Research Statement

Arnaud Maurel

Duke University

April 2019

1 Overview

My research and teaching interests are in labor economics, the economics of education and microeconometrics. Most of my non-methodological work lies at the intersection between the economics of education and labor economics, with a primary focus on post-secondary education demand and occupational choices. Some of my research is also concerned with the production of skills and how these skills are valued in the labor market, as well as how agents respond to changes over time in the labor market environment.

On the methodological side, my research is primarily concerned with the identification and estimation of selection and treatment effect models. I have also recently been exploring the empirical content of belief formation models, developing new tools to assess the sensitivity of behavioral models to departures from the rational expectations hypothesis.

These two lines of research are tightly connected. My work in microeconometrics is motivated by, and generally applied to questions in labor economics and the economics of education. Many of the important questions in labor economics pose fundamental econometric challenges, and much of my research in econometrics is geared toward making progress on some of these important issues.

Another key unifying theme lies in the fact that much of my work is structural. I am indeed particularly interested in modeling and identifying the mechanisms through which individuals make important economic decisions, such as enrolling in college or choosing a

particular occupation. In some of my econometric work, I explore how certain specificities of the decision process can in turn be used to recover some outcomes of interest, in particular in situations where the outcomes are only observed for a subset of individuals who made particular decisions (such as wages for labor market participants).

In the remainder of this statement, I discuss my applied microeconomics research in the areas of post-secondary education demand and occupational choices (Section 2), skill production and returns to skills (Section 3), individual responses to non-stationarity in labor market environment (Section 4), and my econometric research in the areas of identification and estimation of selection and treatment effect models (Section 5), and empirical content of beliefs formation models (Section 6).

2 Post-Secondary Education Demand and Occupational Choices

Understanding how individuals make their schooling decisions at the higher education level was a key focus of my dissertation. In Beffy, Fougère & Maurel (2012), we investigate the relative importance of expected earnings and non-pecuniary factors in the context of sorting across college majors. We consider the determinants of the choice of the major when the length of studies as well as future earnings are uncertain to the individuals at the time of the choice. We estimate on French data a sequential schooling decision model, where we use variation across the business cycle in the wage returns to the different majors to identify the impact of expected earnings on college major choice. We find a small, though statistically significant, positive elasticity of major choice to expected earnings. Our results provide evidence that the choice of college major is mainly driven by non-pecuniary factors. From a policy viewpoint, this research indicates that financial incentives, such as the implementation of major-specific tuition fees, are unlikely to be effective in shifting the skill composition of the economy. Much of the solution appears to lie instead in the formation of skills as well as preferences for different majors.

Several of the research questions I have investigated since then are rooted in my dissertation work. In Arcidiacono, Hotz, Maurel & Romano (2019) we examine the role played by expected earnings and preferences in the subsequent choice of occupation in the labor mar-

ket. Telling apart the role played by monetary incentives versus preferences in the context of sorting across sectors is a challenging task. Doing so generally requires imposing a number of restrictions, such as the existence of exclusion restrictions between monetary returns and preferences (i.e., some characteristics that affect sorting only through preferences), together with assumptions on what individuals know at the time of the decision. In this paper, we use subjective beliefs elicited from a sample of male undergraduates at Duke to relax these assumptions. Remarkably, beliefs college students hold about their future occupations and earnings turn out to be highly predictive of their future labor market choices and earnings some seven years later, even after controlling for their college major. We offer a new perspective on the determinants of sorting across occupations by using the measures of expected earnings associated with all occupations that we elicited from our sample to directly estimate how much income students expect to give up as a result of not choosing the highest paying occupation. These *ex ante* measures of willingness-to-pay provide evidence on the role played by other, in particular non-monetary, factors in the choice of occupation. We find large differences in expected earnings across occupations, and substantial heterogeneity across individuals in the corresponding expected returns. However, while sorting across occupations is partly driven by monetary returns, non-monetary factors do play a key role in this decision.[1]

