A Collaborative Research Project Funded by:

DFID Department for International Development

Implemented by:

HPAI PRO-POOR HPAI RISK REDUCTION
HIGHLY PATHOGENIC AVIAN INFLUENZA

Behavioral Field Experiments for the Study of HPAI Prevention and Control

Angelino Viceisza

Africa/Indonesia Team Working Paper No. 12

Table of Contents

List of Figure

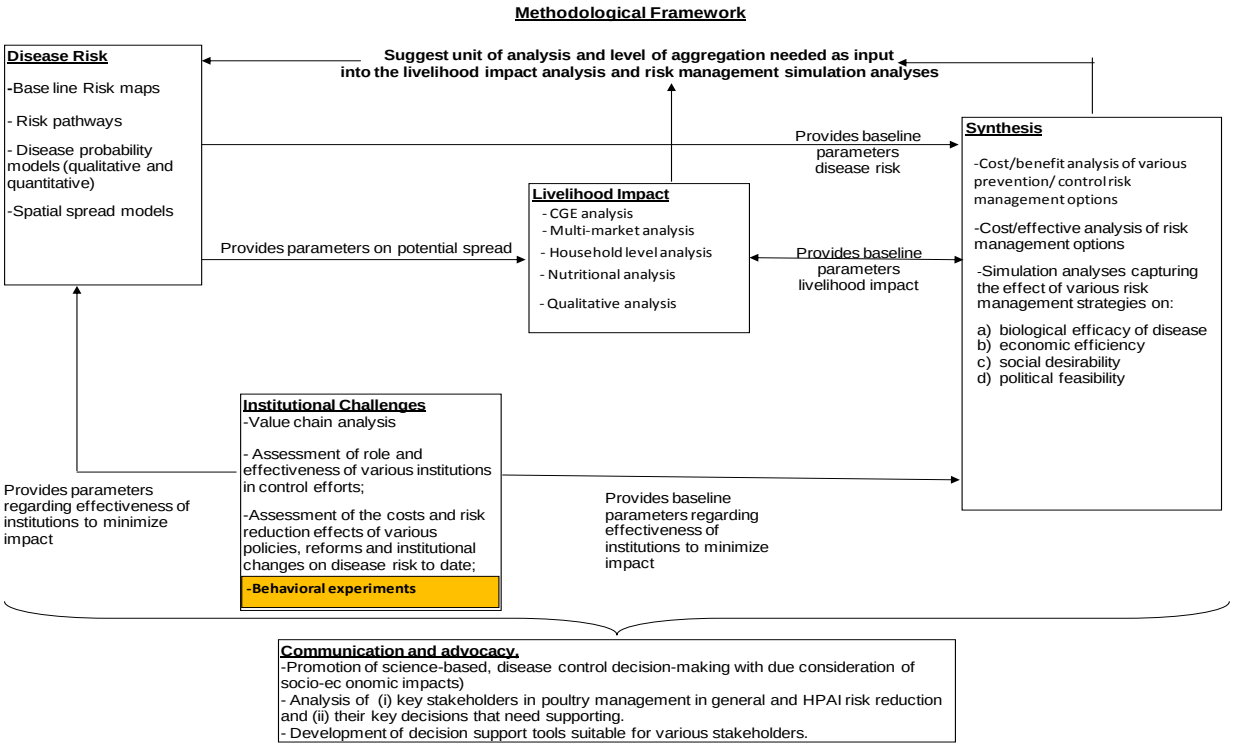
Preface

Since its re-emergence, HPAI H5N1 has attracted considerable public and media attention because the viruses involved have been shown to be capable of producing fatal disease in humans. While there is fear that the virus may mutate into a strain capable of sustained human-to-human transmission, the greatest impact to date has been on the highly diverse poultry industries in affected countries. In response to this, HPAI control measures have so far focused on implementing prevention and eradication measures in poultry populations, with more than 175 million birds culled in Southeast Asia alone.

Until now, significantly less emphasis has been placed on assessing the efficacy of risk reduction measures, including their effects on the livelihoods of smallholder farmers and their families. In order to improve local and global capacity for evidence-based decision making on the control of HPAI (and other diseases with epidemic potential), which inevitably has major social and economic impacts, the UK Department for International Development (DFID) has agreed to fund a collaborative, multidisciplinary HPAI research project for Southeast Asia and Africa.

