1}^k, \ldots, \underline{d}_2^k, \underline{d}_1^k]$$

$$\times \cdots \times \Pr[\underline{d}_{s+1}^k, W_{1s+1}^k, W_{2s+1}^k | \underline{d}_s^k, \ldots, \underline{d}_2^k, \underline{d}_1^k] \times$$

$$\Pr[\mathrm{ES}_s^k | \underline{d}_s^k, \underline{d}_{s-1}^k, \ldots, \underline{d}_2^k, \underline{d}_1^k] \times \Pr[\underline{d}_s^k, W_{1s}^k, W_{2s}^k | \underline{d}_{s-1}^k, \ldots, \underline{d}_2^k, \underline{d}_1^k]$$

$$\times \Pr[\underline{d}_{s-1}^k, W_{1s-1}^k, W_{2s-1}^k | \underline{d}_{s-2}^k, \ldots, \underline{d}_2^k, \underline{d}_1^k] \times \cdots \times$$

$$\Pr[\underline{d}_2^k, W_{12}^k, W_{22}^k | \underline{d}_1^k] \Pr[\underline{d}_1^k, W_{11}^k, W_{21}^k]$$

and s equals the year in which the expectation about year $s + m$ was reported (where for notational convenience we have omitted a superscript k on s and m).

When incorporating the expectations data into the likelihood function, we implicitly assume that the expectations data are consistent with observed individual behavior and with the specified behavioral model. This may in fact not be the case. It may be that respondents did not understand the question or that they provided random responses, thereby invalidating the expectations data, and as discussed later, some respondents may not have reported the most likely future choice in the way assumed. Use of these data in that case could lead to biased estimates.

One way to test for the validity of the reported expectations is to compare the reported expectations with actual realizations. In our case we could simply compare the proportions of individuals expecting to work as teacher, work as a nonteacher, or expecting not to work at age 30 with actual choices at age 30. Table 3 shows that the reported expectations do in fact correspond reasonably well to the actual choices at age 30. The off-diagonal counts can then be explained by the fact that the sample size is relatively small or by reporting errors or by the fact that the predicted choices are based on estimated parameters. However, as

Table 4
Predicted, Expected, and Actual Future Occupation at Age 30

	Predicted Status at Age 30 (Model)			
	Homemaker/ Not Working	School Teacher	Other Specified Occupation	Total
Expected status at age 30:				
Not working	25 (.36, .27)	21 (.30, .06)	24 (.34, .07)	70 (.09)
Teaching job	43 (.12, .46)	235 (.65, .70)	86 (.24, .25)	364 (.47)
Nonteaching job	25 (.07, .27)	81 (.24, .24)	235 (.69, .68)	341 (.44)
Actual status at age 30:				
Not working	71 (.41, .76)	61 (.35, .18)	40 (.23, .12)	172 (.22)
Teaching job	4 (.01, .04)	267 (.89, .79)	28 (.09, .08)	299 (.39)
Nonteaching job	18 (.06, .19)	9 (.03, .03)	277 (.91, .83)	304 (.39)
Total	93 (.12)	337 (.43)	345 (.45)	775

NOTE.—Row and column percentages are given in parentheses.

pointed out by Manski (1990), such validity or rationality tests are invalid in the case of binary intentions data, such as those considered here.[20]

As a second validation test, we can test whether the subjective responses are consistent with the optimal future behavior as implied by the behavioral model and the objective data on actual choices. Using the estimated parameters, we can determine the alternative with the maximum expected future choice probability as explained earlier for each of the J types (unobserved heterogeneity values). Further, given our estimates we can assign type probabilities to each individual. Using Bayes's rule, the probability that an individual k is type j is $q_j \times L_{kj}/L_k$. We can then compare each individual's self-reported expected future choice with that predicted by the model for the individual's most likely type. A good fit would validate the subjective expectations question, under the assumption that our model is correct. Small differences between the reported and predicted choices can be explained by the fact that the prediction was based on an (imprecise) estimate of the individual's type and on estimated parameters, and by the presence of reporting errors.

