NBER WORKING PAPER SERIES

SURVEY EXPERIMENTS ON ECONOMIC EXPECTATIONS

Andreas Fuster
Basit Zafar

Working Paper 29750
http://www.nber.org/papers/w29750

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2022

Survey Experiments on Economic Expectations
Andreas Fuster and Basit Zafar
NBER Working Paper No. 29750
February 2022
JEL No. C83,C93,D84

ABSTRACT

In this chapter, we discuss field experiments in surveys that are conducted with the purpose of learning about expectation formation and the link between expectations and behavior. We begin by reviewing the rationale for conducting experiments within surveys, rather than just relying on observational survey data. We then outline the most commonly used experimental paradigm, randomized information provision, along with some examples. Next, we outline a few methodological issues that are important to consider in the design of such experiments. We also provide a discussion of existing extensions of this paradigm, as well as of alternative approaches.

Andreas Fuster
Swiss Finance Institute @ EPFL
Quartier UNIL-Dorigny
Extranef 216
Lausanne CH-1015
Switzerland
and CEPR
andreas.fuster@gmail.com

Basit Zafar
Department of Economics
University of Michigan
611 Tappan Street
Ann Arbor, Michigan 48109
and NBER
basitak@gmail.com

1 Introduction

In recent years, there has been increased interest in subjective expectations as a driver of economic behaviors, as is reflected by the publication of this *Handbook*. A first key step to understand the role of expectations is the ability to measure expectations, as is commonly done in surveys. Such data then allow the researcher to estimate models of decision-making by relaxing standard assumptions (such as full information rational expectations). While such an approach does not require the researcher to observe the respondent's information set and to understand how expectations are formed, doing so is necessary for counterfactuals, when the researcher wants to understand how behavior and expectations may change in different scenarios. This typically requires some plausibly exogenous variation in expectations. An increasingly popular way to obtain such exogenous variation is via experiments in surveys, which we will review in this chapter.

We begin by outlining the conceptual arguments for why experimentally induced variation in expectations is useful to understand decision-making, and also to understand expectation formation itself. Next, we review the most commonly used experimental paradigm, so-called information provision experiments. Here, we complement a comprehensive recent review paper by Haaland et al. (2021), who show that the number of published papers in top economics journals that use such experiments has strongly increased in recent years. Rather than discussing the wide variety of applications of the paradigm, which will also be covered in other chapters of this book, we restrict ourselves to a few examples in the areas that we have worked in, namely human capital investments and macroeconomics/finance. We also provide an overview of selected methodological issues related to this paradigm.

We then turn to extensions of and alternatives to randomized information provision. One extension is to move beyond exogenously providing a given piece of information and instead allowing survey respondents to choose which information (if any) they want to see. Incorporating this aspect arguably brings the experimental setting closer to everyday decision situations. An alternative approach we discuss is the elicitation of conditional expectations in a number of assumed states of the world (or "vignettes"). This approach can allow researchers to collect much richer data, though may not be equally suitable for all settings. Finally, we provide some directions for future work.

2 Why (field) experiments on expectations?

Survey experiments on expectations are primarily useful for two purposes. One, for understanding decision-making under uncertainty. Second, for understanding the process of expectation formation. In this section, we outline each of these identification issues, in turn.

2.1 Understanding decision-making

Researchers in economics are typically interested in understanding determinants of choices. Specifically, the researcher is interested in recovering the preference parameters and factors of the economic environment that can explain choices. For the sake of exposition, we focus on the choice of college major here; the same arguments apply to any choice under uncertainty. The arguments presented here are a simplified version of those presented in Wiswall and Zafar (2015a); interested readers should refer to the paper for a more in-depth discussion.

Going back to Freeman (1971), economists have been interested in understanding how sensitive educational choices are to potential earnings. Take the following reduced-form model:

$$(\ln \pi_{k,i} - \ln \pi_{\tilde{k},i}) = \beta_0 + \beta_1(\ln w_{k,i} - \ln w_{\tilde{k},i}) + C_i'\delta + \nu_k + \psi_{k,i}, \tag{1}$$

where $\pi_{k,i}$ is individual i's subjective probability of graduating with major k, $w_{k,i}$ are the earnings for individual i in major k, C_i is a vector of individual-specific characteristics, ν_k is a major k fixed effect, and $\tilde{k}$ is the reference major. The residual error $\psi_{k,i}$ equals $\gamma_{k,i} - \gamma_{\tilde{k},i} + \varepsilon_{k,i}$, that is, it consists of unobserved relative taste differences ($\gamma_{k,i} - \gamma_{\tilde{k},i}$) and a component $\varepsilon_{k,i}$ which reflects all other residual components. The goal is to recover β_1, which tells us how sensitive major choice is to the earnings (note that the log-log form of the regression means that β_1 has an "elasticity" interpretation). Without data on subjective expectations, identification requires three additional layers of assumptions:

1. An assumed mapping between revealed or actual earnings in the individual's chosen major to beliefs;
2. An assumed model for counterfactual beliefs about earnings in majors not chosen by the individual; and

3. An assumed distribution of tastes for all majors.

Cross-sectional data on expectations allows the researcher to relax the first two assumptions. In that case, equation (1) becomes:

$$(\ln \pi_{k,i} - \ln \pi_{\tilde{k},i}) = \beta_0 + \beta_1(\ln \hat{w}_{k,i} - \ln \hat{w}_{\tilde{k},i}) + C_i'\delta + \nu_k + \psi_{k,i}, \tag{2}$$

where $\hat{w}_{k,i}$ is i's beliefs about earnings in major k. Using beliefs and choice data at *only* one point in time, a cross-sectional OLS estimation of equation (2) is unbiased only if the residual term—which includes individual components reflecting individual variation in tastes for each of the majors in this example—is uncorrelated with beliefs about earnings, an assumption that is unlikely to hold in reality. In the absence of data on beliefs and choices at multiple points in time, there is little else that a researcher can do. Now, if the researcher had data on choices and beliefs at another point in time, then she could estimate equation (2) in differences and net out the individual taste components ($\gamma_{k,i} - \gamma_{\tilde{k},i}$), that is:

$$[(\ln \pi'_{k,i} - \ln \pi'_{\tilde{k},i}) - (\ln \pi_{k,i} - \ln \pi_{\tilde{k},i})]$$
$$= \beta_0 + \beta_1[(\ln \hat{w}'_{k,i} - \ln \hat{w}'_{\tilde{k},i}) - (\ln \hat{w}_{k,i} - \ln \hat{w}_{\tilde{k},i})] + \epsilon'_{k,i} - \epsilon_{k,i},$$

where $\pi'_{k,i}$ and $\hat{w}'_{k,i}$ are the stated probabilistic choices and earnings beliefs at a second point in time. The estimates of this model are equivalent to adding individual fixed effects (FE) as individual dummy variable indicators to equation (2).

So how does one get data on probabilistic choices and earnings at multiple points in time? One obvious approach is to track individuals over time and collect panel data on beliefs and choices. However, identification requires that the *change* in beliefs about unobserved events or measurement error, given by $\epsilon'_{k,i} - \epsilon_{k,i}$, is mean-independent of the *changes* in observed beliefs about earnings. The issue with longitudinal information on beliefs collected over several months or years is that the individual and major-specific taste parameters $\gamma_{k,i}$ may change. Instead, if the researcher is able to experimentally perturb beliefs and choices *within* a survey, one can credibly claim that the identification arguments hold (that is, tastes are likely to be truly time invariant over the horizon of a survey). Violations of the assumption would occur if the experimental perturbation also changes beliefs about outcomes that the researcher does not ask about. This is not directly testable, but the researcher could collect data on other major-related outcomes both before and after the experimental intervention. This is the approach that Wiswall and Zafar

(2015a) use to estimate the sensitivity of college major choice to earnings. An added advantage of a stylized information experiment embedded in a survey is that the shock to the individual's information set is truly exogenous, and hence the identifying variation is "clean."

An alternative approach to generate arguably exogenous variation in expectations and choices is to rely on "natural experiments." For example, researchers have leveraged cross-sectional variation across counties in the ideological predisposition of constituents to use the election outcome as an exogenous shifter of expectations and sentiment (Mian et al., 2015; Conlon and Zafar, 2017; Benhabib and Spiegel, 2018; De Stefani, 2021). Likewise, surprise monetary policy announcements have been shown to impact consumers' confidence in the state of the economy (Lewis et al., 2019). These studies show a strong immediate impact on sentiment. The impact on expectations, though, generally tends to be weaker. It is also not clear *which* expectations are really affected by such shocks. Hence, it is not surprising that many of these studies find only limited impact on behavior.

Herein lies the trade-off between stylized survey-based information experiments and natural experiments. Stylized experiments are powerful in the sense that, by design, the information is salient and hence likely to impact expectations. Natural experiments accord less control to the researcher. In addition, few individuals may have been attentive to the event. However, natural settings are arguably more reflective of how individuals go about acquiring information in the real world, and are less susceptible to an experimenter demand effect. Later in the chapter, we discuss how survey experiments may incorporate some of this realism.