In my dissertation, in particular in the chapter that lead to Beffy, Fougère & Maurel (2012), I analyzed the determinants of post-secondary schooling choices without accounting for the fact that students may in fact be uncertain and learn about their own abilities throughout the choice process. However, the remarkable fluidity of the transitions between college and the workforce, at least in the U.S, prompted me to examine the role played by imperfect information and learning in this context. Specifically, in Arcidiacono, Aucejo, Maurel & Ransom (2019), my co-authors and I investigate the role that information about one's abilities - both in college and in the labor market - plays in students' higher education and career decisions. A striking fact that has been documented in the literature on higher education is that, among students who enroll in college, the share of those who obtain a Bachelor's degree is lower than 50%, and is even lower today than in the 1970's (Bound, Lovenheim &

---

[1]Nonetheless, it is worth noting that the earnings elasticities are much more sizable than those obtained in my earlier work (Beffy, Fougère & Maurel 2012) as well as in several other papers in the context of college major choice.

Turner 2010). This holds despite large returns to receiving a four-year college degree. One of the potential reasons for the large prevailing dropout rates is learning about one's abilities to succeed in college. In this paper, we explore the extent to which the dynamic process of information acquisition accounts for the observed transitions between college and work, as well as between STEM and non-STEM majors. To do so, we estimate, using data from the National Longitudinal Survey of Youth 1997 (NLSY97), a dynamic model of schooling and work decisions, where students who graduated from high school decide at each period whether to attend college, either in a two- or four-year institution, work part-time or full-time, or engage in home production. A central novelty of our approach lies in the fact that we account for learning over time through college grades and wages, implying that some individuals may decide to leave or re-enter college as they update their beliefs about their schooling ability and work productivity. Our empirical framework allows us to evaluate the importance of learned ability in the decision to leave college, but also how these learned abilities translate into future labor market success. Our results show that the ability components which are revealed over time do account for a sizable share of the observed dispersion in grades and wages. College exit and re-entry decisions are significantly affected by ability learning. However, estimation results show that what students learn about their abilities while in college - beyond what they knew at the time of enrollment - is in fact only weakly correlated with their abilities to succeed in the workforce.[2]

Since I started conducting research on the topic of college major choice, the empirical literature on the causes and consequences of the choice of field of study in post-secondary education has been growing at a fast pace. In Altonji, Arcidiacono & Maurel (2016), my co-authors and I analyze the literature about the determinants as well as the wage effects of the choice of field of study in college and graduate school.[3] In particular, we develop in this work a dynamic model of schooling decisions, in which forward-looking individuals choose to enroll

---

[2]The paper is currently under a soon-to-be-completed revision. The revised paper extends the original paper in several significant ways. In particular, we now i) account for the possibility that students may be credit constrained, ii) extend the model with occupational choice (blue- versus white-collar sector), where white-collar jobs are not always available to job seekers, and iii) formally establish model identification.

[3]I have also written an overview of the literature on post-secondary schooling choices, and the underlying policy implications, for the journal *Regards Croisés sur l'Economie* (Maurel 2015). The main goal of this journal is to expose the general public to academic and policy-relevant research in economics.

in particular fields of study, taking into account the various - monetary and non-monetary - tradeoffs associated with particular schooling options. This model, which encompasses several of the features considered in Beffy, Fougère & Maurel (2012), Arcidiacono, Hotz, Maurel & Romano (2019) and Arcidiacono, Aucejo, Maurel & Ransom (2019), provides a unified framework that can be used to think about the challenges that empirical researchers face when attempting to estimate the causal effects of schooling choices on future labor market outcomes.

Whether, and the extent to which credit constraints matter is arguably one of the most important and challenging questions in the higher education literature. Although I have been interested in this question since the beginning of my Ph.D., I addressed it fairly recently as I obtained access to a dataset that makes it possible to overcome some of the underlying identification issues. In Belzil, Maurel & Sidibe (2018), we contribute to this question by using data from a unique field experiment on the financial barriers to higher education. The experiment was conducted in several high schools throughout Canada, where students had to make a sequence of choices between cash payments, and college loans (at the prevailing market rate), as well as grants of various amounts. In this paper, we use these choices to recover the willingness-to-pay for financial aid packages. To the extent that prospective college students who face liquidity constraints are likely to attach a significant value to the opportunity of receiving a loan at the market rate, while those who do not perceive financial constraints should regard those opportunities as redundant, we can then use the willingness-to-pay for loans to test for the existence of financial constraints in this context. Using the experimental data to identify and estimate a simple model of financial aid acceptance decisions, we find that a large share of students are affected by credit constraints. In particular, the median high school student in our sample would be willing to pay a large (6.6 percentage points) interest rate premium to secure a $1,000 loan.[4]