Table 4 gives a cross tabulation of the reported responses with the predicted choices implied by the model. There is a fairly close correspondence between the two. A chi-square test rejects their equality at the

[20] A simple example will make this clear. If all individuals forecast their future probabilities of choosing the teaching, nonteaching, and not working states to be 0.33, 0.33, and 0.34, then all would report to expect not to work at age 30 (the mode) even though in fact only approximately 34% will turn out doing so.

95% level but not at the 99% level.[21] The second part of the table shows that the predictions implied by the model are always closer to actual behavior at age 30 than the self-reported expectations, which may be an indication of the existence of reporting errors, but should also not be very surprising given that the model was estimated using the actual choice data (including the choices at age 30). Overall, the table shows that the model is able to explain both actual future choices and reported intentions data quite well.

So far, we have assumed that individuals were asked to choose from among the three different choice alternatives considered in our model (not working, teaching, nonteaching occupation). However, in the survey, individuals were provided a larger choice set, which included several different nonteaching professions. It is easy to show that answers may differ when a larger choice set is offered instead of the three alternatives considered in our model. In our case, it may not be unreasonable to assume, however, that in answering the question the individuals in our sample (who are all qualified teachers) adopted a two-stage approach consistent with the model: one where in the first stage the probabilities of working as teacher, nonteacher, and not working are compared and the alternative with the greatest probability is identified. Then, in the second stage, if the individual chose the nonteaching sector (i.e., the probability of working in the nonteaching sector at age 30 is the greatest), the individual selects the most likely alternative from among the 15 different nonteaching occupations.

It is important to stress that this assumption about the way in which an individual provides an answer to a particular question is much less of an ad hoc assumption than it may initially appear. When using the actual choice data (where individuals report their current occupation by choosing from the same list of occupations) to estimate the model, we have similarly implicitly assumed that individuals choose from among the three sectors in the two-stage manner described, and we similarly ignore the second-stage choice decision and the data on the actual nonteaching occupation chosen.[22] For example, if someone reports to be employed as manager in a particular year, we similarly interpret this in the context of our model as though the individual had chosen the nonteaching sector. Thus, both actual choice data and expectations data are treated entirely symmetrically.

Incorporating the expectations data with reporting errors, the likelihood function is exactly that in (18). Estimates are presented in table 5. In general, they are very similar to those in table 2, providing additional

[21] The χ^2 statistic is 7.9, while $\chi^2(2, 0.05) = 5.99$ and $\chi^2(2, 0.01) = 9.21$.

[22] It is interesting to note that while these type of assumptions about the decision process are commonly made in order to match data with a proposed theoretical model, they are almost never explicitly stated.

evidence that the expectations data are consistent with the observed choice data and with the behavioral model. The reporting error variance is 0.33 and is significantly different from zero. In general, the parameter estimates have smaller standard errors than those in table 2 (on average they are 5% smaller), reflecting the efficiency gains obtained from combining subjective expectations data with objective data on actual choice decisions.

Besides the gain in efficiency, a related benefit from using subjective expectations data in the way described in this article concerns an increase in accuracy in forecasting future individual behavior and outcomes. As shown above, data on choice expectations provide valuable information about an individual's unobserved "type" or about unobserved characteristics. Using Bayes's rule, one can derive posterior probabilities of each individual's type, which can be used to improve forecasting of future individual behavior.

The same applies to models in which individuals' unobserved perceptions or beliefs of the state of the economy or state of the world are modeled through some (possibly time-varying) latent variable. Reported expectations can contain information not available to the econometrician through other observables that could help improve forecasting accuracy and in estimating latent variables. A recent illustration of such an approach is Del Negro and Eusepi (2010), who use inflation expectations data from the Survey of Professional Forecasters to improve the accuracy of forecasts generated by their DSGE model. In their case, inflation expectations bring information about the public's unobserved beliefs about the central banks' inflation target.

VI. Conclusion

Many individual- or household-level surveys elicit respondents' expectations about future events or choices. Recently there has been an increased interest in the analysis and collection of such information by economists. Finding that expectations data contain valuable information, there is growing awareness of the potential for such data to make a substantial contribution to our understanding of intertemporal decision making under uncertainty. The most fruitful approach, in my view, is to use such data not only for explaining choice behavior but also for analyzing how expectations are formed. Generally this would require imposing some structure on the expectations formation process and modeling this jointly with current choice behavior and its dependence on expectations.