2.2 Understanding expectation formation

A more common use of information experiments to date has been to study the process of expectation formation. This is relevant not only for policy, in the sense of inferring whether individuals have accurate expectations, but also for informing the modeling of expectations. Returning to the example above, let $i's$ earnings expectations in major k be given by:

$$\hat{w}_{k,i} = f(\Omega_i),$$

where Ω_i is i's information set at a given point in time that she uses to form earnings expectations. The researcher's goal is to say something about the function f and about what the possible components of the information set are, but neither are easily observable. The researcher may have a prior about the elements of the information set. Suppose the researcher thinks that the individual may base her earnings beliefs in major k, in part, on her beliefs about what college graduates earn (let us denote this as z) and that i's beliefs about z may be biased, that is, $\hat{z}_i \neq z$. To test this, the researcher can provide objective information about z to the individual, and re-elicit her earnings expectations $\hat{w}'_{k,i}$. If $\hat{w}'_{k,i}$ differs from the pre-information expectations $\hat{w}_{k,i}$, that tells us that in fact z is an element of the individual's information set *and* that she had biased beliefs about it.

Now, if the re-elicited expectations are exactly the same as baseline expectations, that could imply one of two things: (1) z is not an element of i's information set, or (2) z is in fact an element of i's information set but i has correct beliefs about z. To tell apart these two very different implications, the researcher needs to elicit i's beliefs about z *before* the information is given to her.[1]

3 Information provision experiments

The most common experimental paradigm within surveys has been the randomized provision of information. Expectations can then also be linked to (intended or actual) economic behaviors. In this section, we describe the basic ingredients of this paradigm, and then discuss a few examples from different areas. We focus our discussion mainly on details of the design, less on results (since those will be discussed in later chapters).

3.1 Design basics

Most information provision experiments in the literature consist of three main stages:

1. Measurement of respondents' prior expectations about the variable(s) of interest, y, and/or the variable(s) about which information will be provided, x. We denote i's prior expectations by $\hat{y}_{i,prior}$ and $\hat{x}_{i,prior}$.

[1]There is another possibility: it could be that z is in individual i's information set, and i is misinformed about z but does not find the provided information credible/trustworthy. Empirical work, to date, has not systematically investigated this possibility. We return to this issue in section 4.4.

2. One of K possible pieces of information, or signals, is randomly provided to respondents. We denote the signal received by respondent i as x_i.

3. Respondents' posteriors $\hat{y}_{i,post}$ are measured, as well as potentially (intended or actual) behaviors that $\hat{y}_i$ should plausibly influence.

Various different implementations of the three steps above can be found in the literature. For example:

- Expectations about the variable of interest y are measured either as full subjective distributions, as prominently advocated by Manski (2004), or as point forecasts (but often with a qualitative question about the respondents' confidence in their forecast). The measurement method can also differ between steps 1 and 3.

- x is often information about past realized y (sometimes provided over different past horizons across respondents), some factual information, or expert forecast(s).

- In some cases, all respondents are provided with some piece of information (varying across respondents) while in other cases, only some respondents receive information while a control group does not.

- It is also possible to provide a group of respondents with more than one signal, and compare them to respondents who receive only one or the other signal.

- In addition to respondents' own expectations, higher-order beliefs (beliefs about the expectations of others) can be elicited (Coibion et al., 2021b).

Since information x_i is randomized across respondents, one can in principle directly estimate the treatment effect of a given piece of information on posterior beliefs:

$$\hat{y}_{i,post} = \alpha + \sum_{k=1}^{K} \beta_k I(x_i = x^k) + \varepsilon_i, \tag{3}$$

where $I(x_i = x^k)$ is a dummy variable indicating that respondent i received signal k.

Although this regression allows to cleanly estimate the effect of a signal k on average posterior expectations, it is not necessarily informative about the effects on individual expectations. To see why, imagine that different respondents' priors are symmetrically distributed around a signal. Then, even if respondents' fully align their posteriors with the signal if they receive it, the average posterior would be unaffected, as those with priors above the signal revised their beliefs downwards, and vice versa for those with priors below the signal.

To better capture how receiving a given signal k affects beliefs, it is therefore common to relate a respondent's posterior belief to the difference between the signal and the respondent's prior *about the signal*, often called the "perception gap":

$$\hat{y}_{i,post} = \alpha + \sum_{k=1}^{K} \beta_k \underbrace{I(x_i = x^k)}_{\substack{\text{Treatment}\\ \text{Indicator}}} \underbrace{(x^k - \hat{x}^k_{i,prior})}_{\substack{\text{Perception}\\ \text{Gap}}} + \varepsilon_i. \tag{4}$$

Furthermore, when the prior about the outcome was measured, researchers often directly use the belief *update* (= posterior minus prior) as the dependent variable:

$$(\hat{y}_{i,post} - \hat{y}_{i,prior}) = \alpha + \sum_{k=1}^{K} \beta_k I(x_i = x^k)(x^k - \hat{x}^k_{i,prior}) + \varepsilon_i. \tag{5}$$

A nice feature of this formulation is that it can naturally be linked to a Bayesian learning model with Gaussian distributions, where the posterior is the weighted average of a person's prior, with weight $1 - \omega$, and a signal received, with weight ω. This weight ω in this case corresponds to the coefficient β_k. Theoretically, β_k should then decrease in the dispersion of a respondent's prior (somebody who is already very confident in their forecast will update little in response to a signal). Furthermore, it should increase in the (perceived) precision of the signal. See e.g. Cavallo et al. (2017) for additional discussion.

In practice, equation (5) above is estimated including not only the interaction between the treatment indicator and the perception gap, but also the two terms separately.[2] Simplifying to the case of a single signal x that randomly gets shown to only some respondents, the regression would then be specified as:

$$(\hat{y}_{i,post} - \hat{y}_{i,prior}) = \alpha + \beta I(x_i = x)(x - \hat{x}_{i,prior}) + \gamma_1 I(x_i = x) + \gamma_2(x - \hat{x}_{i,prior}) + \varepsilon_i. \tag{6}$$

The coefficient of interest is still β, but the two additional coefficients are useful in their

[2] One can also estimate a version of equation (5) where the dependent variable is the posterior, instead of the update. In that case, $\hat{y}_{i,prior}$ is included on the right-hand-side. This specification is more flexible. However, most work uses revisions as the dependent variable since the estimates are easier to interpret.

own right: γ_1 allows to test whether receiving a signal that does not differ from a respondent's prior about the signal has a systematic effect on updating (or on posterior beliefs), while γ_2, which is estimated off of the control group, accounts for the possibility that respondents' revisions may be correlated with their priors, for instance due to "noise" in the answers in the prior stage.

In addition to individual updating, researchers also often study how receiving a given signal affects the dispersion of posterior expectations within a group that receives the same signal or across groups that receive different signals.

Finally, the basic design is often augmented by a fourth stage, where expectations are again elicited in a follow-up survey that is run a few weeks (or sometimes a few months) after the original survey. One can then repeat the regressions above using $\hat{y}_{i,followup}$ instead of $\hat{y}_{i,posterior}$ as dependent variable. Such follow-up data is useful to measure the persistence of the effects of the information, and is often used to dismiss demand effects or simple numerical anchoring to the provided information as drivers of effects within the main survey.[3]

3.2 Expectations and behavior

To the extent that the randomized information provision shifts respondents' expectations, this exogenous shift can then be used to study the effect of expectations on economic behaviors that, according to theory, should be driven by expectations. This, of course, requires data on behavior, which is not always readily available. We discuss this issue in the next subsection.

If the randomized information provision differentially affected respondents' beliefs, and denoting by a_i some action by respondent i that is taken after the posterior stage, one can estimate the following model:

$$a_i = \delta + \kappa \widehat{y_{i,post}} + \Gamma Z_i + \nu_i, \tag{7}$$

where the posterior expectation $\widehat{y_{i,post}}$ is instrumented using some variation of equa-

[3] An alternative way to assess the strength of numerical anchoring is to feature a placebo treatment in which a number on an unrelated topic is provided. Coibion et al. (2021c) do so in their study of inflation expectations, providing one group with information on population growth. They find only very small effects of this treatment, suggesting that numerical anchoring is not a big concern in such studies.

tions (4) or (6).[4] Z_i is a vector of other individual characteristics that could also influence the action a_i. Including these controls can help with estimating κ more precisely, though it is not strictly necessary if the provided information was successfully randomized so that it is orthogonal to all relevant individual characteristics.

Estimating equation (7) using the instrumented expectations, rather than just with OLS based on the actual measured expectations, has two advantages. First, it reduces concerns about omitted variables that affect both behaviors and expectations and therefore could lead to a spurious correlation. Second, the instrumental variable strategy can help with measurement error in expectations, which in a standard OLS regression would tend to attenuate the coefficient κ toward zero.

Nevertheless, the exogenous information provision is not a panacea to learn about the causal effect of expectations about a given economic variable on behavior, because to allow for such an interpretation, the exclusion restriction also needs to be satisfied. In other words, the provided information should only affect action a through its effect on expectations about the specific variable y. As we discuss below, this can be challenging especially when asking macroeconomic questions, where expectations about different types of variables are likely correlated and affect many decisions jointly. However, even in cases where a singular interpretation is difficult, the "reduced form" effect of the experimental treatment on behavior is of course still of interest.