The specific purpose of the project is to aid decision makers in developing evidence-based, pro-poor HPAI control measures at national and international levels. These control measures should not only be cost-effective and efficient in reducing disease risk, but also protect and enhance livelihoods, particularly those of smallholder producers in developing countries, who are and will remain the majority of livestock producers in these countries for some time to come.

To facilitate the development of evidence based pro-poor HPAI control measures the project is designed so that there are five work streams: disease risk, livelihood impact, institutional mechanisms, risk communication, and synthesis analysis. Project teams are allocating and collecting various types of data from study countries and employing novel methodologies from several disciplines within each of these work streams. So that efforts aren't duplicated and the outputs of one type of analysis feeds into another the methodologies in each work stream will be applied in a cohesive framework to gain complementarities between them based on uniformity of baselines and assumptions so that policy makers can have consistent policy recommendations. The figure below is the methodological framework used to depict how work stream outputs fit together. This brief discusses the methodologies to be used when conducting the behavioral field experiments highlighted in the methodological framework below.



Author

Angelino Viceisza, Postdoctoral Fellow, Markets, Trade and Institutions Division, International Food Policy Research Institute, 2033 K Street NW, Washington, DC 20006

Acknowledgements

I thank Ekin Birol, Laura Taylor and Máximo Torero for invaluable comments. All errors are my own.

Disclaimer

The views expressed in this report are those of the author(s) and are not necessarily endorsed by or representative of IFPRI, or of the cosponsoring or supporting organizations. This report is intended for discussion. It has not yet undergone editing and does not contain a final design of the field experiments. Please contact author before citing at a.viceisza@cgiar.org.

More information

For more information about the project please refer to www.hpai-research.net.

1. Introduction

Economics is the science that studies human behavior as a relationship between ends and scarce means that have alternative uses (Robbins, 1945). This makes economics a social science. Historically, the method and subject matter of economics have presupposed that it was a non-experimental (or "field observational") science more like astronomy or meteorology than physics or chemistry (Smith, 1987).

At approximately the mid-20th century professional economics began to change with the introduction of the "laboratory experiment" into economic method. This allowed for at least two developments. First, the "laboratory approach" made it possible to introduce demonstrable knowledge into the economists' attempt to understand economic transactions (i.e., markets, social interaction and so on). Second, this approach brought to the economist direct responsibility for an important source of scientific data generated by controlled processes that can be replicated and validated by others.

These developments invited economic theorists (in particular, game theorists) to submit to a new discipline, but also brought an important new discipline and standard of rigor to the data generating process itself.[1] Since the developments of the mid-20th century, the use of experimental methods in economics has grown extensively. In particular, the "laboratory approach" to economics has grown beyond the laboratory to include experiments in the field. For a detailed account on the past, present and future of field experimentation, see Levitt and List (2008).

The purpose of this paper is twofold. First, to discuss the concept of a "field experiment" and its strengths and weaknesses for studying economic behavior. Second, to address a main application (i.e., example) of a field experiment in the context of the DFID-funded project on pro-poor policy options for the prevention and control of Highly Pathogenic Avian Influenza (HPAI). The sample field experiment discussed in this paper only applies to Indonesia even though experiments will also be conducted in Ghana or Nigeria. Since the discussions pertaining to the Africa experiments are at a very preliminary stage, they are not addressed further in this note; however, these experiments are likely to address issues of compensation and/or insurance.

The paper proceeds as follows. Section 2 pins down some terminology. Section 3 motivates the use of experiments. Section 4 discusses the concept of an experiment from a technical standpoint and summarizes some examples of previous experiments. Section 5 discusses the main application for the case of HPAI in Indonesia. Finally, section 6 concludes.

