Learning and Beliefs: Models, Measurement, and Identification*

Arnaud Maurel† Michela M. Tincani‡

September 2, 2026

Abstract

This article discusses the rapidly expanding literature on imperfect information, learning, and belief formation, with a focus on learning models in labor economics. We structure the review around a stylized dynamic learning framework, which we use as a lens to highlight key methodological and empirical contributions. We pay particular attention to issues of measurement and identification. We conclude by outlining promising directions for future research, stressing the importance of tight links between measurements and theory.

*Generative AI tools were used to assist with language editing.

†Duke University, NBER, CESifo, IZA and Rockwool Foundation Berlin, apm16@duke.edu.

‡University College London, IFS, Stone Centre at UCL, CEPR, CESifo, HCEO and INSPIRE, m.tincani@ucl.ac.uk.

1 Introduction

In many contexts that are relevant to labor economics, individuals make choices under imperfect information about the environment and base their decisions on their subjective beliefs about future outcomes. For example, we have evidence from the empirical literature that agents make schooling, occupation, or job search decisions under uncertainty about their own productivities and future labor market opportunities. Similarly, firms typically operate under uncertainty about worker productivity, match quality, and demand conditions. In these different contexts, agents observe some outcomes that are informative about their initially unknown characteristics, such as, e.g., wages for unobserved productivity. They may then use these signals to update their beliefs, which, as time goes by, become more accurate.

In this article, we review the rapidly growing literature on imperfect information, learning, and belief formation, with a focus on applications in labor economics. We provide a common learning framework to unify the literature and highlight the key components that may differ across models, including the information structure and expectation formation process. We then analyze different types of applications of learning models through the lens of this framework. In some cases, learning itself is a key mechanism of interest, and the models help quantify how it shapes high-stakes outcomes such as major choice, college dropout, and sorting across schooling and labor-market paths. In others, belief revisions induced by randomized information provision serve as an identification device, allowing researchers to recover structural parameters under weaker assumptions. The framework we provide helps clarify the distinct roles that learning can play in empirical research.

In the last part of the article, we outline future directions in this literature. An avenue which we think may be particularly fruitful is to build further connections between different strands of the learning literature. This includes, among others, building bridges with behavioral microeconomic theory, or, on the empirical side, leveraging AI tools to elicit new measures of beliefs, as well as building tractable methods to evaluate the sensitivity of learning models to assumptions about belief formation. Strengthening connections between belief measurement and theory is a key direction for future research.

The remainder of this article is structured as follows. In Section 2, we review recent empirical evidence on the prevalence and consequences of imperfect information and learning. Section 3 introduces a stylized dynamic learning model, which we use as a unifying lens for the literature. We then use this framework in Section 4 to discuss measurement and identification challenges, and in Section 5 to review several recent applications of learning models. Finally, Section 6 discusses promising directions for future research.

2 Empirical evidence on imperfect information and learning

We now have ample evidence that beliefs, which may not always be accurate, matter for important economic outcomes by shaping people's choices. Rather than providing an exhaustive review, we highlight a selected set of contributions, mostly recent work in labor economics and the economics of education.[1]

In the context of job search, information frictions affect both where people search and how they evaluate outside options. Belot, Kircher, and Muller (2019) show, using a randomized online advice tool in Scotland, that personalized information about occupational opportunities and skill requirements increases search effort and job interviews, especially among jobseekers who initially search narrowly and among the longer-term unemployed. Jäger, Roth, Roussille, and Schoefer (2024) provide direct evidence that workers hold biased beliefs about outside options: German workers anchor their outside-option beliefs on current wages, and information about comparable workers' wages changes search and wage-negotiation intentions.

Similar patterns arise in education and school choice.[2] In classic evidence on perceived returns, Jensen (2010) shows that eighth-grade students in the Dominican Republic underestimated the returns to education and completed more schooling after receiving information about those returns. Related evidence shows that students' beliefs about academic performance affect high school track choice in Mexico (Bobba, Frisancho, and Pariguana, 2026), and

[1]Large and growing literatures also study learning and expectations in other areas of applied microeconomics, including development economics (Delavande, 2014), industrial organization (Aguirregabiria, 2021, Norris Keiller, de Paula, and Van Reenen, 2026), health economics (Chan and Hamilton, 2006, Dupas, 2011), and public economics (Stantcheva, 2021). Learning models are also popular in other fields such as marketing (see seminal work by Erdem and Keane, 1996).

[2]For a survey of expectations in education, see Giustinelli (2023).

that beliefs about admission chances affect how applicants search for schools in Chile and the United States, with over-optimism leading to under-search and non-optimal placements (Arteaga, Kapor, Neilson, and Zimmerman, 2022).[3] Tincani, Kosse, and Miglino (2026) and Carlana, Chiuri, Miglino, and Tincani (2026) show that students' belief errors about academic performance can shape the short- and longer-run effects of college affirmative action through incentive channels and mismatch in Chile.

Crucially, beliefs are not immutable. Information experiments and panel belief data show that individuals revise beliefs when they receive new signals. In the context of job search, Conlon, Pilossoph, Wiswall, and Zafar (2018) use panel data on workers' expectations and realized search outcomes to show that job seekers update their wage-offer beliefs after offer shocks. Adams-Prassl, Boneva, Golin, and Rauh (2023) similarly show that beliefs about the returns to job-search effort respond to realized job offers. A complementary employer-learning literature shows that firms also learn from signals about workers: in the classic contribution of Altonji and Pierret (2001), wages become increasingly related to productivity measures that are initially difficult for firms to observe, consistent with employers learning about worker productivity as labor-market experience accumulates.[4] In the context of college major choice, Stinebrickner and Stinebrickner (2014b) use rich longitudinal beliefs data elicited among students enrolled in Berea College (Kentucky) to establish that students enter college with inaccurate beliefs about academic match quality and revise those beliefs as grades reveal information about their performance. Wiswall and Zafar (2015) show that students update their beliefs about earnings and academic ability after receiving information, and these revisions affect stated major-choice probabilities. Together, this evidence motivates the dynamic learning framework developed below.