3.3 Examples

Human capital investments. Information frictions have been well-documented in the context of educational choices and health behaviors. Our goal here is not to cover the large literature—here, we summarize a few papers, with a focus on methodology. The chapters in this volume by Delavande (2022) and Giustinelli (2022) provide more extensive overviews of studies in this area.

While several information-based randomized controlled trials have shown significant effects on educational and health outcomes, few of these papers tend to collect data on both expectations and behaviors. One of the most influential information experiments in

[4] One would generally want to instrument for the level of expectations after the experimental intervention, rather than the change between prior and posterior stage as in equation (6). This can be done for instance by simply adding the prior expectation(s) to the vector of controls Z_i.

this area is the study by Jensen (2010). His study design closely follows the paradigm mentioned above: that is, elicitation of priors at the baseline, randomized provision of information about schooling returns to a treatment group, and re-elicitation of beliefs in the short-term as well as data on choices in the long-term. In this particular context, since there were no public data on the relationship between schooling and returns, the author himself collected such data—these formed the basis of the information intervention. He then first documents that eighth-graders in the Dominican Republic underestimate returns to secondary schooling. Next, he shows that a treated group of students who are provided with measured returns (which are simply the average earnings conditional on different years of schooling) have higher perceived returns to schooling four to six months after the intervention, and attain 0.2-0.3 more years of schooling over the next four years, compared to a control group.[5] A nice feature of this study is that it documents systematically biased baseline expectations, provides direct evidence on information impacting beliefs, and shows meaningful impacts on actual choices (which is only possible because longer-term data were collected).

Beliefs about returns is only one side of the picture. Perceived costs, especially at the post-secondary level, are likely a factor in schooling decisions, especially in the US. Bleemer and Zafar (2018) find that households' perceptions of college costs and benefits tend to be systematically biased, with larger biases among lower-income and non-college households. They embed a randomized information experiment within the survey where respondents are randomly exposed to objective information about either average college "returns" or costs. The "returns" experiment leads to an increase in the intended likelihood of college attendance, with the impacts being larger for disadvantaged households (since they tend to be more misinformed ex-ante). The impacts persist in a follow-up survey two months later. On the other hand, they find no impact of the cost information treatment. Their results generally suggest that information frictions might be larger for more disadvantaged households.

The studies above only provided information about average earnings. However, there is a fair amount of dispersion in earnings, even conditional on the level/kind of human capital choice. This (perceived) uncertainty could play a role in individuals' decisions. Wiswall and Zafar (2015a) conduct a multi-stage information experiment on undergraduate students at New York University and collect beliefs about their *self* earnings—both

[5] Importantly, the impacts of actual schooling are observed only for students from higher-income households, suggesting that information frictions are not the only binding constraint in this context.

average earnings and earnings uncertainty—at various points in the life cycle. Specifically, for earnings uncertainty, they ask respondents about the likelihood of earning more than certain thresholds if they were to graduate with different majors. They collect beliefs data on not only earnings but also several other outcomes (such as marriage, fertility, spousal earnings). Importantly, at the baseline, they also elicit *population* beliefs (that is, the students' beliefs about the earnings and other outcomes of current college graduates). After the baseline elicitation, students are provided with detailed objective information about the earnings distribution by college major and employment likelihood of current college graduates. Self beliefs are subsequently re-elicited. In this design, since population beliefs were elicited prior to the revelation of the objective information, it is possible to assess the nature of the errors at the individual level. This rich panel of beliefs (all elicited within one survey) reveals that students have biased beliefs about population earnings and there is considerable heterogeneity in errors, which is uncorrelated with observable characteristics. Unique to their setup, the authors are able to study how students revise their self beliefs, and find substantial heterogeneity in the students' updating heuristics (Wiswall and Zafar, 2015b). They also find large impacts on intended major choice, with the intervention nudging students towards higher-paying majors. This study is somewhat underpowered to investigate impacts on actual major choice.[6]

While these papers provide information about population earnings or costs, there are also field interventions that provide personalized information to students or their parents about the student's (relative) ability (Bobba and Frisancho, 2016; Dizon-Ross, 2019; Franco, 2019). The typical set-up in these papers is as follows: (1) the researcher elicits baseline expectations about relative ability and observes either actual human capital investments or intended choices; (2) a randomized group is provided with a signal about ability; and (3) data on updated beliefs and intended/actual choices are collected. These papers tend to find that individuals are quite misinformed about their ability and that providing such information generally leads to an improvement in the match between students and choices/investments, although impacts are quite heterogeneous.[7]

[6] Conlon (2019), in a field experiment at a public university, finds that students, on average, tend to underestimate mean salaries by majors. Importantly, he finds that students who are treated with earnings information about a given field are significantly more likely to major in it. The importance of information frictions has also been studied by complementing experimental variation with large scale surveys (Hastings et al., 2015).

[7] For example, Franco (2019) finds that feedback discourages low-performing students; Bobba and Frisancho (2016) find that information is processed more efficiently by higher-socioeconomic students; Dizon-Ross (2019) finds that poorer, less-educated parents tend to have less accurate baseline beliefs, and that their

Macroeconomics and finance. The use of information experiments in the context of macroeconomics has exploded in recent years. Initial survey experiments in this area focused on households' or firms' expectations of future inflation, which play a key role in standard macroeconomic models and had also been highlighted by policymakers as an area where more empirical work would be valuable (Bernanke, 2007).

Focusing on households, Armantier et al. (2016) directly follow the basic design above, providing respondents with information either about past price changes (for food and beverages), or about the predictions of professional forecasters. They then study respondents' updating of expected inflation as a function of a respondent's perception gap (i.e., the difference between the signal a respondent received and their prior about the signal). Cavallo et al. (2017) implement similar experiments, comparing estimated treatment effects from the low-inflation US context to those found in Argentina, where inflation is much higher, and also comparing the effects of showing either price changes on selected goods, official statistics, or both. Binder and Rodrigue (2018) study how longer-term inflation expectations are affected when respondents are either informed about the Fed's inflation target or shown a time series (and summary statistics) of past inflation.[8]

More recent work by Coibion et al. (2021c) compares effects on inflation expectations of eight different treatments, using a sample of nearly 20,000 individuals.[9] These authors are also able to link the survey responses to spending data measured by Nielsen, which allows the study of actual (not reported) spending effects of exogenous changes in inflation expectations. A related study, Coibion et al. (2021a), presents information about past inflation to a subset of survey respondents and then measures effects on inflation expectations and on subsequent self-reported spending. These authors provide a thorough discussion of the issue raised above regarding the interpretation of the estimated effects on behavior: providing information about recent inflation may affect subsequent spending not just through the effects on future expected inflation, but also because survey respondents update their perceptions and expectations about other variables.[10]

beliefs/investments respond more to information, relative to richer parents.

[8] Binder and Rodrigue (2018) do not feature a control group and furthermore show both signals to all respondents, but in randomized order and eliciting beliefs after both treatments.

[9] The large sample size further makes it possible to run follow-up surveys with meaningful sample sizes 3 and 6 months after the initial survey.

[10] Another recent study that experimentally manipulates expectations and studies effects on (credit-card) spending is Galashin et al. (2020). Other work has used similar information provision experiments to study the expectations and behaviors of firms (Coibion et al., 2018, 2019). We refer to D'Acunto et al. (2022) and Candia et al. (2022) in this volume for more detailed discussion of work in this area.

Armona et al. (2019) apply an information-provision experiment to the study of individuals' expectations of future house price growth. Respondents receive information about either past 1-year or past 5-year local house price growth (or no information, in the control group), and expectations are similarly measured over the 1-year and 5-year horizon. These different horizons are studied because empirically, house price growth exhibits strong momentum (positive autocorrelation) over the 1-2 year horizon, but mean reversion over longer horizons such as 5 years, and this design allows studying to what extent people's updating after learning about past growth is consistent with these patterns.[11] The exogenous variation in expectations generated by the information provision is then also used to study the effects of expectations on investment behavior within a stylized incentivized experiment, where respondents could allocate some funds either to a risk-free account or to a fund that pays a return equal to local house price growth. Bottan and Perez-Truglia (2020) apply similar information provision to study real-world home selling behavior, detecting substantial effects.[12] See Kuchler et al. (2022) in this volume for further discussion of work in this area.

Roth and Wohlfart (2020) study how beliefs about the likelihood of a recession affect people's expectations of personal unemployment, and how this in turn may affect behavior. After eliciting priors, these authors generate variation in recession expectations by providing a forecast from the Survey of Professional Forecasters—half the respondents randomly get the most optimistic forecaster (lowest recession probability), the others the most pessimistic forecaster (highest probability). This means there is no need for a control group. Qian (2020) uses a similar design to study the effect of expected house price growth on intended spending.

In addition, information provision experiments have also been used to study beliefs about the stock market and effects on investment behavior (Hanspal et al., 2020; Laudenbach et al., 2021), or to study effects of interest rate expectations (Coibion et al., 2020a; Link et al., 2021). Beutel et al. (2021) study how randomized provision of a central bank's

[11] A methodological feature of this study is that respondents are randomly asked all questions either in terms of levels (the price of a typical house in USD) or in growth rates (the percent change in the price of a typical house), based on earlier evidence suggesting that framing could affect inference about extrapolation vs. mean reversion in beliefs (Glaser et al., 2007). In Armona et al. (2019), results are consistent across the two frames.