[1]When economists speak of data generating processes, they typically think of that *part* of the data generating process that is verifiable by means of *action-observable choices*, (see Kőszegi and Rabin, 2008, for definitions). Such data (sometimes referred to as “choice data") are usually analyzed using the revealed preference approach (Samuelson, 1948). The relatively recent “neuroeconomics" literature (i.e., neuroscience applied to economics) has sparked considerable debate on the extent to which economists should also focus on *action-unobservable choices* (also known as “non-choice data") such as smiles, brain images, dopamine levels and so on. Advocates of "neuroeconomics" argue that economists have traditionally only specified and controlled *part* of the data generating process by failing to look at “non-choice data". This methodological debate is at the core of experimental methods in economics. Namely, if “non-choice data" matter, failing to account for them in the data generating process results in a priori lack of control and ex post misspecification. This paper will focus on "choice data" unless otherwise noted. For further discussions on this debate, see Caplin and Schotter (2008); particularly, the lead article by Gul and Pesendorfer (2008).

2. Terminology and related literature

The term "experimental economics" tends to be used interchangeably with "behavioral economics". However, it should be noted that these terms (i.e., subfields of economics) need not be the same. Behavioral economics (as advocated by the Russell Sage Foundation's Behavioral Economics Roundtable) is the combination of psychology and economics that investigates what happens in markets in which some of the agents display human limitations and complications (Mullainathan and Thaler, 2001). Experimental economics constitutes the use of experimental methods (laboratory or field) to address general economic questions that may or may not be related to psychology. The distinction between the two fields has become more nuanced as both fields are increasingly using experimental techniques to address psychology-related questions in economics. The term "field experiments" is used differently in experimental economics (e.g., Harrison and List, 2004) than in (new) development economics (e.g., Duflo, 2006).[2]

Traditionally, experimental economics has been driven by laboratory experiments (also known as "lab experiments"); i.e., experiments conducted with students at academic institutions as subjects. While one of the main strengths of these experiments is their high degree of control over the data generating process, lab experiments have recently been criticized for lack of generalizability, particularly, when addressing questions related to social preferences (Levitt and List, 2007).[3] This critique among others has driven experimental economists—depending on the question under consideration—to conduct lab experiments in conjunction with field experiments. Field experiments in the context of Harrison and List (2004) are classified into three categories: (i) artefactual, (ii) framed and (iii) natural field experiments.

Artefactual experiments tend to be defined as "lab experiments in the field". They are field experiments in the sense that they are conducted with field participants (e.g., farmers), but the experiment tasks are conducted as part of an artefactual environment (similarly to lab experiments). Framed experiments tend to be conducted in actual day-to-day settings with field subjects who know that they are part of an experiment. Finally, natural experiments are the same as framed experiments with the exception that subjects typically do not know that they are part of an experiment. In other words, the experiment tasks are performed as part of the subjects' day-to-day (i.e., natural) environments. These field experiments tend to be behavioral in nature, because they are heavily concerned with the behavioral foundations that drive identified impacts. Furthermore, these are "true experiments" in the sense that the interventions are typically designed *a priori* with the aim of establishing a proper counterfactual for purposes of estimation and inference.

New development economics on the other hand has been driven by the "laboratory approach" applied to the field data and thus, has mainly been concerned with randomized trials conducted *ex post* to assess impacts of interventions that need not have been designed as "experiments" a priori. Furthermore, these trials need not always be concerned with the behavioral processes (i.e., theoretical foundations) driving some of the identified impacts; a critique that has been voiced by for example Mookherjee (2005). To the extent that randomized trials were a priori designed as experiments to identify outcome as well as behavioral impacts, they are the same as field experiments in the sense of the previous paragraph. In fact, the reader will note that the Harrison

[2] The term (new) development economics is discussed further in Basu (2005) and Banerjee (2005).
[3] Lab experiments have been shown to be relatively robust in other contexts (e.g., Brookshire et al., 1987).

and List (2004) taxonomy proposed previously is general enough to capture definitions in either discipline, thus allowing for both behavioral and outcome impacts. The critique that randomized trials lack theoretical foundations has driven new development economics to become more behavioral in nature.