[3] Moreover, belief errors about admission chances reverse the welfare assessment of the Boston and the deferred acceptance allocation mechanisms in the United States (Kapor, Neilson, and Zimmerman, 2020). At the higher education level, students' biased beliefs about their college admission chances have been shown to affect application decisions in Chile and in France (Larroucau, Rios, Fabre, and Neilson, 2026, Hakimov, Schmacker, and Terrier, 2025).

[4] Related evidence includes Pastorino (2024) and Gualdani, Pastorino, de Paula, and Salgado (2026), who study learning about worker productivity and its implications for career dynamics and labor market sorting; Kahn and Lange (2014), who use performance measures to study employer learning and wage dynamics; and Hansen, Hvidman, and Sievertsen (2024), who show that the labor-market value of grades fades as employers learn about worker productivity.

3 A stylized dynamic learning model

We outline in this section a stylized dynamic learning framework, which we will use to highlight the determinants and consequences of learning. We first consider an environment in which agents have rational expectations about their future outcomes. This constitutes a natural benchmark, which we then extend by allowing agents to have biased beliefs. While we focus on the specific context of occupational choice, a similar framework can be used to model learning in different environments, including, for instance, educational choices or health decisions.[5] This model is a simplified version of the framework in Arcidiacono, Aucejo, Maurel, and Ransom (2025), which considers a more general environment where individuals make schooling as well as labor market decisions under imperfect information about their own schooling ability and labor market productivity.

Individuals are assumed to be forward-looking and to choose the sequence of decisions yielding the highest present value of expected lifetime utility. Namely, any given individual i chooses the sequence of actions $(d_{it})_{t=1\ldots T}$ to sequentially maximize the discounted sum of payoffs:

$$E\left[\sum_{t=1}^{T}\beta^{t-1}\sum_{j=1}^{J}(u_j(Z_{it})+\nu_{ijt})1\{d_{it}=j\}\right] \tag{1}$$

where $\beta \in (0,1)$ is the discount factor. Importantly, there are two distinct sources of uncertainty in this environment. The expectation is taken with respect to i) the distribution of the future idiosyncratic preference shocks (ν_{ijt}) and of any stochastic evolution of observable state variables in Z_{it}, and ii) the distribution of signals about unknown heterogeneity components that are associated with different choice paths. The first component is standard in dynamic discrete choice models, while the second arises because agents have imperfect information about a portion of the state that is fixed over time. We focus here on a setup where agents learn about a permanent and possibly vector-valued heterogeneity component, which we denote by θ_i. In the occupational choice environment, it is natural to let θ_i denote the vector of occupation-specific unobserved productivities. When deciding whether

[5] See seminal work by Miller (1984) on worker learning in the context of occupational choice, and subsequent work by Antonovics and Golan (2012) on experimentation and learning in job choice.

to work in a particular occupation, individuals consider the option value associated with the new information acquired along different choice paths. This is in addition to other more standard determinants of occupational choice, including wages and amenities.

Let $V_t(Z_{it})$ denote the ex-ante value function at the beginning of period t, that is, the expected discounted sum of current and future payoffs just before the current period idiosyncratic shock is revealed. The choice-specific value function $v_j(Z_{it})$ is then given by:

$$v_j(Z_{it}) = u_j(Z_{it}) + \beta E_t \left[V_{t+1}(Z_{it+1})|Z_{it}, d_{it} = j\right] \tag{2}$$

where the term $E_t[\cdot|\cdot]$ is indexed by t to highlight the fact that this expectation is conditional on the information set of the individual at the beginning of period t. This information set includes the sequence of signals that the individual has received up until $t-1$.

We assume that individuals have the prior belief that θ_i is normally distributed, with mean zero and covariance matrix $\Lambda(\theta_i)$. They update their beliefs about θ_i through noisy signals that depend on their decisions: working in occupation j provides signals through wages in that occupation. We consider a correlated learning environment, in which productivity is allowed to be correlated across occupations. Signals from a given occupation k may be informative not only about the individual's productivity within that occupation (θ_{ik}), but also within any other occupation $l \neq k$ (θ_{il}).

Upon working in occupation $j \in \{1, \dots, J\}$, realized (log-)wages in that occupation are:

$$w_{ijt} = \delta_t + \gamma_{0j} + X_{ijt}\gamma_{1j} + \theta_{ij} + \varepsilon_{ijt}, \tag{3}$$

where δ_t is a time effect, γ_{0j} is an occupation-specific fixed effect, and X_{ijt} is the observed component of the state vector. The idiosyncratic shocks, ε_{ijt}, are assumed to be distributed $\mathcal{N}(0, \sigma_j^2)$ and are independent over time as well as across individuals and the J sectors, and independent of the other state variables. We also assume that θ_i is independent of the initial observed state variables X_{ij1}. The productivity signal is then given by:

$$S_{ijt} = w_{ijt} - \delta_t - \gamma_{0j} - X_{ijt}\gamma_{1j}. \tag{4}$$

Denote by $\widetilde{S}_{it}$ a $J \times 1$ vector with zeros everywhere except for the element corresponding to the choice in period t. The non-zero element of this vector corresponds to the productivity signal received in this period. We now turn to the updating rules. Let Ω_{it} be a $J \times J$ matrix with zeros everywhere except for the diagonal term corresponding to the choice made by individual i in period t. The diagonal element corresponding to this choice is given by the inverse of the variance of the choice-specific idiosyncratic shock.