[12] This study directly measures the effects of random information provision on real-world behavior, without being able to measure expectations of the home sellers directly. However, to still obtain measures of the elasticity of behavior with respect to expectations, the authors run an auxiliary survey study to measure the effects of the same information provision on expectations only.

warnings about financial stability risks affects risk perceptions and investment behavior in a stylized setting.

The examples mentioned above in the context of human capital investments typically (though not always) elicit expectations about outcomes that the individual has full or partial control over. For instance, the individual chooses whether to attend college or not, and can at least influence future earnings (e.g. via labor supply choices). This is in contrast to most applications in the area of macroeconomics and finance where expectations are generally elicited about events that the individual has no control over (such as the rate of inflation). In this case, heterogeneity in expectations depends on individual-specific beliefs regarding the underlying processes and information sets. This distinction matters for how such data should be incorporated in choice models. For example, if the model includes an expectation over a variable that the individual has control over, then such expectations have to be jointly modeled with the main choice outcome.[13]

It is also worth noting that, while the link between expectations and behavior has been studied quite extensively in the context of human capital investments, fewer survey experiments in the macro area study impacts on actual behavior. This is largely due to data limitations and is gradually changing. In addition, evidence to date suggests that impacts of macro-related expectations on behaviors (stylized or actual) are generally modest. This could be due to several factors. First, macroeconomic expectations are likely not the first-order determinant for certain behaviors—for example, it is plausible that household spending responds primarily to financial constraints and personal income uncertainty rather than inflation expectations. Second, most papers in this area focus on inflation, which has been largely stable and fairly low in the US and other developed economies until recently. That may make households less sensitive to changes in inflation.[14] Third, individuals may simply be more responsive to factors that are somewhat in their own control. These explanations are largely speculative, and merit further investigation.

[13]For illustration, consider a model of university choice which incorporates the subjective likelihood of dropping out as one factor. Dropping out of a given school is likely going to be a function of some of the parameters that directly affect the utility of attending that school. Thus, the dropping out likelihood has to be jointly modeled with the choice of enrolling in that school; readers are referred to Delavande and Zafar (2019) for a discussion of this issue.

[14]Studies that do find effects in experimental or nonexperimental settings include Armantier et al. (2015), Bachmann et al. (2015), Coibion et al. (2021c), or Crump et al. (2021). See the chapter by D'Acunto et al. (2022) for further discussion.

Other applications. Information experiments are increasingly being used in other subfields, including labor economics (for example, to understand the role of information barriers in job search), public economics (for example, to understand drivers of support for various public policies including redistribution), health economics, political behavior, and so on. In addition, information experiments have also been used to explore the role of social norms or social pressure in affecting individual behavior (for example, Bursztyn and Jensen, 2015, and Bursztyn et al., 2020). Interested readers are referred to section 2 in the recent review article by Haaland et al. (2021), which provides a detailed overview of areas within economics where information experiments have been used, and to other chapters in this volume.

4 Methodological issues

This section discusses various methodological issues related to the design of information experiments.

4.1 Within-subject or between-subject design?

The identification argument laid out above in section 2.1 should make it clear that for recovering preference parameters, it is important to have a within-individual design. That is, a setup where expectations and choices are elicited from the *same* individual before and after the provision of information. On the other hand, a between-subject design—where expectations and choices are measured only after different groups have received some information, potentially with one control group not receiving any information—may suffice to shed some light on the expectation formation process, or simply study the treatment effect of a given piece of information on posterior beliefs or choices. However, a within-subject design is still more powerful since it allows the researcher to more extensively study heterogeneity in belief formation. More specifically, while a between-subject design allows one to investigate heterogeneity in belief updating based on observables (for example, by gender, one could simply compare the reported expectations for the treated and control group), it does not allow controlling for heterogeneity based on unobservables. The panel data that is collected through a within-subject design allows the researcher to include an individual fixed effect and, hence, enables the research to conduct a more detailed investigation of heterogeneity.

A within-subject design entails the elicitation of expectations twice, once before the information provision and once afterwards. A concern with eliciting the expectation using the same question twice is that responses may be biased due to an experimenter demand effect. That is, the respondent, upon seeing the same question again after the information provision, may be inclined to answer in a way that is consistent with the provided information only to please the researcher. This would bias the impacts upwards (relative to a between-subject design, where demand effects should, in general, be weaker). While this is hard to rule out entirely, existing evidence suggests that typical experimenter demand effects in surveys tend to be modest (De Quidt et al., 2018).[15] Conversely, respondents may want to appear consistent and not change their posterior due to this bias. This would attenuate the effects of the information downwards. Roth and Wohlfart (2020) find no significant difference in updating of expectations in a robustness check where respondents are randomized to a between- or within-subject design. More work undertaking such comparisons would be useful to determine whether and when one should be concerned about such biases.

Some researchers, in order to avoid asking the respondent the same question twice, instead elicit the prior and posterior expectation using different wordings. For example, Coibion et al. (2021c) use a within-subject design to study how different kinds of inflation-related information impact consumers' inflation expectations. They elicit the prior expectation as a subjective density forecast and the post-information expectation as a point estimate. While such an approach likely reduces consistency bias and survey fatigue, it implicitly assumes that *all* respondents report the same statistic of their subjective density forecast when asked for a point estimate. Unfortunately, this may not be a great assumption—Engelberg et al. (2009) show that there is a lot of variation across forecasters in what percentile of the subjective distribution the point estimate corresponds to. Thus, here again, researchers face potential trade-offs, and it is not clear (yet) what the best approach is.

Since within-subject designs (which are the most common approach to date in the

[15] It is worth noting that the best-known examples of demand effects come from social-psychology experiments in which students received class credits from participating in experiments and might have been informed about the psychological theories their instructors wanted to test. In the case of online surveys of the general population, it is much less plausible that subjects align their behavior with the hypotheses to be tested, since (i) they are less likely to correctly guess the hypotheses to be tested, (ii) they mostly just receive a fixed monetary payment for participation, and (iii) the experimenters are anonymous to them.

literature) rely on within-individual changes in expectations, the issue of measurement error is further exacerbated when using belief revisions to study impact on behaviors. Under classical measurement error, this would then yield a lower bound on the role of beliefs in explaining behavior. While an instrumental variables strategy may partially mitigate some of these concerns (as discussed in section 3.2), more work is needed to understand the nature and extent of measurement error—it is likely that measurement error in certain instances may be non-classical. In fact, the chapter in this volume by Manski (2022) flags the issue of rounding and imprecise probabilities as an important open research question.

4.2 Eliciting perceptions about the provided information

In order to study heterogeneity in belief updating, it is important that researchers elicit perceptions about the information they will eventually provide in the experiment before the information is given to the respondent. This is usually rather straightforward. Take, for example, the set-up in Wiswall and Zafar (2015a). They provide respondents with labor market information of current 30-year-olds with a college degree by major. Before they do so, they ask their respondents about their perceptions about these—this is what they refer to as "population beliefs", i.e., these are respondents' beliefs/perceptions about how much individuals in the population earn.[16]

This raises two issues. First, the mere act of asking people to think about a fact may prime them to factor that information into their beliefs and expectations, and may affect their subsequent behavior (Zwane et al., 2011; Crossley et al., 2017). Absent the survey, the respondent would not have acted in the same way. To bound such effects the researcher would need three groups: a pure control which is not surveyed, a control which is surveyed, and a treatment group. This, of course, imposes demands on the sample size and the researcher's budget. The extent to which one should be concerned about this issue likely depends on the application at hand. But, at the minimum, researchers should be aware of this potential bias and seriously think about how much to survey respondents.

The second issue is related to whether the elicitation of perceptions (or beliefs that can

[16] In other settings, eliciting priors about the information to be provided is not easily possible. For instance, in the Coibion et al. (2021c) paper discussed above, some treatments consist of newspaper articles or official statements. Also, to the extent that there are several possible signals, one would have to elicit every subjects' priors about all of these signals, which would be tedious.

be validated) should be incentivized. As a researcher, the goal is to recover the belief the respondent holds and uses when making decisions in the real-world. Incentives may increase the effort the respondent exerts in the survey, but it is far from clear that incentives increase the likelihood of recovering the true belief. Some researchers strongly argue for the necessity of incentives for belief elicitation and have provided some evidence of differences in beliefs elicited with and without incentives (Harrison and Phillips, 2014). On the other hand, Grewenig et al. (2020) show that incentivizing beliefs can impact the distribution of stated beliefs (relative to an unincentivized group). Their analysis of the data indicates that the incentive effects may be related to the usage of online search engines.[17] Providing incentives, as a result, leads to respondents having more accurate beliefs. But, in most instances, that is *not* the goal—instead the researcher is interested in recovering the true beliefs, whether biased or not, of the respondent. In line with this, Haaland et al. (2021) also conclude that incentives have little effect on beliefs in non-political domains and when responses cannot be readily looked up. Finally, it is worth noting that incentivized elicitation mechanisms can generate biases when respondents are not risk neutral (Murphy and Winkler, 1970). Even if respondents are risk neutral, incentivized belief elicitation techniques are not incentive-compatible when the respondent has a stake in the event that they are predicting (the "no stake" condition in Karni and Safra, 1995). Armantier and Treich (2013) also show that beliefs are less biased (but noisier) in the absence of incentives.