The need for both disciplines to deal with critiques of generalizability of different types has pushed for a "new" approach that I would like to term *"behavioral development economics"*.[4] This terminology is not intended to take away from previous work in either game theory, experimental/behavioral or development economics. In particular, it is not my claim that development has not been behavioral and/or game-theoretic in nature. In fact, as discussed by Bardhan (1993), many novel theoretical foundations and methods were introduced to general economic thought by development economics. Rather, by adopting this terminology I posit that development and experimental economics are increasingly converging to a state in which randomized theory-testing field experiments (possibly complemented by laboratory experiments) are seen as an operationally meaningful way to (i) study economic behavior and mechanism (i.e., institutional) design and (ii) generalize findings to inform policy work. Unless otherwise stated, this is how the term "field experiment" is used throughout this paper.

[4] To a certain extent, this terminology was inspired by Camerer (2003) (titled “behavioral game theory") and Bertrand et al. (2004).

3. Why conduct experiments?

The rationale for conducting field experiments can be traced back to the process of quantitative research. See for example the Mookherjee (2005) modification of Haavelmo's (e.g., Haavelmo, 1958) classification, which constitutes the following stages:

a. Stage 1: Empirical description of the relevant phenomenon, consisting of exploratory data analysis aimed at helping identify empirical regularities that need to be explained by a suitable theory, and in addition the nature of assumptions that such a theory can make without gross violation to the empirical patterns.
- Stage 2: The formulation of a relevant theory, including derivation of potentially observable (hence falsifiable) implications.
- Stage 3: The estimation and testing of theories, a stage which may lead back to modification or replacement of the previous theories, in an iterative back-and-forth process with Stage 2.
- Stage 4: Use of the least unsuccessful theory--from the standpoint of empirical verification—for purposes of prediction and policy implementation and evaluation.

The unifying theme underlying this process is that economists (i) observe choice (and perhaps, non-choice) data and (ii) seek to explain them. In some sense, this process is hierarchical. If we do not observe data, there is nothing to be explained.[5] So, it is sensible to ask where the observed data come from; i.e., how were they generated and collected? This is mainly how the experimental approach differs from the classical field observational approach to economics.

Suppose we live in a first-best world in which we (i) know and (ii) observe the complete data generating process for all variables, $V^* \subseteq \mathbb{R}^n$ for $n \in \mathbb{N}$. In such a world, we have perfect "predictive ability", since we know/observe all relevant variables and the complete data generating process. However, we do not live in a first-best world. In fact, we live in a second-best world in which we (i) do not observe all variables and (ii) do not know the true process that generated the data. We only know/observe a proper subset of all variables ($V \subset V^*$) and somehow, we must use these to infer a relationship between such variables. So, how do we proceed?

Well, suppose that V can be partitioned into X and Y. Then, the classical (non-Bayesian) way of proceeding is to:

1. View X and Y as random variables.
2. Assume existence of some relationship between X, Y and some error term ε (also a random variable); i.e., $Y \sim (X, \varepsilon)$.
3. Assume some form of exogeneity on X, such as $E(\varepsilon \mid X) = 0$.
4. Explore the relationship between X and Y (possibly, non-parametrically).
5. Draw inference therefrom for purpose of prediction.

In our second-best world (i.e., when working with real data) point 3 fails under many circumstances mainly due to unobserved data (i.e., the reduction of dimensionality when moving from V^* to V). The question thus becomes how to proceed when this assumption is violated. The experimental approach differs from the traditional approach mainly in how and when one deals with this violation.

[5] For the sake of argument, the exposition here is deliberately one-sided. However, it is generally accepted that there is continuous two-way feedback between empirics and theory (e.g., Samuelson, 2005). I.e., both inform each other. An example is Mookherjee's Stage 3 described previously.

While both approaches attempt to control for this violation, experimental approaches do so by using *a priori* techniques—such as, control of the "data generating process" by design of the experiment—as opposed to *ex post* methods such as instrumental variables (e.g., Wooldridge, 2002) or "randomized trials" (e.g., Duflo, 2006; Banerjee et al., 2006).[6]

A carefully designed experiment rids ex post econometric analysis of most common problems faced by *field observational* data sets. This is the principle rationale for conducting field experiments. However, in order to carefully design a field experiment, the experimenter needs a solid understanding of the first-best (i.e., true) data generating process, which brings us back to Mookherjee's four-stage process for conducting quantitative research. Namely, in order to carefully design an experiment, it is essential to have a theoretical model in mind. And, in order to have a theoretical model in mind, it is important to have undergone Stage 1. So, experiments should not be seen as stand-alone objects. Rather, experiments should be seen as meaningful theory-testing instruments that fit well into the four-stage process of quantitative research. Furthermore, if linked to a well designed model, they can provide generalizable results that inform policy-making.