It follows from the normality assumptions on the initial prior productivity distribution and on the idiosyncratic shocks that the posterior productivity distributions are also normally distributed. Specifically, denoting by $E_t(A_i)$ and $\Lambda_t(A_i)$ the posterior productivity mean and variance-covariance matrix at the end of period t, the updating rules for the mean and the variance-covariance of the posterior distribution are given by (see DeGroot, 1970):

$$E_t(A_i) = (\Lambda_{t-1}^{-1}(A_i) + \Omega_{it})^{-1}(\Lambda_{t-1}^{-1}(A_i)E_{t-1}(A_i) + \Omega_{it}\widetilde{S}_{it}) \quad (5)$$

$$\Lambda_t(A_i) = (\Lambda_{t-1}^{-1}(A_i) + \Omega_{it})^{-1}. \quad (6)$$

To build some intuition, it is useful to consider a special case with unidimensional unobserved productivity. Denoting by σ_ε^2 the variance of the idiosyncratic productivity shocks, assumed to be identical across occupations, the posterior productivity mean (m_{it}) and variance (σ_{At}^2) are in this case given by:

$$m_{it} = \left(\frac{1}{\sigma_{At-1}^2} + \frac{1}{\sigma_\varepsilon^2}\right)^{-1}\left(\frac{m_{it-1}}{\sigma_{At-1}^2} + \frac{S_{it}}{\sigma_\varepsilon^2}\right) \quad (7)$$

$$\sigma_{At}^2 = \left(\frac{1}{\sigma_{At-1}^2} + \frac{1}{\sigma_\varepsilon^2}\right)^{-1} \quad (8)$$

As is clear from these expressions, the posterior variance decreases over time towards zero as individuals receive additional productivity signals, thus giving more weight to the $t-1$ posterior productivity mean and less to the current-period signal. In this environment, workers who have accumulated many years of labor market experience will end up with arbitrarily accurate beliefs about their own productivity.

Initially known vs. unknown heterogeneity We have remained silent so far about the model objects that are observed or unobserved by the econometrician. It is reasonable to think that individuals know more about their own productivity than the econometrician. For instance, individuals who enter the labor market likely already have some idea about their own productivity, in part through signals received while in school. There may be two productivity components, both unobserved to the econometrician, the first one being initially unknown to the agent, the other one being known from the outset.

The framework presented above allows for such an environment, with the first component being given by θ_i and the second being part of the state variables Z_{it}. We return to this issue below when discussing identification, for which the co-existence of both types of unobserved heterogeneity factors will give rise to additional challenges.

Departures from rational expectations We have so far maintained the assumption that agents have rational expectations about their future outcomes. While this is a very standard assumption in economics, a growing body of empirical evidence shows that, in various contexts, agents may instead hold biased beliefs (see, e.g., Manski, 2004, D'Haultfoeuille, Gaillac, and Maurel, 2021, Crossley, Gong, Stinebrickner, and Stinebrickner, 2024, Tincani, Kosse, and Miglino, 2026). Along similar lines, individuals may also update their beliefs in ways that depart from Bayesian updating (see Ortoleva, 2024, for a recent survey). The dynamic learning setup considered above can be extended in these directions.

A natural approach is to relax rational expectations, while maintaining the assumption that the (conditional) subjective expectations operator shares the same mathematical properties as the conditional expectation operator. Standard examples of departures from rational expectations that have been considered in the literature include over-optimism (or over-pessimism) and myopic expectations. The main difference is that subjective expectations integrate over the subjective—as opposed to the true—conditional outcome distribution. Agents choose the sequence of actions that maximize the subjectively expected discounted sum of payoffs, given by:

$$\mathcal{E}\left[\sum_{t=1}^{T}\beta^{t-1}\sum_{j=1}^{J}(u_j(Z_{it})+\nu_{ijt})1\{d_{it}=j\}\right] \tag{9}$$

where $\mathcal{E}[.]$ denotes the subjective expectations operator, which may differ from the rational expectations counterpart, $E[.]$. For instance, for the case of over-optimistic individuals, one can write beliefs about occupation-specific utilities as $\mathcal{E}[u_j(Z_{it})] = E[u_j(Z_{it})] + b(Z_{it})$, where $b(Z_{it}) > 0$ denotes the bias.

Individuals may update their beliefs, based on the accumulated signals, in ways that differ from the standard Bayesian updating rule described above. For instance, individuals may overreact to signals, and place a higher weight on the signals than would be optimal given their precision. Alternatively, individuals may exhibit confirmation bias, placing a higher weight on signals that confirm their prior beliefs.

We focus for simplicity on the case of unidimensional unobserved productivity. The posterior productivity mean and variance at the end of period t can more generally be written as a function of prior productivity mean and variance and of the signals:

$$m_{it} = f_{m,i}(m_{it-1}, S_{it}, \sigma^2_{Ait-1}) \tag{10}$$

$$\sigma^2_{Ait} = f_{\sigma,i}(\sigma^2_{Ait-1}) \tag{11}$$

Note that m_{it} and σ^2_{Ait} correspond here to the mean and variance of the *subjective* beliefs distributions. In the standard rational expectations - Bayesian updating setup discussed above, the updating process $(f_{m,i}, f_{\sigma,i})$ is the same for all individuals. In general though, some individuals may put a higher weight on the signals (relative to their prior beliefs) than others, resulting in heterogeneous updating rules.