This article represents a first exploration in this direction by presenting a methodology for the incorporation of subjective data on choice expectations in the estimation of stochastic dynamic choice models. While applied to a study of teacher career decisions, it is generalizable to other life cycle decisions. Using information about self-reported career expec-

Table 5
Estimates of the Life Cycle Model Using Expectations Data

	Variable	Estimate	SE
Utility function parameters:			
α	C_t	.435*	.065
β_{10}	Constant	.546	.573
β_{111}	Nonwhite	−.474*	.231
β_{112}	BEd	−.441*	.155
β_{113}	MEd	.929*	.397
β_{114}	MA	.623*	.288
β_{115}	Science	−1.043	.657
β_{116}	Female	.415*	.143
β_{12}	age_t	.121*	.028
β_{13}	$\exp_{1t}$	.390*	.060
β_{20}	Constant	3.198*	.875
β_{211}	Nonwhite	−.208	.174
β_{212}	BEd	−.220	.121
β_{213}	MEd	.396	.370
β_{214}	MA	.156	.406
β_{215}	Science	.371	.526
β_{216}	Female	−.127	.116
β_{22}	age_t	−.020	.023
β_{23}	$\exp_{2t}$	.105	.059
Arrival rate teaching jobs:			
ω_1	Constant	−.783	.869
ω_2	$\exp_{1t}$	−.085*	.018
ω_3	yr_t	−.235*	.045
ω_4	age_t	.071	.041
ω_5	Nonwhite	.106	.092
ω_6	BEd	.199*	.080
ω_7	MEd	−.109	.260
ω_8	MA	−.288	.521
ω_9	Female	.033	.083
Error covariance matrix:			
$\text{Cov}(\omega_t, \xi_{1t})$		−4.142*	.287
$\text{Var}(\xi_{1t})$		26.054*	1.949
$\text{Cov}(\omega_t, \xi_{2t})$		−6.974*	.311
$\text{Cov}(\xi_{1t}, \xi_{2t})$		21.521*	2.156
$\text{Var}(\xi_{2t})$		54.725*	3.647
r	Reporting error variance	.328*	.024
Teacher earnings equation:			
γ_{111}	Constant	9.399*	.860
γ_{112}	Nonwhite	.846*	.268
γ_{113}	BEd	1.291*	.232
γ_{114}	MEd	2.274*	.546
γ_{115}	MA	−.411	1.446
γ_{116}	Science	3.181*	.844
γ_{117}	SAT	−1.062*	.377
γ_{118}	Female	−.984*	.197
γ_{12}	$\exp_{1t}$	.591*	.238
γ_{13}	$\exp_{1t}^2$	.517	.578
γ_{14}	$\exp_{1t}^3$	−.960*	.424
γ_{15}	yr_t	−1.333*	.162
γ_{16}	yr_t^2	1.245*	.130
γ_{17}	Manufacturing wage	1.429*	.263
Nonteacher earnings equation:			
γ_{211}	Constant	2.110	1.167
γ_{212}	Nonwhite	.719	.545
γ_{213}	BEd	.152	.356
γ_{214}	MEd	4.047*	.647
γ_{215}	MA	4.698*	.530
γ_{216}	Science	3.291*	.985

Table 5 (*Continued*)

	Variable	Estimate	SE
γ_{217}	SAT	2.380*	.484
γ_{218}	Female	−4.659*	.257
γ_{22}	$\exp_{2t}$	1.655*	.228
γ_{23}	$\exp_{2t}^2$	−1.036*	.364
γ_{24}	$\exp_{2t}^3$	.494*	.147
γ_{25}	$\exp_{1t}$	−.234*	.074
γ_{26}	yr_t	−1.506*	.192
γ_{27}	yr_t^2	1.305*	.159
γ_{28}	Manufacturing wage	3.100*	.494
Heterogeneity distribution:			
ρ_1		2.412*	.217
ρ_2		.285*	.143
κ_1		1.911*	.214
κ_2		26.933*	13.726
q_2		.381*	.022
q_3		.171*	.019
q_4		.277*	.023
δ	Discount factor	.90	
L	Log likelihood	−17,240.8	

* Significant at the 5% level.

tations, it was shown that such data could be readily incorporated in the estimation of the model under similar assumptions required to analyze objective choice data. While the efficiency gain from incorporating data from a single expectations question in our application was rather modest, one can expect this gain to become more substantial as the number of incorporated expectations increases.