Finally, three comments. First, theory-testing experiments are not the only types of experiments. In fact, Smith (1982) draws the distinction between theory-testing (i.e., nomothetic) and heuristic experiments—among others. The latter are not theory-testing in the sense that they tend to be exploratory; e.g., to inform the development of new theory when certain models are scientifically underdeveloped. While these types of experiments have their place in the four-stage process, they are not the type we have in mind for policy-making. So, our main focus is on theory-testing experiments.

Second, the above discussion illustrates the complementarity between surveys and experiments. A good method for gaining an understanding of the status quo in Stage 1 is to conduct baseline surveys (and perhaps, rapid appraisals). These methods inform the stylized facts that drive the model-building in Stage 2, which in turn informs the design of the experiments. So, experiments and surveys tend to be meaningful complements, particularly, in the field. Of course, surveys do not only serve a purpose in Stage 1 of the research process.

Third, experiments should be conducted in the context of the four-stage process. Since the design of an experiment controls several aspects of the data generating process, an experiment in isolation can lead to non-generalizable results. Namely, if an experimenter does not carefully go through Stage 1 and Stage 2 and thus, gains a wrong understanding of the data generating process, the experiment can suffer from *"ex ante model misspecification"*. This term is used to distinguish from standard *ex post model misspecification*, which occurs when specifying an econometric model once the data have been generated.

[6] We crudely use the term "randomized trials" to refer to the (new) empirical techniques for ex post impact evaluation such as the use of propensity score matching to establish counterfactuals (typically, a control group), which in turn is used to do for example difference-in-difference estimation. For a textbook reference, see Wooldridge (2002).

4. What are experiments?

4.1. Generic discussion

This section provides a micro-theoretic "definition" of an experiment. The discussion is generic and applies to any type of experiment regardless of the research question under consideration. We find it useful to set these theoretical foundations in generic terms, so that we can appeal to them when discussing the application. The discussion draws heavily on Smith (1987).

In defining a microeconomic system two components will be identified: (i) an environment and (ii) an institution. The environment consists of a set of N economic agents $\{1,...,N\}$, a set of $K+1$ commodities (including resources) $\{0,1,...,K\}$ and certain characteristics of each agent i, such as the agent's preference relation $\succ^i$, technology (knowledge) endowment T^i, and a commodity endowment vector ω^i. Hence, the ith agent is characterized by the vector $e^i = (\succ^i, T^i, \omega^i)$ for which the components are assumed to be defined on the $K+1$ dimensional commodity space $\mathbb{R}^{K+1}$. Hence, a microeconomic *environment* is defined by the collection of characteristics $e = (e^1,...,e^N)$.

The institution specifies:

1. A *language* $M = (M^1,...,M^N)$ consisting of messages $m = (m^1,...,m^N)$, where m^i is an element of M^i, the set of messages that can be sent by agent i. Note two things. First, m may refer to the final allocation-determining message. Second, the set of allowable messages M^i need not be unique across all i.
2. A set of allocation rules for each i, $H = (h^1(m),...,h^N(m))$. The rule $h^i(m)$ states the final commodity allocation to each i as a function of the messages sent by all agents which precedes the allocation.
3. A set of cost imputation rules $C = (c^1(m),...,c^N(m))$. The rule $c^i(m)$ states the payment to be made by each agent in numeraire units (money) as a function of the messages sent by all agents.[7]
4. A set of adjustment process rules, $G = (g^1(t_0,t,T),..., g^N(t_0,t,T))$. These rules consist of a (i) starting, (ii) transition and (iii) stopping rule.

Each agent i's property rights in communication and in exchange are defined by

$$I^i = (M^i, h^i(m), c^i(m), g^i(t_0,t,T)),$$

which specifies the messages that i has the right to send; the starting, transition and stopping rules which govern these communication rights; and finally the right to claim commodities or payments in accordance with the outcome rules that apply to messages.