The papers that I have discussed so far open up a number of important research avenues, and some of my current research is focused on making progress on these questions. For instance, in Beffy, Fougère & Maurel (2012) and Arcidiacono, Aucejo, Maurel & Ransom (2019), we examine how individuals make their schooling decisions, in particular which field

---

[4]Tuition fees are, on average, substantially lower in Canada than in the U.S., so that one might expect even higher value to accessing college loans in the U.S.

of study to pursue, in part as a function of the expected labor market outcomes associated with each alternative. In Arcidiacono, Hotz, Maurel & Romano (2019), we focus on sorting across occupations, and on the importance of earnings beliefs versus other - in particular non-pecuniary - factors in the allocation of workers across occupations. However, much remains to be known about the interplay between college major and occupational choices. I expect that this is going to be an important part of my research agenda over the next few years. In particular, this motivated a research project (Maurel & Shephard 2019), in which we incorporate schooling choice, both in terms of number of years of schooling and college major, within an empirical equilibrium search model. In this framework, students make their schooling decisions by taking into account the fact that the labor market is frictional. For instance, a STEM college graduate may end up only receiving a job offer in a non-STEM occupation, resulting in a form of mismatch between her skill set and occupation. Firms may in turn adjust their demand for different types of skills to the skill composition of the workforce. In this environment, wage returns to college majors evolve endogenously with the composition and prevalence of each field of study. We next plan on using this empirical framework to evaluate the labor market consequences of various policies aimed at increasing the number of college students graduating in STEM.

In many countries, students are allocated across different colleges and programs based on a centralized admission system. While these centralized mechanisms are common in both developing and developed countries, much remains to be known about the properties - in terms of efficiency, as well as equity - of these mechanisms as implemented in practice. Some of my future research will examine this question, with a focus on the role of information and students' beliefs in the allocation process. In a project with Olivier De Groote and Margaux Luflade (De Groote, Luflade & Maurel 2019), we plan on investigating this question in the context of the French centralized higher education admission system. Of particular interest in this environment is the fact that there has been multiple changes over time in the process through which students are allocated across different college options. Notably, we intend to leverage variations in the information that was provided to the applicants about their admission prospects to the different programs to evaluate the impact of expected admission chances on application behavior of prospective students. We will then also exploit these variations to estimate a model of higher education demand and centralized college admission,

and use the model to evaluate counterfactual admission policies.[5] We recently obtained access to the application and admission data from the French Ministry of Higher Education for the year 2016, and are currently in the process of obtaining access for the years 2012 to 2017.

3   Skill Production and Returns to Skills

Another line of my research is concerned with the production of skills, and how those skills are valued on the labor market. In Buscha, Maurel, Page & Speckesser (2012), we investigate the effect of working while in high school on schooling performance, as measured by test scores in mathematics and reading by the end of grade 12. Most of the previous papers in the literature have attempted to deal with the fact that students who hold a job while in school tend to have different observed and unobserved characteristics than those who do not work by using an instrumental variable approach. However, credible instruments that only affect test scores through high school employment are particularly difficult to find in this context. In this paper we follow another path, and attempt to identify the effect of interest by taking advantage of the panel dimension of the National Education Longitudinal Study of 1988 together with the availability of a rich set of covariates. We implement a propensity score matching estimator combined with difference-in-differences, and find evidence of a negligible academic cost from high school employment during grade 12.[6]

More recently, I have been particularly interested in the sources of heterogeneity in the wage returns to schooling. While there is now clear empirical evidence from the literature that returns to schooling are highly heterogeneous across individuals, relatively little is still known about the factors that are driving this heterogeneity. In Clark, Joubert & Maurel (2017),

---

[5]By doing so, this project will extend some of my earlier work by i) investigating the simultaneous choice of college and field of study, and ii) modeling the centralized application and admission process to the various college programs.