4 Econometric and measurement challenges

Identification invariably depends on the type of data that is available to the analyst. We distinguish in what follows between common situations where subjective beliefs are not elicited from the individuals (see, e.g., Stange, 2012, Arcidiacono, Aucejo, Maurel, and Ransom, 2025, Bunting, Diegert, and Maurel, 2024), and others where subjective beliefs are available to the analyst (see, e.g., Stinebrickner and Stinebrickner, 2014a, Wiswall and Zafar, 2015, Attanasio, Sancibrian, and Ambrosio, 2025).

While our discussion below focuses on measurement and identification issues, the estimation of learning models also raises its own challenges. This applies in particular to the case of dynamic learning models with large state spaces, especially when agents learn about multiple, potentially correlated attributes over time. Conditional choice probability-based estimation methods

(Arcidiacono and Ellickson, 2011) make it possible to estimate these models at a modest computational cost by avoiding repeatedly solving the agents' dynamic optimization problem within the estimation procedure.

4.1 Learning models without subjective beliefs data

Without data on agents' beliefs about their productivities, a key challenge is to tell apart the portion of unobserved heterogeneity that is known by the agents but unobserved to the econometrician, from the one that is initially unknown to the agents as well but may be gradually revealed over time. Different strategies can be used depending on the available data. Arcidiacono et al. (2025) consider an environment where noisy measurements on the known portion of unobserved heterogeneity—ASVAB ability measurements from the NLSY97 in the context of their application—are available to the researcher. With access to a sufficient number of measurements, this makes it possible to identify the distribution of the initially known heterogeneity. Intuitively, the extent to which agents learn about the initially unknown heterogeneity, and make their decisions based on it, is informed by empirical patterns such as leaving college upon getting lower grades than predicted, or returning to school upon receiving lower wages than anticipated. A simple but important insight is that a pure learning model where agents do not have private information reduces to a problem of selection on observables. Namely, conditional on observed signals (e.g. grades or wages), choices are independent of latent productivity as the agent and the econometrician share the same information set.

Noisy measurements of the initially known component of unobserved heterogeneity provide a powerful source of identifying information, but such measurements are unavailable in many empirical settings. Recent work by Bunting, Diegert, and Maurel (2024) analyzes the identification within a broad class of learning models, focusing on situations where researchers have access to a short panel on choices and realized outcomes. This work illustrates a trade-off between access to ability measurements and strength of the identifying assumptions. When such measurements are not available, identification relies on stronger restrictions. In particular, the model is identified under a compact support restriction on the initially known heterogeneity together with distributional assumptions such as normality of the initially unknown heterogeneity and of the idiosyncratic shocks.

4.2 Value and challenges of subjective beliefs data

Elicited beliefs data, when available, provide a valuable source of information for identifying and estimating learning models while relaxing some of the underlying assumptions. Stinebrickner and Stinebrickner (2014b), which we discuss in detail in Section 5.1, provides an important illustration of the value of such data. Their role also depends on which features of subjective belief distributions are elicited. For example, measuring uncertainty, as in Gong, Stinebrickner, and Stinebrickner (2019), and incorporating it into structural learning models can help identify respondent-specific updating weights and characterize heterogeneity in belief updating. But subjective beliefs data do not, by themselves, resolve all the identification challenges discussed above. Elicited beliefs are best viewed as noisy measurements of the latent beliefs entering individuals' choice problems, and using them requires explicit assumptions about measurement error.

Another measurement challenge relates to the type of outcomes about which beliefs are elicited. Various surveys collect beliefs about realized outcomes, including wages, grades, or longevity. However, individuals make decisions based on their beliefs about potential outcomes associated with alternative sequences of decisions. This requires directly eliciting beliefs about potential outcomes (see, e.g., Wiswall and Zafar, 2015, Arcidiacono, Hotz, Maurel, and Romano, 2020, for college major and occupational choice; Cunha, Elo, and Culhane, 2022, Tincani, Kosse, and Miglino, 2026, for skill formation; and Andrew and Adams, 2025, for the marriage-market returns to education), or imposing additional structure to recover these beliefs based on the available beliefs data.

Beyond their role in identification and estimation, expectations data can also help as diagnostic tools for structurally estimated models.[6] Seminal work by Wolpin and Gonul (1985) proposes comparing stated (probabilistic) choices to the choice probabilities implied by a fully specified model, which reflect assumptions about both preferences and beliefs. Discrepancies can indicate that expectations data contain information about choice determinants not captured in the model, such as belief biases when the model imposes rational expectations. In a more recent application of a similar idea, Delavande and Zafar (2019) estimate a university-choice model under students' existing financial constraints and show that it closely predicts sep-

[6] See Kosar and O'Dea (2023) for a survey of the literature combining beliefs data with structural models.

arately elicited stated choices when financial constraints are lifted, providing an out-of-sample-style validation of the estimated model. While applications of elicited beliefs data for this purpose remain scarce, this is an important avenue for future research involving structural models, particularly dynamic learning models that require additional assumptions about belief formation and updating.

5 Applications

5.1 Learning about academic match quality

Stinebrickner and Stinebrickner (2014b) study how college students learn about their academic match quality and how this learning affects major choices and college dropout. Students enter college with imperfect information about their future performance in different fields. Over time they observe signals about their academic performance through grades and update their beliefs accordingly. These updated beliefs in turn affect decisions such as whether to remain in a field, switch majors, or leave college. The paper combines administrative transcript data with panel survey data on students' subjective beliefs and does not impose rational expectations.