A microeconomic *institution* is defined by the collection of all these individual property right characteristics $I = (I^1,...,I^N)$. A microeconomic environment (e) and a microeconomic institution (I) together define a *microeconomic system*, $S = (e, I)$. Agent i's outcome behavior is defined by a

[7]Note that C is redundant in that it could be included in the definition of H, but it will be convenient in many applications (as when there are no income effects) to distinguish between commodity allocations by H and payment imputations by C.

function $\beta^i(e^i \mid I)$ which yields the allocation-determining message m^i sent by agent i with characteristic e^i, given the property rights of all agents defined by I.

Ultimately, an experiment becomes an empirical exercise. So, in order to think about how the above discussion can be applied to construct and conduct an experiment, we must address two questions. The first question is what do we hope to learn from conducting the experiment? This question is relevant because it dictates (i) those observable characteristics that we want to control or (ii) those unobservables that we might want to elicit as part of the experiment. The second question is which of the above-mentioned elements are observable in the field?

Among the observable elements of an economy are: (i) the list of agents, (ii) the list of physical commodities and resources, (iii) the physical commodity and resource endowments of individual agents, (iv) the language and property right characteristics of institutions, and (v) outcomes (e.g., choices). What is unobservable are (vi) preference relations, (vii) technological (knowledge, human capital) endowments, and (viii) agent message behavior $\beta^i(e^i \mid I)$, $i = 1,..., N$.

The fact that certain variables in the microeconomic system are unobserved can pose a problem in applications depending on the research question the experiment seeks to answer. For example, consider experiments that test game-theoretic predictions. The usual way of constructing such experiments is to (i) define a model with parametric preferences, (ii) apply the appropriate solution concept to derive testable hypotheses and (iii) conduct an experiment to test such predictions.

However, as Ray and Zhou (2001) and Cox (2004)—among others—have pointed out, if the theoretical prediction differs from the observed outcome, there could be two possible explanations: either individuals do not play according to the theory, or preferences are misspecified. Notice that the extent to which this is problematic depends on the research question under consideration. For example, if the intention of the experiment is to observe behavior under alternative institutions, "stability" of preferences—which would be an additional assumption/premise—would reduce the concern.

On the other hand, if the aim of the experiment is to test a particular game-theoretic solution concept, this becomes a serious concern. One of the ways of dealing with this issue is to proceed by modeling preferences non-parametrically; i.e., to consider a revealed preference approach to games as discussed by Ray and Zhou (2001). However, this approach can have its limitations, (see Weibull, 2004, for the so-called "Weibull critique"). Finally, another approach is to attempt to empirically validate preferences by conducting (possibly tangential) experiments that elicit preferences such as experiments on social, risk and time preferences.

Typically, experiments will try to observe and explain outcome behavior (i.e., a proxy for $\beta^i(e^i \mid I)$, call it $\hat{\beta}^i(e^i \mid I)$) in the form of choice data, under alternative institutions (i.e., by changing I). In some cases, the experiment may be specifically designed to elicit characteristics of the environment (e) such as individual preferences. In other cases, preferences (e) are assumed to be stable and the experiment seeks to characterize changes in outcome behavior (i.e., the shape of $\hat{\beta}^i(.)$) across changing institutions (I).

4.2 Some examples of experiments

Many experimental applications have appeared in the economics literature. One study that is most closely related to the experiments proposed in this paper is by Bennett and Willis (2007). Bennett and Willis report choice experiments that estimate the value that society places on changes to the size of the badger population. The study was undertaken in the context of the possible need to reduce the badger population by culling to help control bovine tuberculosis in cattle. The study found that people were concerned about the problem of bovine tuberculosis in cattle, as reflected in their willingness to pay to control the disease; however, people did not like the idea of a policy that intentionally killed large numbers of badgers, as reflected in their relatively high willingness to pay not to have such a policy.

To the extent of our knowledge, very few experiments in the published literature have explored people's preferences towards and the effect thereof on the prevention and/or control of animal diseases (in particular, HPAI). However, there have been *numerous* studies assessing (i) agents' preferences and (ii) interactions between preferences and alternative institutions that can be applied to the context under consideration. Below, we discuss further examples of experiments in two broad categories. The first is risk (and time) experiments and the second is contingent valuation (choice) experiments. The choice of these two categories is not random, as these are the types of experiments we plan to conduct in the context of this project.