[6]I also have investigated the effect of college employment on graduation in the French context. In Beffy, Fougère & Maurel (2009), we use geographical variations in low-skill youth unemployment rates to identify the effect of college employment on graduation. Our results provide evidence of a large detrimental effect of holding a part-time job on the probability of graduating from college. We argue that this may be due to the lack of flexibility of the French university system, where evening classes and continuous assessment of students are uncommon.

we shed light on this question by estimating the wage penalties of working in an occupation which requires less than the attained schooling level, e.g. "overeducated" (or underemployed) college graduate working in a high-school job, as opposed to a college job. Overeducation appears to be an important factor behind the heterogeneity in the returns to schooling. Combining data from the National Longitudinal Survey of Youth 1979 (NLSY79) with the Current Population Survey, we find that overeducated employment is a pretty persistent phenomenon, with 66% of workers remaining overeducated after one year. Overeducation is also more frequent and more persistent among blacks and low ability individuals. Besides, not only is it hard for many workers to transition out of overeducated employment, but overeducation is also found to have sizable negative effects on current as well as future wages. Interestingly, this finding bears similarity with the scarring effects that have been found to accompany prolonged unemployment spells.[7] While important, overeducation is only one among various dimensions through which returns to schooling vary across individuals. Some of my future research will focus on the explanatory power of occupations and sectors in accounting for the heterogeneity in returns to college majors.

Some of my research has also focused on the evolution of the returns to skills over time. In Ashworth, Hotz, Maurel & Ransom (2018), we investigate the change in the wage returns to schooling and early work experiences over the past twenty years. Using data from the NLSY79 and NLSY97 cohorts, we estimate a life-cycle model of joint schooling and work decisions where individuals differ based on their observed characteristics (such as test scores and family background), and unobserved ability. This framework allows us to account for the endogenous nature of changes in schooling and work experiences across cohorts, making it possible to isolate price effects (i.e., changes in the value of skills in the labor market) from composition effects (e.g. decline in the average quality of college graduates). Our results show that composition effects are an important driving force behind the evolution of schooling and work experiences wage premia over the last two decades.[8]

4 Individual Responses to Non-Stationarity in Labor Market Environment

Skill prices, along with many other features of the labor market, tend to vary over time. Understanding how individuals respond to these changes is an important and challenging issue. In my research on higher education demand - in particular Beffy, Fougère & Maurel (2012) - I found evidence that wage returns to, and choices of college majors vary across the business cycle. In my subsequent work Arcidiacono, Aucejo, Maurel & Ransom (2019), I also account for the fact that some of the transitions between college and work may result from aggregate labor market shocks. While this was not the primary question of interest in this earlier work, some of my current research is focused on making progress on this question, which I have been tackling from two main different angles.

In a work in progress (Arcidiacono, Gyetvai, Ekaterina & Maurel 2019), we propose a new empirical framework to estimate continuous-time job search models in a non-stationary environment. Non-stationary search models provides a valuable framework for analyzing the behavior of forward-looking job seekers in an uncertain environment, where some of the state variables vary over time (e.g. unemployment benefits varying with time spent unemployed). These types of environment are pervasive in practice. However, taking these models to the data has proven to be challenging from a computational viewpoint, and only relatively few papers have attempted to do so. In this project, we adapt the conditional choice probability estimation methods that have found many applications in the dynamic discrete choice literature to a continuous-time job search environment. To do so, the proposed framework incorporates preference shocks into the search framework, resulting in a tight connection between value functions and conditional choice probabilities. We show that this empirical framework makes it possible to estimate rich job search models in a simple and tractable fashion. In particular, in contrast with the existing literature on this question, we derive a constructive identification strategy that can be used to estimate non-stationary job search models without having to solve any differential equation, and in some cases even avoiding any optimization.[9] The methodological part of the paper is now complete, and we are currently applying the

---

[9]We are also in the process of writing a user-friendly companion Stata package that will allow applied researchers who may not be familiar with the estimation of this type of models to easily estimate non-stationary search models.