The environment in Stinebrickner and Stinebrickner (2014b) can be mapped to our stylized occupational choice framework. Let $j \in \mathcal{J}$ denote the set of possible choices at decision date $t^* > 1$ (e.g., science major, non-science major, or dropout). Utility from each major depends, among other factors, on academic performance in that major, θ_{ij}. At time 1 students do not know the value of θ_{ij} they will experience at time t^* and must therefore form beliefs about it. The authors use a panel dataset of beliefs, with survey measures of expected academic performance elicited at times 1 and t^*, and administrative data on grade signals, to empirically estimate the belief updating process. They then estimate a static representation of a dynamic learning problem in which the major and dropout decisions occur at date t^*. At college entry ($t = 1$), students report their subjective probability of each future outcome at t^* and their expected future grade signals up to t^*. Using the estimated belief updating process, the authors estimate the utility-function parameters by maximum likelihood. The likelihood is constructed by matching students' stated choice probabilities to model-implied ones.

In one of the empirical specifications, the signal vector S_i is multidimensional. It includes GPA by subject, allowing grades in one field to affect beliefs about expected performance in another. The updating equation can be written as

$$\mathcal{E}_{t^*}(\theta_{ij}) = \alpha_j + \rho_j \mathcal{E}_1(\theta_{ij}) \cdot f(CL_{ij}) + \sum_k \gamma_{jk} GPA_{ik},$$

where $\mathcal{E}_{t^*}(\theta_{ij})$ denotes the posterior belief, $\mathcal{E}_1(\theta_{ij})$ the prior belief, $f(CL_{ij})$ a function of the number of classes the student took in field j, and GPA_{ik} the grade signal in subject k. The updating specification is motivated by a Bayesian learning framework in which posterior beliefs depend on the prior and the signal, with the weight on the prior depending on signal informativeness, proxied by course-taking intensity $f(CL_{ij})$. The cross-effects across disciplines are consistent with a correlated learning environment. The coefficients γ_{jk}, however, are estimated directly from the data rather than derived from an explicit model of correlated latent productivities.

The paper finds that grades in science courses substantially affect beliefs about performance in science majors, while beliefs about performance in non-science majors respond more to grades from a broader set of subjects. This provides a novel policy insight. If students begin college with overly optimistic beliefs, greater exposure to science coursework may accelerate learning about comparative advantage and ultimately reduce the number of science graduates. Policies designed to increase science participation by encouraging students to take more science courses may therefore have unintended effects. At the same time, such policies may improve sorting according to comparative advantage, potentially leading to a more efficient allocation of students across majors.

A further application to academic match connects directly to the option-value component of our stylized framework. Stange (2012) quantifies this channel in the context of college enrollment: enrollment generates grade signals about academic aptitude, allowing students to condition later continuation or dropout decisions on what they learn. Using this dynamic learning framework, the paper estimates that the option to revise schooling decisions after learning accounts for 14% of the total value of the opportunity to enroll in college.

5.2 Learning, schooling and occupational choices

Arcidiacono, Aucejo, Maurel, and Ransom (2025) investigate the role of imperfect information and learning in schooling and occupational choices. An important distinction with the papers discussed in the previous section is that they explore the extent to which transitions between college and the workplace are driven by learning about schooling ability and labor market productivity.

The framework in that paper nests the stylized dynamic learning model discussed in Section 3. Individuals learn from their grades while enrolled in college, wages while in the workforce, and possibly both if holding a job while in college. Relative to the stylized model, the dynamic framework in this paper is extended to allow for search frictions: in the white-collar sector job offers arrive stochastically.

Estimation results yield two main findings. First, imperfect information and learning play an important role. A sizeable share of the observed variation in grades and wages is due to differences in initially unknown ability. Second, interesting correlated-learning patterns emerge.[7] Grades for a given college type and major are informative about abilities associated with the other college options; wages in a given sector (blue- or white-collar) are informative about productivities in the other. In contrast, grades reveal little new information about future labor market performance. An important implication is that college acts as an effective sorting mechanism in terms of academic ability, but less so in terms of labor market productivity.

With the model estimates in hand, Arcidiacono et al. (2025) consider counterfactual environments in which individuals have perfect information *ex ante* about their abilities and productivity. Four-year college graduation would increase, primarily because fewer students would drop out once enrolled. Information provision also generates strong composition effects, with enhanced sorting into college based on comparative advantage in labor market productivity, leading to a large increase in the college wage premium. The paper also documents distributional consequences as information provision weakens the relationship between family income and college graduation.

5.3 Learning as an identification device

Wiswall and Zafar (2015) study the determinants of college major choice using an information experiment and panel data on subjective beliefs. While

[7] See also related work by James (2011) for evidence of correlated learning in the context of occupational choice.

still enrolled, students first report their beliefs about their future earnings and their academic ability in each major, and their probability of graduating in each major. They then receive accurate information about population earnings and employment outcomes by major and report these beliefs again. The paper uses the resulting pre/post belief variation to identify a structural model of major choice. In this application, learning serves as an identification device: the experiment induces belief revisions that allow the authors to separate the effects of expected earnings and academic ability from unobserved tastes in major choice. This illustrates another valuable role of learning, especially when combined with belief elicitation.

The environment in Wiswall and Zafar (2015) also maps naturally into our stylized choice framework. Let $j \in \mathcal{J}$ denote the set of possible college major choices (including dropout as an alternative). Students form beliefs about future major-specific payoffs, including expected earnings, perceived academic ability, and other post-graduation outcomes. In the language of our model, these beliefs enter the subjective expected payoff $\mathcal{E}[u_j(Z_{it})]$ and reflect uncertainty about payoff-relevant heterogeneity, θ_{ij}. By contrast, major-specific tastes are treated as *fixed* latent components known by the student but unobserved by the econometrician. The information treatment can therefore be interpreted as an exogenous signal that shifts beliefs about future payoffs, moving subjective major-choice probabilities while leaving the underlying taste component unchanged.