4.3 Eliciting higher-order moments

As mentioned in section 3.1, it is common to elicit the full subjective distribution from respondents. Not only does this allow the respondent to express uncertainty, it also provides a richer analysis of belief updating. For example, under Bayesian updating, respondents with more uncertain priors should be more responsive to the provided information.[18] Having information on baseline uncertainty allows the researcher to investigate this directly.

[17] They, in fact, conduct a treatment that explicitly encourages online-search activity without providing incentives. They find that this encouragement treatment produces effects that are similar to the incentive treatment, suggesting that altered online-search behavior is responsible for the impacts in the incentive treatment.

[18] Note that this is only true if prior uncertainty is not systematically correlated with other information-processing preferences/costs of the respondent. See Fuster et al. (2020b) for a detailed discussion.

Methodologically, there are multiple ways of eliciting subjective densities (see Bruine de Bruin et al., 2022, in this volume for further discussion). One way entails providing respondents with pre-determined bins, asking them to assign probabilities (on a 0-100 scale) to the bins such that they sum to 100. The density questions in the NY Fed's Survey of Consumer Expectations (SCE), for example, use this approach. Alternatively, since respondents are typically first asked for a point estimate in many surveys, one could make the bins individual-specific, by anchoring the variation around the individual's point estimate. For example, this approach is used by Armona et al. (2019), who provide bins that are anchored to the respondent's baseline expectation.

A potential shortcoming with the former approach is that the point estimate may lie outside the support of the bins which are presented to the respondent, and hence the researcher will not have quantitative measures of respondent uncertainty in such a case. The SCE provides bins that go from -12 to +12 percent for year-ahead inflation and home price changes. Using the individual-level publicly available SCE data since 2020, we find that the point forecasts for inflation and housing price changes are outside this interval for more than 15% percent of the respondents.

An alternative approach has been to use a qualitative measure of uncertainty, where respondents are asked for the certainty in their response on a Likert-scale. While there are well-known issues with the interpretation of such data, especially for comparisons across individuals and groups, such data are easier to elicit, and may still be useful to document descriptive patterns of heterogeneity in belief-updating.

4.4 Information content and presentation

While information provision experiments routinely vary the quantitative content of the information that is provided—indeed, that is usually the purpose of the randomization—not very much is known about how other "dimensions" of the provided information matter. For instance, does it matter who provides the information, meaning what source for a given piece of information is shown to respondents? There is some work that shows that the identity of the messenger has an impact on how different groups of respondents incorporate the same information, and what information they choose (D'Acunto et al., 2021). Likewise, different ways of communicating similar information can have differential impact on expectations: for example, Coibion et al. (2021c) find that an intervention that provides respondents with a news article about the most recent Federal Open Market Committee (FOMC) meeting leads to an impact on expectations that is only half as strong as reading the FOMC statement itself. These two pieces of information differ in terms of their technicality, content, and credibility, and so it is difficult to fully disentangle what leads to these differential impacts. For example, to assess the role of the source, the ideal experiment would entail providing the same identical information but with different sources.[19] That, of course, is generally not feasible.

Related to this, there is so far limited evidence on how credible and relevant respondents find the information to be, and to our knowledge there have not been studies that exogenously vary the credibility of the information source to study how that affects the extent to which it is incorporated into subjective beliefs.[20]

Researchers also have to decide how complex and detailed the provided information is. There is a trade-off: one wants the information to be as comprehensible and simple as possible for the respondent, without it being misleading. For example, most experiments give respondents information about the average of a variable (for example, average earnings of college graduates in the population, or the average inflation forecast of experts). But of course, these are just averages and there is a confidence interval around these estimates. Should the researcher just provide a point estimate or convey some sense of the distribution (or the uncertainty in the estimate)? While the former is easier to communicate, there is a risk that doing so may inadvertently make the respondent think of the outcome as being deterministic.

Finally, one may ask whether respondents could be provided with false (made-up) information, either out of convenience (in order to get more variation in the provided information) or to study whether people are able to distinguish false from true information. However, it is important to keep in mind that in experimental economics, there is a strong norm against using deception (although there are different opinions on what constitutes deception; see Charness et al., 2021 for discussion). Thus, we advise to only use factual information.

In short, more systematic work is needed to understand the most effective way of pre-

[19] Coibion et al. (2021c) show that respondents tend to assign lower credibility to newspapers as a source of information about the economy than e.g. to the government or social media, making it likely that differential credibility played a role in their results.

[20] Similarly, to our knowledge it is an open question whether information provided by one person (one forecaster) versus (the average from) a group of forecasters has differential effects.

senting information. We suspect the answer will depend on the application and context.

4.5 Where and how to run these surveys?

Most of the studies that use information provision experiments are run online. The most commonly used platform to run surveys is Qualtrics, although some survey panels have separate dedicated platforms. Aside from the platform on which the survey is run, one then needs access to a subject pool. Haaland et al. (2021) provide an overview of different types of samples: (i) probability-based samples (e.g., RAND American Life Panel); (ii) representative online samples (based on certain observable characteristics; e.g., Dynata, Lucid, Prolific, or Qualtrics); or (iii) online labor markets (Amazon MTurk). The representativeness declines across these options, but so does the cost.

Especially for relatively low-cost options like Amazon MTurk, it is important to have a few attention checks/screens in the survey, as the data can otherwise be quite noisy.[21] However, one has to be careful not to exaggerate the number of such checks, as this could otherwise introduce selection (with experienced survey takers being better at spotting such checks) or alienate respondents. Reading comprehension checks, recall questions or open-ended questions can also be useful.

A very useful reference on such implementation issues is Bergman et al. (2020). They highlight for instance the need to think carefully about question wording and survey length, and the best practice of running a small-scale pilot survey before launching the full study, in order to detect any potential issues that would compromise the usefulness of the data.[22] They also provide example Qualtrics files.

Some central banks and other public institutions maintain high-quality probability-based samples (often with a panel dimension), which have been used for many of the studies discussed in this chapter. To the extent possible, it is desirable for such institutions to allow outside researchers to add their own questions to their surveys, e.g. via a "call for proposals" as was done by the Bundesbank when launching their online pilot survey on consumer expectations (see Beckmann and Schmidt, 2020).

[21] In some cases it is easy to spot noisy responses—e.g. if a respondent always chooses the first option in multiple choice answers or is extremely fast in completing the survey—but it is also possible that some respondents effectively respond "randomly."

[22] A more powerful but also expensive type of pilot is to have a few respondents recorded as they take the survey and "think aloud" as they go through each question.

5 Extensions and alternative approaches

5.1 Moving beyond exogenously provided information

There are two important limitations to the information provision experiments discussed in section 3. The first one is that the signals provided to respondents are picked by the researcher, while of course in reality there is a very large number of signals that an economic agent could consult when forming expectations about a given variable.

The second limitation is that signals are simply "dropped" exogenously on survey respondents, often in a way that makes the signals quite salient and ensures that respondents are paying attention to them. However, in their everyday lives, respondents would themselves choose whether to pay attention to any signal, which they generally have to expend some effort to obtain (e.g. by searching for information online), and they would choose which of the many possible signals to look for.

A few recent studies have extended the survey-experimental paradigm to better understand the information acquisition decision. While the set of available signals is still chosen ex-ante by the researchers, this allows studying which signals respondents with different characteristics or priors tend to prefer, and how much they value the information.

Fuster et al. (2020b) study information acquisition within the context of home price expectations. Their experimental design has four stages. First, prior expectations about the future national median home price are elicited. Second, in a later part of the survey, respondents are informed that they will again be asked for their forecast, but that (i) they can now win a monetary reward if their forecast turns out to be accurate;[23] and (ii) that they have the option to choose among different pieces of information that they could consult and that could be useful for their forecast. Respondents were then asked to express their preference over the following four signals: the national home price change over the past 1 year, the change over the past 10 years, the average expert forecast of future home prices, or seeing no information at all. Third, respondents' willingness to pay for their most preferred signal is elicited (unless they stated they preferred not seeing any information), using an incentive-compatible "multiple price list" method. Then, depending

[23] This reward was randomized to be either \$10 or \$100 (in each case with probability 10% of being eligible for the reward), in order to study whether incentives matter for the choice, valuation and use of information.

on the willingness to pay and randomness, some respondents are shown their preferred piece of information. Finally, in the fourth stage, posterior expectations are elicited.

Thus, the first and last stage are as in a standard information-provision experiment, but in between, additional information is collected that allows to study respondent preferences across signals as well as the valuation of information. Results indicate substantial heterogeneity in respondents' rankings of the available signals, and more sophisticated respondents (measured by education or numeracy) are more likely select the arguably more informative signals (the expert forecast or past 1-year growth). Respondents display high valuation for the signals (when compared to the possible prizes from a successful forecast) and incorporate the information, if they receive it, into their beliefs.