proposed methods to analyze the effect of unemployment benefit expiration on the duration of unemployment spells and on accepted wages, using rich longitudinal administrative data from Hungary.[10]

The level of economic uncertainty individuals face may also vary over time. Understanding how individuals respond to fluctuation in uncertainty is crucial to anticipate the consequences of economic shocks, and devise policies that are most effective at mitigating the adverse effects of such shocks. In an ongoing project with Edwin Leuven and Sturla Lokken, Leuven, Lokken & Maurel (2019), we examine the impact of the uncertainty shock associated with the fall of the Berlin Wall on fertility behavior. The fall of the Wall was associated with a drastic and sudden increase in economic uncertainty faced by the inhabitants of former East Germany, who underwent a fast transition from a centrally planned communist economy with full employment to a market economy. Fertility rates among East German women dropped very sharply after the reunification, by as much as 50% over the three years following the fall of the Berlin Wall. Using data from the German Socio-Economic Panel (GSOEP), our identification strategy further exploits variation in individual characteristics that shield women from the effects of uncertainty, such as having a civil servant job. Estimation results exploiting this variation indicate that the economic uncertainty channel does play an important role in accounting for the decline in fertility that occurred after the reunification. Using subjective expectations data available from the GSOEP data, we also find that fertility responses to the fall of the Wall tend to be less pronounced for women who report being more shielded from economic uncertainty (and, e.g., report a lower subjective probability of losing their job). In future work, we plan on leveraging the uncertainty shock associated with the fall of the Berlin Wall to estimate a dynamic structural model of fertility. By doing so, we will be able to predict how fertility would respond to different types of economic shocks.[11]

---

[10]The paper has already been presented in various conferences and seminars.

[11]This research benefited from a Duke Population Research Institute - NIH Developmental Award, and is currently being presented in seminars and conferences.

5 Identification and Estimation of Selection and Treatment Effect Models

My research in econometrics has primarily been concerned with the identification and estimation of selection and treatment effect models. Since the seminal articles by Gronau (1974) and Heckman (1976) forty years ago, endogenous selection has been recognized as one of the key methodological issues in the analysis of microeconomic data. In the canonical example of female labor force participation, the selection issue arises from the fact that the analyst only observes the wages of the individuals who choose to participate to the labor market. To the extent that the decision to enter the workforce is endogenous, as it depends on both observed (e.g. schooling level) and unobserved (e.g. labor market productivity) characteristics which also affect wages, naive regressions of observed wages on observed characteristics typically result in invalid inference regarding the determinants of wages for all women. In D'Haultfœuille & Maurel (2013a), we propose a novel identification strategy for endogenous selection models. Unlike prior methods available in the literature, our approach does not require any instrument for selection (i.e. a covariate affecting selection but not the wage) nor a large support regressor shifting the selection probability to zero or one. In most applications, valid instruments and large support regressors are hard to find. In this paper, we prove identification of the outcome equation, up to a location-scale normalization, under the key assumption that selection becomes independent of the covariates, for arbitrarily large values of the outcome. This assumption is mild provided that selection is indeed endogenous. For instance, it seems reasonable to assume that highly productive individuals are going to enter the workforce with a probability approaching one.

In D'Haultfœuille, Maurel & Zhang (2018), we build on the identification analysis of D'Haultfœuille & Maurel (2013a) and develop a new semiparametric estimation method for selection models that does not require an instrument for selection or a large support regressor. We propose to estimate the covariates effects on an outcome of interest using quantile regressions in the upper tail of the outcome distribution (extremal quantile regressions). We show that the proposed estimator is consistent and asymptotically normal, extending existing results from the extremal quantile literature to a context with sample selection. Importantly for the practical usefulness of our method, we show that our estimation procedure performs well even in small samples. We apply our method to estimate the black-white wage gap among

males from the NLSY79 and NLSY97 cohorts. Accounting for selection into the workforce is key in this context since black males are more likely to dropout from the labor market than white males, and this difference has increased over time. Our estimation results show that pre-labor market factors, such as test scores (AFQT) and parental background, play an important role in explaining the black-white wage gap. Our results also indicate a lack of convergence between both groups, with a (selection-corrected) black-white wage gap of similar magnitude in 2007 and 1990. In order to make the method accessible to a broad set of researchers, we have written and posted online a ready-to-use Stata package (`eqregsel`) that makes it possible to easily implement the method proposed in this paper.[12]