This application showcases a different use of learning. In Stinebrickner and Stinebrickner (2014b), learning is itself an economic mechanism that shapes later choices. In Wiswall and Zafar (2015), by contrast, learning-induced belief changes mainly serve to recover model primitives. The contribution of measuring learning therefore concerns identification: observing how subjective beliefs respond to exogenous information allows the researcher to separate beliefs from preferences and identify the parameters of a richer structural model under weaker assumptions. Students revise their beliefs about major-specific future payoffs in logical ways after receiving information, and these belief revisions are linked to changes in their stated likelihood of graduating in each major. Once tastes are separated from beliefs, a key finding is that the elasticity of major choice with respect to expected earnings is much smaller than cross-sectional OLS estimates would imply, because unobserved tastes are the dominant determinant of major choice.

6 Future directions

Over the past two decades, the labor economics literature has made impressive advances in measuring and modeling imperfect information, learning, and beliefs, and incorporating them into choice models for policy analysis. In these concluding remarks, we discuss avenues for future research that we find especially promising.

6.1 Measurements and econometrics

In the stylized framework, individuals are allowed to hold imperfect information about their own labor market productivity, but they know the causal structure of the choice problem: how actions $(d_{it})_{t=1,..T}$ and the latent heterogeneity component θ_i affect (log-)wages in each occupation. An ambitious but natural next step is to move from inferring beliefs within a prespecified model to inferring beliefs about the models themselves: which variables individuals consider relevant, which causal links they perceive, and how strong they believe those links to be. While research on subjective causal structures is expanding rapidly, real-world labor economics applications remain scarce (Ambuehl, Bhui, and Thysen, 2026). Open-ended surveys are a promising way to elicit mental models in a non-primed way, and recent developments in AI permit collecting unstructured data at scale (Haaland, Roth, Stantcheva, and Wohlfart, 2025). While such data pose new econometric challenges (see, e.g., Christensen and Hansen, 2026), eliciting causal structures could allow researchers to relax the assumption that individuals optimize within a common, researcher-imposed model.

Another important avenue is to develop tractable methods for assessing the sensitivity of structural learning models to assumptions about belief formation and updating. A natural starting point is to build on the recent econometrics literature on sensitivity analysis (Bonhomme and Weidner, 2022, Christensen and Connault, 2023), taking into account the fact that some belief formation and updating processes may be ruled out given available elicited beliefs data. Since subjective beliefs are often not jointly observed with choices and outcomes of interest, this would also require leveraging some of the recent econometrics literature on data combination (D'Haultfoeuille, Gaillac, and Maurel, 2021, 2025, Meango, Henry, and Mourifie, 2025). A natural approach would be to focus on policy-relevant counterfactuals and derive bounds on counterfactual outcomes (see, e.g., Kalouptsidi, Kitamura, Lima, and Souza-Rodrigues, 2026) that are robust to the specific type of belief formation process.

6.2 Models

In our model, signals automatically enter the belief-updating process. In practice, individuals may endogenously acquire payoff-relevant information, and may do so in heterogeneous ways (Fuster, Perez-Truglia, Wiederholt, and Zafar, 2022). The theoretical framework could therefore be extended to model information acquisition explicitly, potentially drawing on microeconomic and behavioral theories of motivated reasoning and rational inattention (Bénabou and Tirole, 2002, 2016, Mackowiak, Matejka, and Wiederholt, 2023). Linking repeated belief surveys to newly available platform-use data that record search and information acquisition (see, e.g., Arteaga, Kapor, Neilson, and Zimmerman, 2022) may help identify and estimate such learning processes.

Individuals may learn not only about outcomes but also about causal structures. The literature on behavioral causal inference offers frameworks for modeling how agents adopt and revise perceived causal structures (Spiegler, 2020). Andre, Haaland, Roth, Wiederholt, and Wohlfart (2026) take an early step toward quantification by incorporating perceived causal structures, or narratives, into an otherwise standard New Keynesian macroeconomic model and showing through simulations that they can have sizable effects on inflation and output. Adapting these approaches to labor economics would allow researchers to quantify how beliefs about deeper objects—not only outcomes, but the model linking actions to outcomes—shape high-stakes choices and identify the mechanisms through which they do so.

A final avenue that we view as challenging but especially valuable is to incorporate peer influence into structural models of learning. Experimental evidence shows that individuals learn about unknown states of the world from peers' information and actions (Bursztyn, Ederer, Ferman, and Yuchtman, 2014, BenYishay and Mobarak, 2019, Chandrasekhar, Larreguy, and Xandri, 2020). Social interactions are also relevant for broader beliefs, such as stereotypes and social norms, that shape the environment in which economic decisions are made (Corno, La Ferrara, and Burns, 2022, Meluzzi, 2024). A promising next step is to incorporate these channels into estimable structural models, building on existing Bayesian and non-Bayesian approaches to learning in networks (Mobius and Rosenblat, 2014, Golub and Sadler, 2016). Peer signals could enter agents' information sets with weights that depend on peer identity or network position. Combining these models with repeated belief measurements and data on choices would allow researchers to quantify how peers shape learning and, in turn, high-stakes education and labor-market decisions.

References

Adams-Prassl, Abi, Teodora Boneva, Marta Golin, and Christopher Rauh. 2023. "Perceived returns to job search." *Labour Economics* 80:102307.

Aguirregabiria, Victor. 2021. "Identification of firms' beliefs in structural models of market competition." *Canadian Journal of Economics/Revue canadienne d'économique* 54 (1):5–33.

Altonji, Joseph G and Charles R Pierret. 2001. "Employer learning and statistical discrimination." *The Quarterly Journal of Economics* 116 (1):313–350.

Ambuehl, Sandro, Rahul Bhui, and Heidi C Thysen. 2026. "Mental models of causal structure in economics and psychology." *arXiv preprint arXiv:2603.29070* .