An issue not explicitly addressed in this article concerns the general quality of subjective expectations data. While the interpretation of the answers to the expectation question used here seems logical, it is clear that there is a need for more carefully worded and more detailed expectations questions. For example, to avoid any ambiguity about whether a question or response relates to a mean, median, or mode of a variable (and also to measure uncertainty about future outcomes), it would be preferable to elicit information about each individual's complete subjective probability distribution of future realizations as in, for example, Dominitz and Manski (1997b), Engelberg, Manski, and Williams (2009), and Bruine de Bruin et al. (2011). As effectively illustrated by Blass et al. (2010), elicited choice probabilities are much more informative than stated choices (which represent simple point forecasts), allowing respondents to express their uncertainty about their future behavior.

It also would be useful if the question spelled out in more detail what an expected probability or the expectation should be conditioned on. For example, when asking someone whether or not they expect to work at age 65 (or the probability of such an event), it may not be obvious to the interviewee whether the question is conditional or unconditional on surviving to age 65, especially for individuals with an illness. More generally, these issues and the application presented in this article point to importance of gathering expectations data in a model specific context.

Finally, an important topic for future research is to study how and whether expectations data could be used to relax some of the assumptions inherent in most structural dynamic models of decision behavior under uncertainty, about the way in which expectations are formed. While in this article we have maintained the assumption that expectations are rational, expectations data could help identify the ways in which different agents form and update expectations. For example, they could be used in estimating models that incorporate adaptive learning or models with heterogeneity across individuals in expectations formation, with some being rational, others adaptive or using another boundedly rational approach, as in the rationally heterogeneous expectations model of Branch (2004).

Data Appendix

The National Longitudinal Study of the High School Class of 1972 (NLS-72), surveyed over 22,000 high school seniors in 1972 and surveyed this group until 1986, when most members were in their early thirties. After the first base year questionnaire in 1972, five follow-up surveys were held, in 1973, 1974, 1976, 1979, and 1986. In addition, the final survey included a special teacher supplement, which focused on the 1,517 individuals in the sample who, during the 1972–86 period, had taught or had become qualified to teach. The NLS-72 surveys combined provide a valuable source for the study of the early career decisions and mobility patterns of a cohort of high school and college graduates. It also contains detailed information on wages and educational background, including measures of academic ability and course subjects. Most important for our study, the NLS-72 population includes a relatively large national sample of school teachers, thereby representing one of the most comprehensive sources of information on the labor market experiences of school teachers.

I restrict my analysis to the subsample of individuals who were part of the fifth follow-up survey and who became eligible or qualified to teach, that is, graduated from a teacher training program during the 1976–79 period. I define the latter group to be all individuals who received at least one of the following: (*a*) a bachelor's degree in education, (*b*) a master's degree in education, or (*c*) a teaching certificate. The first observation year for each individual in the sample is then the year in which the individual became qualified to teach and had left full-time education. The final observation year for most individuals is the final survey year 1986, but for a small number it is the year after which information about their career history was missing or incomplete. For the resulting unbalanced panel of 817 individuals, the average number of years available is about 9 years per individual (see table 1).

An individual is defined to teach in a particular year ($P_t = 1$) if he or she was teaching in October of that year and did not also report to be in full-time education that month.[23] Similarly a person is defined to be employed in a nonteaching job ($P_t = 2$) in a year if the person was employed in such a job in October of that year and not enrolled full time in college. Those not working in a particular year include all individuals working at home, enrolled in full-time education, or unemployed (although in our sample very few individuals reported being unemployed). No distinction is made between full-time and part-time work. Yearly earnings in each occupation are defined as 2,000 times the real (in 1982 dollars) hourly wage rate. The latter was obtained by dividing the reported weekly, monthly, or yearly earnings in the job occupied in October of that year by the reported number of hours worked in that time interval.