These two papers open up several interesting research avenues, which we plan on pursuing in future work. In particular, we plan on extending our analysis to a context with multiple ($K > 2$) sectors, applying our identification and estimation strategy to estimate state-specific returns to schooling, taking into account the fact that workers are likely to choose their location partly based on their location-specific potential earnings. In another future project, we are also interested in building on the identification strategy of D'Haultfoeuille & Maurel (2013a) to derive and estimate bounds on quantile treatment effects. In D'Haultfoeuille, Maurel & Zhang (2018), we were assuming that the explanatory variables of interest had an homogenous effect across the distribution of the outcome. Here we would relax this assumption, which would result in a lack of point identification in general. But the bounds may still be tight in practice, and would also hold under very mild assumptions. Again, this method should be applicable in a wide range of empirical settings as it will not require any exclusion restriction nor large support regressor.

The identification strategy proposed and used in D'Haultfoeuille & Maurel (2013a) and D'Haultfoeuille, Maurel & Zhang (2018) is very natural in the context of the Roy model of self-selection, where individuals choose between two sectors by maximizing their earnings. In D'Haultfoeuille & Maurel (2013b) we consider the identification and estimation of an extended Roy model, which includes a non-pecuniary component in the selection equation and allows for uncertainty on the potential earnings in each sector. In this paper, we are particularly interested in estimating the non-pecuniary component of the choice equation. Recovering

---

[12]We have also written a paper, D'Haultfoeuille, Maurel, Qiu & Zhang (2019), that describes the Stata package `eqregsel`. This article will be submitted to the Stata journal soon.

this component is crucial to be able to tell apart the relative importance of monetary returns versus preferences in the context of sorting across sectors, which is a question that I have examined in some of my other work on college major choice (Beffy, Fougère & Maurel 2012) and occupational choice (Arcidiacono, Hotz, Maurel & Romano 2019). By making the most of the structure of the selection equation, we show that the non-pecuniary component is identified from the knowledge of the covariate effects on earnings, as long as one covariate is continuously distributed. Together with D'Haultfoeuille & Maurel (2013a), this result implies that identification can be achieved without any instrument for selection nor large support covariates. From this analysis, we further derive bounds on the distribution of the expected monetary returns. We propose a three-stage semiparametric estimation procedure for this model, which yields consistent and asymptotically normal estimators. Finally, we apply our method to quantify the relative importance of non-pecuniary factors and expected returns to schooling in the decision to enroll in higher education in France, and find that non-pecuniary factors play a dominant role in this decision. Notably, our estimates imply that the higher education attendance rate would fall from 83.1% to 72% in the absence of non-pecuniary factors. This decrease is eight times larger than the one associated with an hypothetical 10% decrease in labor market earnings of higher education attendees. This article received the 2015 Dennis J. Aigner Award for the most significant contribution in empirical econometrics published by the *Journal of Econometrics* in 2013 or 2014.

In a related line of research, I have also been working on developing new identification strategies and estimation methods for treatment effects. When selection into treatment is endogenous, as is in practice the case in most situations, then one cannot simply infer the average treatment effect by comparing the mean outcome of treated individuals with that of the untreated individuals. The case of the wage returns to college is a classical example, where the treatment (college attendance or graduation) is typically endogenous with respect to future wages, so that the raw college wage premium tends to substantially differ from the causal effect of college on wages. In a work in progress (Khan, Maurel & Zhang 2019), we consider the identification and estimation of a treatment effect model where the treatment status and the outcome are both binary, and modeled as non-linear functions of observed and unobserved characteristics. In the absence of restrictive large-support type assumptions on the distribution of the covariates, these models are generally only partially identified,

and inference can be fragile even if these support conditions hold. In this project, we show that imposing a linear factor structure on the unobservables makes it possible to point-identify the parameters of interest, including the coefficient associated with the endogenous regressor in the outcome equation, under weaker conditions. In particular, we show that an exclusion restriction, requiring an explanatory variable in the outcome equation to be excluded from the treatment equation, is no longer necessary for identification. Furthermore, we show that support conditions on the included covariates in the outcome equation can also be substantially weakened. Finally, we propose an estimation procedure for this model that, by standard regularity conditions, yields root-n consistent and asymptotically normal estimators.[13]