Some of the most celebrated risk experiments in a field context are those of Binswanger (1980). Binswanger conducted risk experiments with households in rural India using real monetary payoffs and found that at high stakes, most individuals are risk averse—despite wealth levels. Since, risk experiments have become extremely important in economics and with good reason: most models view risk preferences as an important determinant of behavior. Consequently, it is important to elicit risk preferences in order to be able to control for them when doing estimation and inference.[8] In the case of the HPAI project, we see risk as an important factor as well, since occurrence of the disease in itself is random and thus, risky. Furthermore, an HPAI outbreak puts household livelihoods at risk. This is another way in which risk impacts people's behavior during an outbreak.

The second set of experiments are in the class of stated preference methods (SPMs), which encompass both contingent valuation and choice experiments. SPMs have been applied in many contexts to assess consumers' willingess to pay (i.e., demand) for attributes of different goods such as food safety, food quality and generic product characteristics. One such study was conducted by Alfnes et al. (2006).[9] These authors looked at consumers' willingness to pay for the color of salmon. They found that consumers use color of salmon as a quality indicator and that they are willing to pay significantly more for salmon fillets with normal or above-normal redness, as compared with paler salmon fillets. Our experiments will be similar in the sense that they will gauge consumers' willingness to pay for market cleanliness. Furthermore, we will also explore the extent to which producers are willing to lose some surplus (during periods of market closure) in order to provide the attribute of market cleanliness.

[8] Recently, Andersen et al. (2008) showed that it is important to elicit both risk and time preferences.

[9] For additional references, see the review by Birol et al. (2008).

5. An application

5.1 Motivation

This section discusses preliminary details for a potential field experiment in Indonesia related to the DFID-funded project on pro-poor policy options for the prevention and control of HPAI. We start by motivating the main question in terms of proposed policy options in Indonesia. Then, we discuss the behavioral model underlying the experiments as well as some details on the experimental design and protocol. The basic information for the "market closures" is based on discussions with several Indonesian stakeholders, including but not limited to Dr. Elly Sawitri Siregar, head of the HPAI Campaign Management Unit that is part of the Indonesian Department of Agriculture (DEPTAN). These discussions are ongoing.

As part of their National Strategic Plan for the containment and eradication of HPAI, the Indonesian government is planning to implement market closures. The basic idea behind a market closure is the following: At a regularly scheduled pre-announced date and time, the market in question will be closed for thorough cleaning. For example, in the case of a live-bird market this would entail removing all birds and cages from the market and thoroughly cleaning both the premises and the cages prior to re-installation.

The government is concerned with the economic ramifications of the proposed policy. Namely, once the policy is introduced, it will affect both consumers and producers. Specifically, three key economic questions arise:[10]

1. How will this policy affect consumers' and producers' surplus?
2. Given the magnitude of these economic losses, what are the expected changes in behavior both on the consumer and the producer side?
3. If the policy changes induce incentives towards perverse behavior, how can the government secure that these are minimal in terms of economic disturbance?

The standard (i.e., neoclassical) approach to markets is to model consumers and producers whose aggregate demand and supply schedules give rise to an equilibrium price in the market (e.g., Mas-Colell et al., 1995). The typical graph that most readers are familiar with from introductory economics textbooks is given in figure 1.[11] However, the basic idea is the following: If we assume that consumers and producers have well-defined preferences, then rationality and optimality manifest themselves in aggregate form as well-defined demand and supply functions. Each point on the demand curve indicates consumers' aggregate willingness to pay for a particular quantity of the commodity in question. Similarly, the supply curve indicates producers' aggregate willingness to accept at any given quantity.