In addition to the information on the sample members' complete work and earnings history from the date of graduation until 1986, the analysis includes information about their educational attainment at the time of graduation, as well as a number of other individual characteristics, such as race, gender, age and state of residence. Nonwhite is defined as zero if the person is white and one otherwise. Female equals one if the individual is a female. BEd and MEd equal one if the individual has a bachelor's or master's degree in education, respectively, and equal zero if not. MA equals one if the individual has a master's degree in another subject. If the individual received a bachelor's degree in one of the sciences, Science is equal to one, and zero if not. SAT represents the individual's total SAT scores, and Manufacturing wage is the mean state manufacturing wage earnings, in thousands of 1982 dollars, averaged over the 1975–85 period. Age, $\exp_{1t}$, and $\exp_{2t}$ represent the individual's age in the first period, the individual's total teaching experience, and total years of work experience in the nonteaching sector.

The means and standard deviations of the variables are shown in table 1. Because of oversampling of various subgroups (including oversampling of school teachers), the NLS-72 sample does not constitute a nationally representative random sample of the population of all school teachers in this cohort. Therefore sample weights were applied in all estimations.

To obtain an idea of the extent of occupational mobility in the sample, table A1 shows the frequency counts of various career patterns. The table shows that only 244 individuals (30%) remained in the same labor force state throughout the sample period, 126 (15%) changed labor force status

[23] While information is available about the individual's work status in all other months as well, this information was found to be somewhat less reliable than that for the status in October. The first four follow-up surveys were all conducted in October or shortly thereafter, and individuals were asked about their status in that month specifically, reducing potential recall errors.

once (i.e., they had exactly two spells), and 185 (23%) had three spells. The remaining 262 individuals (32%) experienced more than three different spells.

Table A1
Frequencies of Observed Occupational Choice Sequences

LF Status Sequence	Number of Observations	LF Status Sequence	Number of Observations	LF Status Sequence	Number of Observations
O	6	T	146	N	92
OT	12	TO	50	NT	30
ON	18	TN	64	NO	54
OTO	5	TOT	49	NOT	5
OTN	1	TON	30	NON	43
ONO	7	TNO	8	NTO	16
ONT	4	TNT	17	NTN	0

Note.—Each letter represents a spell occurring over one or more years. LF stands for labor force status, O for out of labor force, T for teaching, and N for employment in the nonteaching sector. Observed sequences end either at the end of the sample period (1986) or in the first year in which the occupation status is unknown. The first spell starts in the first year after graduation from a teacher training program in which the individual is no longer engaged in full-time study.

References

Arcidiacono, Peter, Joe Hotz, and Songman Kang. 2010. Modeling college major choice using elicited measures of expectations and counterfactuals. Working Paper no. 15729, National Bureau of Economic Research, Cambridge, MA.

Armantier, Olivier, Wändi Bruine de Bruin, Giorgio Topa, Wilbert van der Klaauw, and Basit Zafar. 2011. Inflation expectations and behavior: Do survey respondents act on their beliefs? Staff Report no. 509, Federal Reserve Bank of New York.

Bellemare, Charles, Sabine Kröger, and Arthur van Soest. 2008. Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica* 76, no. 4:815–39.

Benitez-Silva, Hugo, and Debra Dwyer. 2005. The rationality of retirement expectations and the role of new information. *Review of Economics and Statistics* 87, no. 3:587–92.

Bernheim, B. Douglas. 1988. Social Security benefits: An empirical study of expectations and realizations. In *Issues in contemporary retirement*, ed. Edward Lazear and Rita Ricardo-Campbell, 312–48. Stanford, CA: Hoover Institution.

———. 1989. The timing of retirement: A comparison of expectations and realizations. In *The economics of aging*, ed. D. A. Wise, 335–55. Chicago: University of Chicago Press.

———. 1990. How do the elderly form expectations? An analysis of responses to new information. In *Issues in the economics of aging*, ed. D. A. Wise, 259–83. Chicago: University of Chicago Press.

Bernheim, B. Douglas, and Lawrence Levin. 1989. Social Security and personal savings: An analysis of expectations. *American Economic Review* 79, no. 2:97–102.

Blass, Asher A., Saul Lach, and Charles F. Manski. 2010. Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability. *International Economic Review* 51, no. 2:421–40.

Blau, Francine, and Marianne Ferber. 1991. Career plans and expectations of young women and men. *Journal of Human Resources* 26, no. 4:581–607.

Branch, William A. 2004. The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal* 114, no. 497:592–621.