The heterogeneity in preferences across signals has an important implication: as the price of information is lowered, the cross-sectional dispersion in posterior beliefs does not necessarily decrease. This is because a decrease in dispersion *within* a group that sees a certain signal—which is a typical result obtained in standard information provision experiments—is offset by an increase in dispersion *across* groups that chose different signals.[24] This second channel would be missed by standard information provision experiments.

Roth et al. (2021) also focus on information acquisition. This study tests whether people become more likely to select a given piece of information (in this case, a professional forecast about future recession risk) when their (perceived) personal benefit from the information is higher. To manipulate this personal benefit, the authors use variation in the unemployment rates from a past recession measured across different official statistics. The main result is that indeed, exogenously increased perceived exposure to macroeconomic risk leads respondents to become more likely to want to be informed about recession risk (relative to other macro variables such as inflation).

Other recent studies that consider endogenous information acquisition within surveys are Faia et al. (2021), who let respondents choose between different newspaper headlines about topics related to the COVID-19 pandemic, and D'Acunto et al. (2021), who study how the choice of newspaper articles about the Federal Reserve is affected by identity of the Fed policymaker that is featured in the article, and how these effects differ across demographic groups. Capozza et al. (2021) review the broader literature on information

[24]This result continues to hold in an extension of the survey experiment where respondents can obtain more than one signal, which of course would also be the case in reality.

acquisition in field settings.

5.2 Alternatives to information provision experiments

An attractive feature of information experiments is that they allow the researcher to generate shocks to an individual's information set in order to study resulting effects on expectations and behavior. An alternative approach to generate within-individual variation in the economic environment is to elicit *conditional* expectations and/or planned behaviors in so-called vignettes, i.e., hypothetical scenarios.

To illustrate the usefulness of such data, let us go back to the example of college major choice from section 2.1. The ideas presented here are a condensed version of those in Wiswall and Zafar (2021). Consider an individual i who at time τ is deciding which college major to choose. i's perception of the expected utility from major k is given by:

$$E_{i,\tau}(V_k) = \sum_{t=\tau+1}^{T} \beta^{t-\tau} \int u_t(X) \, dG_{i,\tau}(X|k,t),$$

where $\beta \in (0,1)$ is the discount rate, $u(X)$ is the post-graduation utility function that provides the mapping from the finite vector of post-graduation events X to utility, and $G_{i,\tau}(X|k,t)$ is the individual's *beliefs* at period τ about the probability of the vector of future outcomes X occurring in all future periods $t > \tau$ *if* she were to complete major k. Given her beliefs at period τ, individual i chooses the major that provides the highest expected present discounted utility.

Surveys that collect data on subjective expectations directly elicit individual's beliefs $G_{i,\tau}(X|k,t)$. For example, Wiswall and Zafar (2021) ask each individual what they believe the distribution of X would be in some future period t *if* they were to complete different majors. That is, students are asked their expectations about post-graduation events *conditional* on different majors. In their case, X includes a wide range of events such as earnings, labor supply, marriage, and spousal earnings.

From the collection of self beliefs data, one can analyze *individual-level* differences in belief distributions:

$$\Delta_{G,i\tau}(k,k') = G_{i,\tau}(X|k,t) - G_{i,\tau}(X|k',t),$$

where $\Delta_{G,i\tau}(k, k')$ can be viewed as the *ex ante* (i.e., prior to the choice of major) "treatment effect" on the distribution of future events X. Specifically, it reflects the "causal effect" individual i expects if she chooses major k rather than k'. As an example, consider beliefs about future earnings at age $t = 30$, where earnings w would be one element of the X vector, $w \subset X$. By eliciting each individual's beliefs about their expected future earnings at age $t = 30$ *if* they were to complete major k, and if they were not to complete a major k', one can compute the difference in the individual's point forecast of expected earnings:

$$\delta_i = E_i(w|k, 30) - E_i(w|k', 30).$$

Wiswall and Zafar (2021) extend this idea to many other potential outcomes beyond expected earnings, including earnings uncertainty, labor supply, marriage, fertility, spousal characteristics, and labor supply. Arcidiacono et al. (2020) use the same approach to study the causal impact of majors on perceived returns to occupations. Similarly, by collecting data on conditional probabilities of work given health, Giustinelli and Shapiro (2019) estimate the subjective ex ante treatment effect of health on retirement.[25] The chapters in this volume by Kézdi and Shapiro (2022) and Kosar and O'Dea (2022) provide additional discussion and references.

The studies above measure both conditional expectations and actions in a given hypothetical state of the world, but it is also possible to measure only expectations or only actions. An example of the former in the context of macroeconomics is Andre et al. (2021). These authors present respondents (both a representative sample from the general population and a sample of experts) with vignettes in which respondents are asked to predict future unemployment and inflation under different hypothetical macroeconomic shocks, for instance about the monetary policy rate or government spending.[26] The resulting data allow them to study heterogeneity in respondents' "subjective models" of how the economy responds to such shocks, and the source of such heterogeneity (which they link to

[25]Note that the concept of ex-ante treatment effects is distinct from ex-post treatment effects of choices. In the case of ex-post treatment effects, the alternative outcomes are counterfactual and unobserved. A large econometrics and statistics literature studies how to identify these potential outcomes and moments of the potential outcomes (such as average treatment effects) from realized choice data (Heckman and Vytlacil, 2005).

[26]They are explicitly providing information about the assumed source of the shocks, in order to make respondents think of them as truly "exogenous" (as opposed to an endogenous response to other economic developments).

personal experiences).[27]

Other work uses vignettes to study conditional intended behaviors only, without measuring conditional expectations.[28] For instance, in Fuster et al. (2020a), respondents are presented with scenarios that provide hypothetical one-time income gains or losses of different sizes. Respondents are then asked for their propensity to spend. This allows the authors to estimate, for example, how the size of a one-time cash transfer causally impacts the intended marginal propensity to spend, an object that is of policy relevance (see also Jappelli and Pistaferri, 2014; Christelis et al., 2019). In a different context, Fuster and Zafar (2021) use vignettes to study survey respondents' willingness-to-pay for a home under different configurations of financing conditions (interest rate, downpayment requirement).

The elicitation of conditional expectations hinges on two implicit assumptions. First, that individuals have well-formed expectations about the outcomes conditional on the state of the world that they are being asked about. This is not directly testable and is really at the discretion at the researchers. For example, it is hard to argue that this assumption is not satisfied in the case of asking college students about outcomes conditional on college major—a decision that the students are actively thinking about. Second, that there is no systematic bias in the reporting of expectations. This is an assumption that is implicitly made when using any survey data, and is not specific to (conditional) expectations data.

When eliciting choices in hypothetical scenarios, likewise, an implicit assumption is that stated choices are reflective of what respondents would do in actual scenarios. Although there has historically been concern about the plausibility of this assumption (Diamond and Hausman, 1994), there is growing evidence that the two approaches of using stated choices or actual choices yield similar preference estimates (see Fuster et al., 2020a, and references therein), and that the stated approach yields meaningful responses when the counterfactual scenarios presented to respondents are realistic and relevant for them.

[27]Other examples of conditional expectation elicitation in macro are Coibion et al. (2020b), where respondents are asked about expected paths for the economy depending on the outcome of the presidential election, and Dibiasi et al. (2021), where firm managers are asked how their expected investment, employment and production would change in the face of a (mean zero) uncertainty shock, in this case linked to a popular vote.

[28]There is also work that elicits intended behaviors in different conditional states, but expectations in only one state of the world. This then requires the researcher to also model how expectations may change in counterfactual states. For example, Delavande and Zafar (2019) use such data to estimate the causal impact of credit constraints and pecuniary and non-pecuniary factors on the choice of type of higher education institution.

A challenge with the vignette approach is that the researcher has to decide how much information to include in each scenario. Should the scenarios be fully specified or incomplete? The results may depend on these decisions because respondents may "fill the blanks" for unspecified features. The answer really depends on what the goal is. If the goal is to use the vignette to estimate a specific model, then fully specifying the model assumptions (obviously, in a comprehensible way) makes sense. This, for example, is the approach used by Ameriks et al. (2020) to estimate preferences for motives for late-in-life saving in a lifecycle model with incomplete markets. Such an approach could make the vignettes cognitively demanding for respondents, who may perceive the questions to be artificial, but allows the researcher to nicely map the vignette to the model framework.

In short, vignettes appear to be a promising method to obtain richer information than is possible with information provision alone.

6 Directions for future work

The arguments in this chapter have hopefully made it clear that survey experiments are a powerful tool to help researchers understand decision-making and belief formation. Empirical evidence overwhelmingly suggests that, across domains, individuals are quite misinformed and that simple information provision is often effective at nudging expectations and, in some cases, behaviors.