Andre, Peter, Ingar Haaland, Christopher Roth, Mirko Wiederholt, and Johannes Wohlfart. 2026. "Narratives about the Macroeconomy." *Review of Economic Studies* Forthcoming.

Andrew, Alison and Abi Adams. 2025. "Revealed beliefs and the marriage market return to education." *The Quarterly Journal of Economics* 140 (3):2107–2162.

Antonovics, Kate and Limor Golan. 2012. "Experimentation and job choice." *Journal of Labor Economics* 30 (2):333–366.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom. 2025. "College attrition and the dynamics of information revelation." *Journal of Political Economy* 133 (1):53–110.

Arcidiacono, Peter and Paul Ellickson. 2011. "Practical methods for estimation of dynamic discrete choice models." *Annual Review of Economics* 3 (1):363–394.

Arcidiacono, Peter, V. Joseph Hotz, Arnaud Maurel, and Teresa Romano. 2020. "Ex ante returns and occupational choice." *Journal of Political Economy* 128 (12):4475–4522.

Arteaga, Felipe, Adam J Kapor, Christopher A Neilson, and Seth D Zimmerman. 2022. "Smart matching platforms and heterogeneous beliefs in centralized school choice." *The Quarterly Journal of Economics* 137 (3):1791–1848.

Attanasio, Orazio, Victor Sancibrian, and Federica Ambrosio. 2025. "New measures for richer theories: Some thoughts and an example." *Journal of Political Economy Microeconomics* Forthcoming.

Belot, Michele, Philipp Kircher, and Paul Muller. 2019. "Providing advice to jobseekers at low cost: An experimental study on online advice." *The Review of Economic Studies* 86 (4):1411–1447.

Bénabou, Roland and Jean Tirole. 2002. "Self-confidence and personal motivation." *The Quarterly Journal of Economics* 117 (3):871–915.

———. 2016. "Mindful economics: The production, consumption, and value of beliefs." *Journal of Economic Perspectives* 30 (3):141–164.

BenYishay, Ariel and A Mushfiq Mobarak. 2019. "Social learning and incentives for experimentation and communication." *The Review of Economic Studies* 86 (3):976–1009.

Bobba, Matteo, Veronica Frisancho, and Marco Pariguana. 2026. "Perceived ability and school choices." *Journal of Political Economy Microeconomics* Forthcoming.

Bonhomme, Stephane and Martin Weidner. 2022. "Minimizing sensitivity to model misspecification." *Quantitative Economics* 13 (3):907–954.

Bunting, Jackson, Paul Diegert, and Arnaud Maurel. 2024. "Heterogeneity, uncertainty and learning: Semiparametric identification and estimation." NBER Working Papers 32164, National Bureau of Economic Research.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman. 2014. "Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions." *Econometrica* 82 (4):1273–1301.

Carlana, Michela, Sara Chiuri, Enrico Miglino, and Michela M Tincani. 2026. "How far can inclusion go? The long-term impacts of preferential college admissions." NBER Working Papers 32525, National Bureau of Economic Research.

Chan, Tat Y and Barton H Hamilton. 2006. "Learning, private information, and the economic evaluation of randomized experiments." *Journal of Political Economy* 114 (6):997–1040.

Chandrasekhar, Arun G, Horacio Larreguy, and Juan Pablo Xandri. 2020. "Testing models of social learning on networks: Evidence from two experiments." *Econometrica* 88 (1):1–32.

Christensen, Timothy and Benjamin Connault. 2023. "Counterfactual sensitivity and robustness." *Econometrica* 91 (1):263–298.

Christensen, Timothy and Stephen Hansen. 2026. "Performing Valid Inference with AI/ML-Generated Covariates: A Guide for Empirical Practice." In *AEA Papers and Proceedings*, vol. 116. American Economic Association, 92–97.

Conlon, John J, Laura Pilossoph, Matthew Wiswall, and Basit Zafar. 2018. "Labor market search with imperfect information and learning." NBER Working Papers 24988, National Bureau of Economic Research.

Corno, Lucia, Eliana La Ferrara, and Justine Burns. 2022. "Interaction, stereotypes, and performance: Evidence from South Africa." *American Economic Review* 112 (12):3848–3875.

Crossley, Thomas F., Yifan Gong, Ralph Stinebrickner, and Todd Stinebrickner. 2024. "Examining income expectations in the college and early post-college periods: New distributional tests of rational expectations." *Journal of the European Economic Association* 22 (6):2700–2747.

Cunha, Flávio, Irma Elo, and Jennifer Culhane. 2022. "Maternal subjective expectations about the technology of skill formation predict investments in children one year later." *Journal of Econometrics* 231 (1):3–32.

DeGroot, Morris. 1970. *Optimal Statistical Decisions*. New York: McGraw Hill.

Delavande, Adeline. 2014. "Probabilistic expectations in developing countries." *Annu. Rev. Econ.* 6 (1):1–20.

Delavande, Adeline and Basit Zafar. 2019. "University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints." *Journal of Political Economy* 127 (5):2343–2393.

D'Haultfoeuille, Xavier, Christophe Gaillac, and Arnaud Maurel. 2021. "Rationalizing rational expectations: Characterizations and tests." *Quantitative Economics* 12 (3):817–842.

———. 2025. "Partially linear models under data combination." *Review of Economic Studies* 92 (1):238–267.

Dupas, Pascaline. 2011. "Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya." *American Economic Journal: Applied Economics* 3 (1):1–34.

Erdem, Tulin and Michael Keane. 1996. "Decision-making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." *Marketing Science* 15 (1):1–20.