Bruine de Bruin, Wändi, Charles F. Manski, Giorgio Topa, and Wilbert van der Klaauw. 2011. Measuring consumer uncertainty about future inflation. *Journal of Applied Econometrics* 26, no. 3:454–78.

Bruine de Bruin, Wändi, Simon Potter, Robert Rich, Giorgio Topa, and Wilbert van der Klaauw. 2010. Improving survey measures of household inflation expectations. *Current Issues in Economics and Finance* 16, no. 7. http://www.newyorkfed.org/research/current_issues/ci16-7.html.

Delavande, Adeline. 2008a. Measuring revisions to subjective expectations. *Journal of Risk and Uncertainty* 36, no. 1:43–82.

———. 2008b. Pill, patch, or shot? Subjective expectations and birth control choice. *International Economic Review* 49, no. 3:999–1042.

Delavande, Adeline, and Hans-Peter Kohler. 2011. HIV/AIDS-related expectations and risky sexual behavior in Malawi. Working Paper 7–29–2011, Populations Studies Center, University of Pennsylvania.

Delavande, Adeline, and Susann Rohwedder. 2011. Individuals' uncertainty about future Social Security benefits and portfolio choice. *Journal of Applied Econometrics* 26, no. 3:498–519.

Del Negro, Marco, and Stefano Eusepi. 2010. Fitting observed inflation expectations. Staff Report no. 476, Federal Reserve Bank of New York.

Dominitz, Jeff. 1998. Earning expectations, revisions, and realizations. *Review of Economics and Statistics* 80, no. 3:374–88.

Dominitz, Jeff, and Charles F. Manski. 1997a. Perceptions of economic insecurity: Evidence from the Survey of Economic Expectations. *Public Opinion Quarterly* 61, no. 2:261–87.

———. 1997b. Using expectations data to study subjective income expectations. *Journal of the American Statistical Association* 92, no. 439: 855–67.

———. 2005. Measuring and interpreting expectations of equity returns. Working Paper no. 11313, National Bureau of Economic Research, Cambridge, MA.

Eckstein, Zvi, and Kenneth I. Wolpin. 1989. The specification and estimation of dynamic stochastic discrete choice models: A survey. *Journal of Human Resources* 24, no. 3:562–98.

Engelberg, Joseph, Charles F. Manski, and Jared Williams. 2009. Comparing the point predictions and subjective distributions of professional forecasters. *Journal of Business and Economic Statistics* 27, no. 1:30–41.

Flinn, Christopher, and Daniela Del Boca. 1984. Self-reported reservation wages and the labor market participation decision. *Richerche Economiche* 38 (July/September): 363–83.

Galati, Gabriele, Peter Heemeijer, and Richhild Moessner. 2010. How do inflation expectations form? Evidence from a high frequency survey. DNB Working Paper no. 285, De Nederlandsche Bank.

Gotz, Glenn A., and John McCall. 1985. A dynamic retention model for Air Force officers. Report no. R-3028-AF, Rand Corporation, Santa Monica, CA.

Griliches, Zvi. 1980. Expectations, realizations, and the aging of young men. In *Research in labor economics*, ed. R. G. Ehrenberg, 3:1–21. Greenwich, CT: JAI.

Gronau, Reuben. 1988. Sex-related wage differentials and women's interrupted labor careers—the chicken or the egg? *Journal of Labor Economics* 6, no. 3:277–301.

Hamermesh, Daniel S. 1985. Expectations, life expectancy, and economic behavior. *Quarterly Journal of Economics* 100, no. 2:389–408.

Honig, Marjorie. 1996. Retirement expectations: Differences by race, ethnicity, and gender. *Gerontologist* 36, no. 3:373–82.

Horowitz, Joel L. 1992. A smoothed maximum score estimator for the binary response model. *Econometrica* 60, no. 3:505–32.

Houser, Daniel, Michael Keane, and Kevin McCabe. 2004. Behavior in a dynamic decision problem: An analysis of experimental evidence using a Bayesian type classification algorithm. *Econometrica* 72, no. 3:781–822.

Hurd, Michael D., and Kathleen McGarry. 1995. Evaluation of the subjective probabilities of survival in the Health and Retirement Study. *Journal of Human Resources* 30, suppl.: S268–S292.