We end by providing a discussion of possible directions for future work. First, the fact that information experiments—which largely provide information that is readily available in the public domain—shift individuals' expectations is rather puzzling. That individuals respond to such information implies that they find such information useful. Then *why* are individuals choosing to stay misinformed? Several models in the literature have attempted to explain this, either based on barriers to information selection or information acquisition (see references in Fuster et al., 2020b). With the exception of a few empirical studies mentioned in section 5.1, most information experiments to date have little to tell us about heterogeneity in information selection and information acquisition. Instead, they primarily focus on heterogeneity in information processing. A natural extension of prevalent information experiments is to bring in these dimensions.

Second, as alluded to earlier in section 3.3, the literature that investigates the link between macroeconomic expectations and behavior has mostly relied on elicited intended behaviors or stylized behaviors within incentivized decision situations. Only in rare cases do researchers have access to actual behaviors as reported in follow-up surveys or direct measures of behavior (for instance in administrative data) from the same individuals from whom expectations data are elicited. With increasing ease of collecting data and ability to link to administrative data (on behaviors), we expect this data limitation to relax over time. The few papers in the macro space that have such rich data typically find modest impacts of expectations on behavior. Work that sheds light on why that may be the case would be useful.

Third, information experiments have been quite useful in documenting heterogeneity in updating rules or mental models. Going forward, it will be useful for theory and experimental evidence to evolve interactively—with new theories being proposed based on this heterogeneity, and then being tested again more systematically in new settings.

Fourth, beyond heterogeneity based on observables, there is also evidence that individuals process information in a biased manner, especially when updating over variables that they have control over or that can impact utility directly, such as ability (see Benjamin, 2019 for an overview of biases in the belief-updating literature). In the context of ability updating, the tendency of people to be conservative in updating is fairly common. On the other hand, evidence on whether individuals tend to update more in response to "good" news versus "bad news"—a bias referred to as "asymmetric updating" (Mobius et al., 2014) or the "good news-bad news effect" (Eil and Rao, 2011)—is quite mixed (see discussion in Benjamin, 2019). In addition, we know even less about the dynamics of motivated beliefs (Zimmermann, 2020). More work is needed to better understand these biases and the circumstances in which particular biases may emerge. Much of the existing evidence comes from lab studies, but incorporating the analysis of such motivated-belief biases in the context of larger-scale survey experiments also seems a fruitful avenue for future research to us.

Finally, as mentioned in section 4.4, there has been little systematic work on how to design the content and presentation of information interventions. While we do not expect a unified theory of how information should be presented and communicated, and the answer likely depends on the context and application at hand, work that sheds further light on these issues would certainly be useful.

References

AMERIKS, J., J. BRIGGS, A. CAPLIN, M. D. SHAPIRO, AND C. TONETTI (2020): "Long-Term-Care Utility and Late-in-Life Saving," *Journal of Political Economy*, 128, 2375–2451.

ANDRE, P., C. PIZZINELLI, C. ROTH, AND J. WOHLFART (2021): "Subjective Models of the Macroeconomy: Evidence from Experts and Representative Samples," *Review of Economic Studies*, forthcoming.

ARCIDIACONO, P., V. J. HOTZ, A. MAUREL, AND T. ROMANO (2020): "Ex Ante Returns and Occupational Choice," *Journal of Political Economy*, 128, 4475–4522.

ARMANTIER, O., W. BRUINE DE BRUIN, G. TOPA, W. KLAAUW, AND B. ZAFAR (2015): "Inflation expectations and behavior: Do survey respondents act on their beliefs?" *International Economic Review*, 56, 505–536.

ARMANTIER, O., S. NELSON, G. TOPA, W. VAN DER KLAAUW, AND B. ZAFAR (2016): "The Price is Right: Updating Inflation Expectations in a Randomized Price Information Experiment," *Review of Economics and Statistics*, 98, 503–523.

ARMANTIER, O. AND N. TREICH (2013): "Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging," *European Economic Review*, 62, 17–40.

ARMONA, L., A. FUSTER, AND B. ZAFAR (2019): "Home price expectations and behaviour: Evidence from a randomized information experiment," *Review of Economic Studies*, 86, 1371–1410.

BACHMANN, R., T. O. BERG, AND E. R. SIMS (2015): "Inflation expectations and readiness to spend: cross-sectional evidence," *American Economic Journal: Economic Policy*, 7, 1–35.

BECKMANN, E. AND T. SCHMIDT (2020): "Bundesbank online pilot survey on consumer expectations," Technical Paper 01/2020, Deutsche Bundesbank.

BENHABIB, J. AND M. M. SPIEGEL (2018): "Sentiments and Economic Activity: Evidence from US States," *The Economic Journal*, 129, 715–733.

BENJAMIN, D. J. (2019): "Chapter 2 - Errors in probabilistic reasoning and judgment biases," in *Handbook of Behavioral Economics - Foundations and Applications*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, North-Holland, vol. 2, 69–186.

BERGMAN, A., A. CHINCO, S. M. HARTZMARK, AND A. B. SUSSMAN (2020): "Survey Curious? Start-Up Guide and Best Practices For Running Surveys and Experiments Online," Working Paper, University of Chicago - Booth School of Business.

BERNANKE, B. S. (2007): "Inflation Expectations and Inflation Forecasting," Speech at the Monetary Economics Workshop of the NBER Summer Institute, Cambridge, MA.

BEUTEL, J., N. METIU, AND V. STOCKERL (2021): "Toothless tiger with claws? Financial stability communication, expectations, and risk-taking," *Journal of Monetary Economics*, 120, 53–69.

BINDER, C. AND A. RODRIGUE (2018): "Household informedness and long-run inflation expectations: Experimental evidence," *Southern Economic Journal*, 85, 580–598.

BLEEMER, Z. AND B. ZAFAR (2018): "Intended college attendance: Evidence from an experiment on college returns and costs," *Journal of Public Economics*, 157, 184–211.

BOBBA, M. AND V. FRISANCHO (2016): "Perceived Ability and School Choices," TSE Working Papers 16-660, Toulouse School of Economics (TSE).

BOTTAN, N. L. AND R. PEREZ-TRUGLIA (2020): "Betting on the House: Subjective Expectations and Market Choices," Working Paper 27412, National Bureau of Economic Research.

BRUINE DE BRUIN, W., A. CHIN, J. DOMINITZ, AND W. VAN DER KLAAUW (2022): "Household surveys and probabilistic questions," in *Handbook of Economic Expectations*, chap. 1.

BURSZTYN, L., A. L. GONZÁLEZ, AND D. YANAGIZAWA-DROTT (2020): "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American Economic Review*, 110, 2997–3029.

BURSZTYN, L. AND R. JENSEN (2015): "How Does Peer Pressure Affect Educational Investments?" *The Quarterly Journal of Economics*, 130, 1329–1367.

CANDIA, B., O. COIBION, AND Y. GORODNICHENKO (2022): "The Macroeconomic Expectations of Firms," in *Handbook of Economic Expectations*, chap. 14.

CAPOZZA, F., I. HAALAND, C. ROTH, AND J. WOHLFART (2021): "Studying Information Acquisition in the Field: A Practical Guide and Review," CEBI working paper series 21-15, University of Copenhagen.

CAVALLO, A., G. CRUCES, AND R. PEREZ-TRUGLIA (2017): "Inflation Expectations, Learning and Supermarket Prices: Evidence from Survey Experiments," *American Economic Journal: Macroeconomics*, 9, 1–35.

CHARNESS, G., A. SAMEK, AND J. VAN DE VEN (2021): "What is considered deception in experimental economics?" *Experimental Economics*.

CHRISTELIS, D., D. GEORGARAKOS, T. JAPPELLI, L. PISTAFERRI, AND M. VAN ROOIJ (2019): "Asymmetric Consumption Effects of Transitory Income Shocks," *Economic Journal*, 129, 2322–2341.

COIBION, O., D. GEORGARAKOS, Y. GORODNICHENKO, AND M. VAN ROOIJ (2021a): "How Does Consumption Respond to News about Inflation? Field Evidence from a Randomized Control Trial," Working Paper.

COIBION, O., D. GEORGARAKOS, Y. GORODNICHENKO, AND M. WEBER (2020a): "Forward Guidance and Household Expectations," Working Paper 26778, National Bureau of Economic Research.

COIBION, O., Y. GORODNICHENKO, AND S. KUMAR (2018): "How do firms form their expectations? New survey evidence," *American Economic Review*, 108, 2671–2713.

COIBION, O., Y. GORODNICHENKO, S. KUMAR, AND J. RYNGAERT (2021b): "Do You Know that I Know that You Know...? Higher-Order Beliefs in Survey Data," *Quarterly Journal of Economics*, 136, 1387–1446.

COIBION, O., Y. GORODNICHENKO, AND T. ROPELE (2019): "Inflation Expectations and Firm Decisions: New Causal Evidence," *Quarterly Journal of Economics*, 135, 165–219.

COIBION, O., Y. GORODNICHENKO, AND M. WEBER (2020b): "Political Polarization and Expected Economic Outcomes," Working Paper 28044, National Bureau of Economic Research.

——— (2021c): "Monetary Policy Communications and Their Effects on Household Inflation Expectations," *Journal of Political Economy*, forthcoming.

CONLON, J. AND B. ZAFAR (2017): "Measuring Americans' Expectations Following the 2016 Election," Federal Reserve Bank of New York Liberty Street Economics Blog.