6 Empirical Content of Beliefs Formation Models

How much information individuals have, and how individuals incorporate the available information to form their beliefs about future outcomes, is a fundamental and difficult question that arises whenever one seeks to understand decision making under uncertainty. Several of my papers have highlighted the central role played by imperfect information, and beliefs updating, in the context of post-secondary schooling choices and occupational choices (see, in particular, Arcidiacono, Aucejo, Maurel & Ransom, 2019, Arcidiacono, Hotz, Maurel & Romano, 2019, and my work-in-progress on the French centralized admission system).[14]

Much of the existing research on this question relies on strong assumptions about beliefs formation, as well as about the type of information that is available to the individuals at the time of their decisions. A standard assumption in economics, known as rational expectations, is that individuals have expectations that do not systematically differ from the realized outcomes, and efficiently process all private information to form these expectations.

---

[13]This paper has been presented in various seminars and conferences. We are currently in the process of finalizing a draft.

[14]In my work on the Roy model and the determinants of sorting across sectors (D'Haultfoeuille & Maurel, 2013), a distinctive feature of the framework considered in that paper - which allows for imperfect information - is that we remain agnostic on the information set of the agents at the time of the choice. In other contexts though, the information set itself may be of interest.

In D'Haultfoeuille, Gaillac & Maurel (2018), we analyze the empirical content of the rational expectations assumption. We first develop a new method to test for rational expectations. The test can be used in a variety of data environments, including in frequent cases where realized outcomes and subjective beliefs - e.g., realized income and expected income - are observed in two different datasets that cannot be matched. In this context, we show that the rational expectations hypothesis is equivalent to the distribution of realizations being a mean-preserving spread of the distribution of beliefs. This condition is easy to test using recent statistical tools from the moment inequality testing literature. Then, in situations where rational expectations are violated, we propose a way to quantify the magnitude of the deviations from rational expectations. Specifically, using recent advances from the optimal transport literature, we introduce a concept of, and propose to estimate the minimal departures from rational expectations that are consistent with the data. This naturally gives rise to a way to conduct a sensitivity analysis on the assumed form of expectations. We apply these methods to test for and measure deviations from rational expectations about future earnings using data from the Survey of Consumer Expectations (New York Fed), and examine the consequences of such departures in the context of a model of consumption over the lifecycle. We find in particular that departures from rational expectations account for some of the over-insurance to permanent income shocks, as well as the excess sensitivity of consumption to transitory shocks that have been documented in the empirical consumption literature.

Importantly for practical purposes, the test of rational expectations, and the estimation of the departures from rational expectations can be performed at a modest computational cost. In order to make the method accessible to a broad set of researchers, we have written and posted online on the Comprehensive R Archive Network a ready-to-use R package `RationalExp` that makes it possible to easily implement these methods.

While subjective expectations data have become more frequently available in recent years, many of the standard datasets used in labor economics, economics of education as well as in other fields in applied microeconomics such as industrial organization lack such data. As a consequence, it is also important to understand what can be inferred from standard observational data, in terms of information set as well as beliefs formation. Recently, I have started to explore these questions with a Ph.D. student from our program, Jackson Bunting (Bunting & Maurel 2019). In this project, we are particularly interested in providing conditions under

which one can identify, and estimate from longitudinal data on choices and realized outcomes (e.g. choice of occupation and realized earnings) the relative importance of i) predictable heterogeneity, ii) initially unpredictable heterogeneity that is gradually revealed over time, and iii) uncertainty, in accounting for the variation in outcomes. To do so, our approach uses recent results from the measurement error literature in econometrics and statistics, combined with behavioral restrictions regarding how individuals incorporate the available information to form their beliefs, along with how choices depend on those beliefs.