———. 2002. The predictive validity of subjective probabilities of survival. *Economic Journal* 112, no. 482:966–85.

Hurd, Michael D., James P. Smith, and Julie M. Zissimopoulos. 2004. The effects of subjective survival on retirement and Social Security claiming. *Journal of Applied Econometrics* 19, no. 6:761–75.

Juster, Thomas. 1966. Consumer buying intentions and purchase probability: An experiment in survey design. *Journal of the American Statistical Association* 61 (September): 658–96.

Keane, Michael P., and Kenneth I. Wolpin. 1997. Career decisions of young men. *Journal of Political Economy* 105, no. 3:473–522.

Lancaster, Tony, and Andrew Chesher. 1983. An econometric analysis of reservation wages. *Econometrica* 51, no. 6:1661–76.

Lochner, Lance. 2007. Individual perceptions of the criminal justice system. *American Economic Review* 97, no. 1:444–60.

Magnac, Thierry, and David Thesmar. 2002. Identifying dynamic discrete decision processes. *Econometrica* 70, no. 2:801–16.

Manski, Charles F. 1975. Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics* 3, no. 3:205–28.

———. 1985. Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimator. *Journal of Econometrics* 27 (1985): 313–34.

———. 1990. The use of intentions data to predict behavior: A best case analysis. *Journal of the American Statistical Association* 85, no. 412:934–40.

———. 1993. Adolescent econometricians: How do youth infer the returns to schooling? In *Studies of supply and demand in higher education*, ed. Charles Clotfelter and Michael Rothschild, 43–60. Chicago: University of Chicago Press.

———. 2002. Identification of decision rules in experiments on simple games of proposal and response. *European Economic Review* 46, no. 4–5:880–91.

———. 2004. Measuring expectations. *Econometrica* 72, no. 5:1329–76.

Rust, John. 1991. Estimation of dynamic structural models: Problems and prospects; part 1, discrete decision processes. In *Advances in econometrics: Proceedings of the sixth world congress of the Econometric Society*, ed. Christopher Sims and Jean-Jacques Laffont. New York: Cambridge University Press.

———. 1994. Structural estimation of Markov decision processes. In *Handbook of econometrics*, vol. 4, ed. Robert F. Engle and Daniel L. McFadden, 3081–43. North-Holland: Elsevier.

———. 1996. Numerical dynamic programming in econometrics. In *Handbook of computational economics*, vol. 1, ed. Hans M. Amman, David A. Kendrick, and John Rust, 619–729. North-Holland: Elsevier.

Sandell, Steven H., and David Shapiro. 1980. Work expectations, human capital accumulation, and wages of young women. *Journal of Human Resources* 15, no. 3:335–53.

Shaw, Lois B., and David Shapiro. 1987. Women's work plans: Contrasting actual expectations and actual work experience. *Monthly Labor Review* 100, no. 11:7–14.

Stinebrickner, Todd, and Ralph Stinebrickner. 2008. Learning about academic ability and the college drop-out decision. Working paper, Department of Economics, University of Western Ontario.

van der Klaauw, Wilbert. 1996. Female labor supply and marital status decisions: A life cycle model. *Review of Economic Studies* 63, no. 2: 199–235.

van der Klaauw, Wilbert, and Kenneth I. Wolpin. 2008. Social Security and the retirement and savings behavior of low-income households. *Journal of Econometrics* 145, nos. 1–2:21–42.

Wolpin, Kenneth I. 1999. Commentary on "Analysis of choice expectations in incomplete scenarios" by C. F. Manski. *Journal of Risk and Uncertainty* 19, no. 1–3:67–69.

Wolpin, Kenneth I., and Fusun Gonul. 1985. On the use of expectations data in micro surveys: The case of retirement. Report on the Employment and Training Administration (March). US Department of Labor, Washington, DC.

Zafar, Basit. 2009. College major choice and the gender gap. Staff Report no. 364. Federal Reserve Bank of New York.

———. 2011a. Can subjective expectations data be used in choice models? Evidence on cognitive biases. *Journal of Applied Econometrics* 26, no. 3:520–44.

———. 2011b. How do college students form expectations? *Journal of Labor Economics* 29, no. 2:301–48.