Fuster, Andreas, Ricardo Perez-Truglia, Mirko Wiederholt, and Basit Zafar. 2022. "Expectations with endogenous information acquisition: An experimental investigation." *Review of Economics and Statistics* 104 (5):1059–1078.

Giustinelli, Pamela. 2023. "Expectations in education." *Handbook of economic expectations* :193–224.

Golub, Benjamin and Evan Sadler. 2016. "Learning in Social Networks." In *The Oxford Handbook of the Economics of Networks*, edited by Yann Bramoullé, Andrea Galeotti, and Brian W. Rogers. Oxford University Press, 504–542. URL https://doi.org/10.1093/oxfordhb/9780199948277.013.12.

Gong, Yifan, Todd Stinebrickner, and Ralph Stinebrickner. 2019. "Uncertainty about future income: Initial beliefs and resolution during college." *Quantitative Economics* 10 (2):607–641.

Gualdani, Cristina, Elena Pastorino, Aureo de Paula, and Sergio Salgado. 2026. "Identifying Uncertainty, Learning about Productivity, and Human Capital Acquisition: A Reassessment of Labor Market Sorting and Firm Monopsony Power." NBER Working Papers 34973, National Bureau of Economic Research.

Haaland, Ingar, Christopher Roth, Stefanie Stantcheva, and Johannes Wohlfart. 2025. "Understanding economic behavior using open-ended survey data." *Journal of Economic Literature* 63 (4):1244–1280.

Hakimov, Rustamdjan, Renke Schmacker, and Camille Terrier. 2025. "Confidence and college applications: Evidence from a randomized intervention." Tech. rep., WZB Discussion Paper.

Hansen, Anne Toft, Ulrik Hvidman, and Hans Henrik Sievertsen. 2024. "Grades and employer learning." *Journal of Labor Economics* 42 (3):659–682.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer. 2024. "Worker beliefs about outside options." *The Quarterly Journal of Economics* 139 (3):1505–1556.

James, Jonathan. 2011. "Ability matching and occupational choice." Working paper, FRB of Cleveland.

Jensen, Robert. 2010. "The (perceived) returns to education and the demand for schooling." *The Quarterly Journal of Economics* 125 (2):515–548.

Kahn, Lisa B and Fabian Lange. 2014. "Employer learning, productivity, and the earnings distribution: Evidence from performance measures." *The Review of Economic Studies* 81 (4):1575–1613.

Kalouptsidi, Myrto, Yuichi Kitamura, Luca Lima, and Eduardo Souza-Rodrigues. 2026. "Counterfactual Analysis for Structural Dynamic Discrete Choice Models." *Review of Economic Studies* Forthcoming.

Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman. 2020. "Heterogeneous beliefs and school choice mechanisms." *American Economic Review* 110 (5):1274–1315.

Kosar, Gizem and Cormac O'Dea. 2023. "Expectations data in structural microeconomic models." *Handbook of Economic Expectations* :647–675.

Larroucau, Tomás, Ignacio Rios, Anaïs Fabre, and Christopher Neilson. 2026. "College application mistakes and the design of information policies at scale." *Journal of Political Economy* Conditionally accepted.

Mackowiak, Bartosz, Filip Matejka, and Mirko Wiederholt. 2023. "Rational inattention: A review." *Journal of Economic Literature* 61 (1):226–273.

Manski, Charles F. 2004. "Measuring expectations." *Econometrica* 72 (5):1329–76.

Meango, Romuald, Marc Henry, and Ismael Mourifie. 2025. "Combining stated and revealed preferences." *arXiv preprint arXiv:2507.13552* .

Meluzzi, Federica. 2024. "The College Melting Pot: Peers, Culture and Women's Job Search." *arXiv preprint arXiv:2409.20225* .

Miller, Robert A. 1984. "Job matching and occupational choice." *Journal of Political Economy* 92 (6):1086–1120.

Mobius, Markus and Tanya Rosenblat. 2014. "Social learning in economics." *Annual Review of Economics* 6 (1):827–847.

Norris Keiller, Agnes, Áureo de Paula, and John Van Reenen. 2026. "Estimating Production Functions Using Subjective Expectations Data." .

Ortoleva, Pietro. 2024. "Alternatives to Bayesian Updating." *Annual Review of Economics* 16:545–70.

Pastorino, Elena. 2024. "Careers in firms: The role of learning about ability and human capital acquisition." *Journal of Political Economy* 132 (6):1994–2073.

Spiegler, Ran. 2020. "Behavioral implications of causal misperceptions." *Annual Review of Economics* 12 (1):81–106.

Stange, Kevin M. 2012. "An empirical investigation of the option value of college enrollment." *American Economic Journal: Applied Economics* 4 (1):49–84.

Stantcheva, Stefanie. 2021. "Understanding tax policy: How do people reason?" *The Quarterly Journal of Economics* 136 (4):2309–2369.

Stinebrickner, Ralph and Todd Stinebrickner. 2014a. "Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model." *Journal of Labor Economics* 32 (3):601–644.

Stinebrickner, Ralph and Todd R Stinebrickner. 2014b. "A major in science? Initial beliefs and final outcomes for college major and dropout." *Review of Economic Studies* 81 (1):426–472.

Tincani, Michela M, Fabian Kosse, and Enrico Miglino. 2026. "Beliefs and the incentive effects of preferential college admissions: Evidence from an experiment and a structural model." *Journal of Political Economy* Conditionally accepted.

Wiswall, Matthew and Basit Zafar. 2015. "Determinants of college major choice: Identification using an information experiment." *The Review of Economic Studies* 82 (2):791–824.

Wolpin, Kenneth I and Fusun Gonul. 1985. "On the Use of Expectations Data in Micro Surveys: The Case of Retirement." Working paper, ERIC.