References

Altonji, J. G., Arcidiacono, P. & Maurel, A. (2016), The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects, *in* E. A. Hanushek, S. Machin & L. Woessmann, eds, 'Handbook of Economics of Education', Vol. 5, Elsevier, pp. 305–396.

Arcidiacono, P., Aucejo, E., Maurel, A. & Ransom, T. (2019), College attrition and the dynamics of information revelation. Working paper.

Arcidiacono, P., Gyetvai, A., Ekaterina, J. & Maurel, A. (2019), Conditional choice probability estimation of continuous-time job search models. In progress.

Arcidiacono, P., Hotz, V. J., Maurel, A. & Romano, T. (2019), Ex ante returns and occupational choice. Working Paper.

Ashworth, J., Hotz, V. J., Maurel, A. & Ransom, T. (2018), Changes across cohorts in wage returns to schooling and early work experiences. Working paper.

Beffy, M., Fougère, D. & Maurel, A. (2009), 'The effect of part-time work on post-secondary educational attainment(in french)', *Economie et Statistique* **422**, 31–50.

Beffy, M., Fougère, D. & Maurel, A. (2012), 'Choosing the field of studies in postsecondary education: Do expected earnings matter?', *Review of Economics and Statistics* **94**, 334–347.

Belzil, C., Maurel, A. & Sidibe, M. (2018), Estimating the value of higher education financial aid: Evidence from a field experiment. Working Paper.

Bound, J., Lovenheim, M. F. & Turner, S. (2010), 'Why have college completion rates declined? An analysis of changing student preparation and collegiate resources', *American Economic Journal: Applied Economics* **2**(3), 129–57.

Bunting, J. & Maurel, A. (2019), Heterogeneity, uncertainty and learning: A semiparametric identification analysis. In progress.

Buscha, F., Maurel, A., Page, L. & Speckesser, S. (2012), 'The effect of employment while in high school on educational attainment: a conditional difference-in-differences approach', *Oxford Bulletin of Economics and Statistics* **74**, 380–396.

Clark, B., Joubert, C. & Maurel, A. (2017), 'The career prospects of overeducated americans', *IZA Journal of Labor Economics* **6**, –.

De Groote, O., Luflade, M. & Maurel, A. (2019), The effects of informational interventions in a centralized college-major admission system: Evidence from admission post-bac and parcoursup. In progress.

D'Haultfoeuille, X., Gaillac, C. & Maurel, A. (2018), 'Rationalizing rational expectations? Tests and deviations', *NBER Working Paper* **25274**.

D'Haultfoeuille, X. & Maurel, A. (2013*a*), 'Another look at the identification at infinity of sample selection models', *Econometric Theory* **29**(1), 213–224.

D'Haultfoeuille, X. & Maurel, A. (2013*b*), 'Inference on an extended Roy model, with an application to schooling decisions in France', *Journal of Econometrics* **174**, 95–106.

D'Haultfoeuille, X., Maurel, A., Qiu, X. & Zhang, Y. (2019), Estimating selection models without instrument with Stata. Working Paper.

D'Haultfoeuille, X., Maurel, A. & Zhang, Y. (2018), 'Extremal quantile regressions for selection models and the black-white wage gap', *Journal of Econometrics* **203**(1), 129–142.

Gronau, R. (1974), 'Wage comparisons - a selectivity bias', *Journal of Political Economy* **82**(6), 1119–1143.

Heckman, J. J. (1976), 'The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models', *Annals of Economic and Social Measurement* **5**, 475–492.

Khan, S., Maurel, A. & Zhang, Y. (2019), Informational content of factor structures in simultaneous discrete response models. In progress.

Leuven, E., Lokken, S. & Maurel, A. (2019), Fertility and uncertainty. In progress.

Maurel, A. (2015), 'Quels sont les determinants des choix d'orientation dans l'enseignement superieur', *Regards Croises sur l'Economie* **1**(16), 69–79.

Maurel, A. & Shephard, A. (2019), Skills, mismatch and labor market search. In progress.