CONLON, J. J. (2019): "Major Malfunction: A Field Experiment Correcting Undergraduates' Beliefs about Salaries," *Journal of Human Resources*, 0317–8599R2.

CROSSLEY, T. F., J. DE BRESSER, L. DELANEY, AND J. WINTER (2017): "Can Survey Participation Alter Household Saving Behaviour?" *The Economic Journal*, 127, 2332–2357.

CRUMP, R. K., S. EUSEPI, A. TAMBALOTTI, AND G. TOPA (2021): "Subjective intertemporal substitution," *Journal of Monetary Economics*, forthcoming.

D'ACUNTO, F., A. FUSTER, AND M. WEBER (2021): "Diverse Policy Committees Can Reach Underrepresented Groups," Working Paper 2021-95, University of Chicago Becker-Friedman Institute.

D'ACUNTO, F., U. MALMENDIER, AND M. WEBER (2022): "What Do the Data Tell Us About Inflation Expectations?" in *Handbook of Economic Expectations*, chap. 7.

DE QUIDT, J., J. HAUSHOFER, AND C. ROTH (2018): "Measuring and bounding experimenter demand," *American Economic Review*, 108, 3266–3302.

DE STEFANI, A. (2021): "House Price History,Biased Expectations and Credit Cycles: the Role of Housing Investors," *Real Estate Economics*, forthcoming.

DELAVANDE, A. (2022): "Expectations in Development," in *Handbook of Economic Expectations*, chap. 11.

DELAVANDE, A. AND B. ZAFAR (2019): "University Choice: The Role of Expected Earnings, Nonpecuniary Outcomes, and Financial Constraints," *Journal of Political Economy*, 127, 2343–2393.

DIAMOND, P. A. AND J. A. HAUSMAN (1994): "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives*, 8, 45–64.

DIBIASI, A., H. MIKOSCH, AND S. SARFERAZ (2021): "Uncertainty Shocks, Adjustment Costs and Firm Beliefs: Evidence From a Representative Survey," Working Paper, KOF, ETH Zurich.

DIZON-ROSS, R. (2019): "Parents' Beliefs about Their Children's Academic Ability: Implications for Educational Investments," *American Economic Review*, 109, 2728–65.

EIL, D. AND J. M. RAO (2011): "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself," *American Economic Journal: Microeconomics*, 3, 114–38.

ENGELBERG, J., C. F. MANSKI, AND J. WILLIAMS (2009): "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters," *Journal of Business & Economic Statistics*, 27, 30–41.

FAIA, E., A. FUSTER, V. PEZONE, AND B. ZAFAR (2021): "Biases in Information Selection and Processing: Survey Evidence from the Pandemic," Working Paper 28484, National Bureau of Economic Research.

FRANCO, C. (2019): "How does relative performance feedback affect beliefs and academic decisions?" Working paper.

FREEMAN, R. (1971): *The Market for College Trained Manpower*, Harvard University Press, named one of "Outstanding Books in Industrial Relations and Labor Economics, 1970-1979." Chapter 6 reprinted in Labor Market Analysis. ed by J. Burton, L. Benham, W. Vaughn, and R. Flanagan, 1971.

FUSTER, A., G. KAPLAN, AND B. ZAFAR (2020a): "What Would You Do with $500? Spending Responses to Gains, Losses, News, and Loans," *The Review of Economic Studies*, 88, 1760–1795.

FUSTER, A., R. PEREZ-TRUGLIA, M. WIEDERHOLT, AND B. ZAFAR (2020b): "Expectations with Endogenous Information Acquisition: An Experimental Investigation," *Review of Economics and Statistics*, forthcoming.

FUSTER, A. AND B. ZAFAR (2021): "The Sensitivity of Housing Demand to Financing Conditions: Evidence from a Survey," *American Economic Journal: Economic Policy*, 13, 231–65.

GALASHIN, M., M. KANZ, AND R. PEREZ-TRUGLIA (2020): "Macroeconomic Expectations and Credit Card Spending," Working Paper 28281, National Bureau of Economic Research.

GIUSTINELLI, P. (2022): "Expectations in Education," in *Handbook of Economic Expectations*, chap. 9.

GIUSTINELLI, P. AND M. D. SHAPIRO (2019): "SeaTE: Subjective ex ante Treatment Effect of Health on Retirement," Working Paper 26087, National Bureau of Economic Research.

GLASER, M., T. LANGER, J. REYNDERS, AND M. WEBER (2007): "Framing effects in stock market forecasts: The difference between asking for prices and asking for returns," *Review of Finance*, 11, 325–357.

GREWENIG, E., P. LERGETPORER, K. WERNER, AND L. WOESSMANN (2020): "Incentives, search engines, and the elicitation of subjective beliefs: Evidence from representative online survey experiments," *Journal of Econometrics*, forthcoming.

HAALAND, I., C. ROTH, AND J. WOHLFART (2021): "Designing Information Provision Experiments," *Journal of Economic Literature*, forthcoming.

HANSPAL, T., A. WEBER, AND J. WOHLFART (2020): "Exposure to the COVID-19 Stock Market Crash and Its Effect on Household Expectations," *The Review of Economics and Statistics*, forthcoming.

HARRISON, G. W. AND R. D. PHILLIPS (2014): *Subjective Beliefs and Statistical Forecasts of Financial Risks: The Chief Risk Officer Project*, London: Palgrave Macmillan UK, 163–202.

HASTINGS, J., C. A. NEILSON, AND S. D. ZIMMERMAN (2015): "The effects of earnings disclosure on college enrollment decisions," Working Paper 21300, National Bureau of Economic Research.

HECKMAN, J. J. AND E. VYTLACIL (2005): "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73, 669–738.

JAPPELLI, T. AND L. PISTAFERRI (2014): "Fiscal Policy and MPC Heterogeneity," *American Economic Journal: Macroeconomics*, 6, 107–136.

JENSEN, R. (2010): "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 125, 515–548.

KARNI, E. AND Z. SAFRA (1995): "The impossibility of experimental elicitation of subjective probabilities," *Theory and Decision*, 38, 313–320.

KÉZDI, G. AND M. D. SHAPIRO (2022): "Retirement Expectations," in *Handbook of Economic Expectations*, chap. 12.

KOSAR, G. AND C. O'DEA (2022): "Expectations Data in Structural Microeconomic Models," in *Handbook of Economic Expectations*, chap. 19.

KUCHLER, T., M. PIAZZESI, AND J. STROEBEL (2022): "Housing Market Expectations," in *Handbook of Economic Expectations*, chap. 8.

LAUDENBACH, C., A. WEBER, AND J. WOHLFART (2021): "Beliefs About the Stock Market and Investment Choices: Evidence from a Field Experiment," Working Paper.

LEWIS, D. J., C. MAKRIDIS, AND K. MERTENS (2019): "Do Monetary Policy Announcements Shift Household Expectations?" Working Papers 1906, Federal Reserve Bank of Dallas.

LINK, S., A. PEICHL, C. ROTH, AND JOHANNESWOHLFART (2021): "Information Frictions among Firms and Households," Working paper.

MANSKI, C. (2022): "Looking Ahead to Research Enhancing Measurement of Expectations," in *Handbook of Economic Expectations*, chap. 26.

MANSKI, C. F. (2004): "Measuring expectations," *Econometrica*, 72, 1329–1376.

MIAN, A., A. SUFI, AND N. KHOSHKHOU (2015): "Government Economic Policy, Sentiments, and Consumption," Working Paper 21316, National Bureau of Economic Research.

MOBIUS, M., M. NIEDERLE, P. NIEHAUS, AND T. ROSENBLAT (2014): "Managing Self-Confidence," Working paper.

MURPHY, A. H. AND R. L. WINKLER (1970): "Scoring rules in probability assessment and evaluation," *Acta Psychologica*, 34, 273–286.

QIAN, W. (2020): "House Price Expectations and Consumption – A Survey-based Experiment," Working Paper, University of Notre Dame.

ROTH, C., S. SETTELE, AND J. WOHLFART (2021): "Risk Exposure and Acquisition of Macroeconomic Information," *American Economic Review: Insights*, forthcoming.

ROTH, C. AND J. WOHLFART (2020): "How do expectations about the macroeconomy affect personal expectations and behavior?" *Review of Economics and Statistics*, 102, 731–748.

Wiswall, M. and B. Zafar (2015a): "Determinants of College Major Choice: Identification using an Information Experiment," *The Review of Economic Studies*, 82, 791–824.

——— (2015b): "How do college students respond to public information about earnings?" *Journal of Human Capital*, 9, 117–169.

——— (2021): "Human Capital Investments and Expectations about Career and Family," *Journal of Political Economy*, 129, 1361–1424.

Zimmermann, F. (2020): "The Dynamics of Motivated Beliefs," *American Economic Review*, 110, 337–61.

Zwane, A. P., J. Zinman, E. Van Dusen, W. Pariente, C. Null, E. Miguel, M. Kremer, D. S. Karlan, R. Hornbeck, X. Giné, et al. (2011): "Being surveyed can change later behavior and related parameter estimates," *Proceedings of the National Academy of Sciences*, 108, 1821–1826.