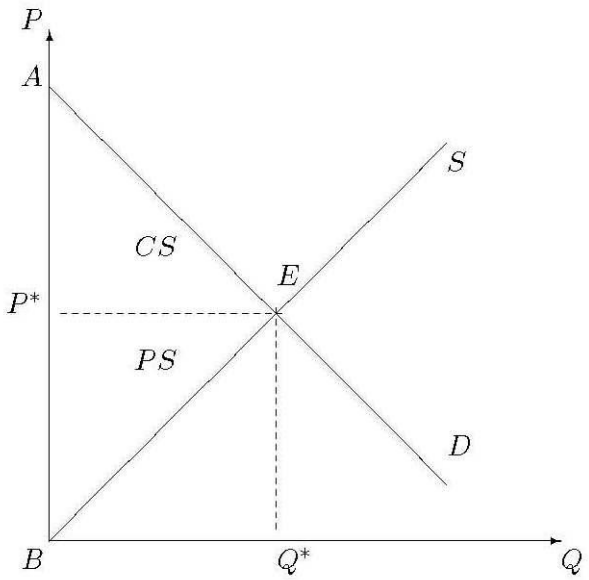
Markets exist because they have benefits (in terms of increased welfare) associated with them. One way to measure welfare is by means of consumers' and producers' surplus. These concepts are best explained using figure 1. Consider point A in figure 1. Suppose that with this point there is associated the following price-quantity combination, (P_A, Q_A). Compare P_A with P^*. Clearly, $P_A > P^*$. How can we use this to formulate a welfare measure? Well, note that even though the aggregate consumer pays a price of P^* in equilibrium, the demand curve indicates that at quantity

[10] The DFID-funded project as a whole is concerned with other consequences such as epidemiological ones.

[11] The schedules in this figure can rigorously be derived; see Mas-Colell et al. (1995).

Q_A, the aggregate consumer is willing to pay P_A. So, the aggregate consumer gets a surplus since in equilibrium she pays a lower price than she would have been willing to pay for this particular quantity. We can reason similarly for quantities to the right of Q_A up to the equilibrium quantity Q^*. *Consumers' surplus* is thus indicated by the area marked CS; i.e., the area of the triangle (Δ) named AEP^*. A similar argument with regard to the aggregate supply curve yields the concept of *producers' surplus* (PS). A market is associated with these two welfare measures, which together may be referred to as *total surplus* ($TS = CS + PS$).[12]

Figure 1: Market Equilibrium (*Q*, P)**



Now, what happens if a market is closed? Suppose a market operates over time (e.g., on a daily basis). Then, we can associate with it a daily surplus of TS_t, where the subscript t denotes a particular day. A daily market closure implies a surplus loss of TS_t for every t. Two things are worth noting here. First, the daily surplus need not be the same across all time periods. Second, the total surplus consists of two components: consumers' and producers' surplus. In other words, when there is a closure, both parties lose.

Under the assumption of scarcity, lost surplus induces economic agents to make up for economic losses in other markets. In the case of market closures in Indonesia, we might expect agents to find informal (i.e., unobserved) markets in which to engage in alternative economic activity. In other words, the government policy might induce agents to resort to informal economic activity, which may be even less hygienic than those that the government is trying to avoid. It is in this sense that the proposed policy action can induce perverse economic incentives, since it may reduce the probability of disease containment.

How then we can we deal with this behavioral issue? Recall that as discussed by Smith (1982), message behavior $\beta^i(.)$ is a function of two variables, the environment (e, i.e., preferences) and the institution (I). So, accordingly if we observe $\hat{\beta}^i(.)$, we can focus on one or both of its

[12] See MWG95 for a formal definition of these surpluses and how to calculate them.

determinants. We can try to build institutions such that perverse incentives from the government policy are minimized. Alternatively, we can study the extent to which agents' preferences will lead to minimal perverse behavior. On the institutional side, one option is to close the market on a date and at a time when the surplus loss is expected to be minimal. For example, a period during which there is typically minimal activity in the market. However, this may be infeasible. A second option is to increase monitoring of informal activity on days of market closure. This in turn may be costly. So, it is worth asking to what extent agents' preferences may be such that the resulting perverse behavior will be minimal. This is where our field experiments will play a role.

The proposed field experiments ask whether producers have an incentive to comply with the market closure due to consumers' willingness to pay for "clean" markets. Namely, if consumers are willing to pay for cleaner markets, this will be reflected in a price premium for producers. This has two implications: First, consumers will be willing to incur loss of consumers' surplus (during periods of market closure) in order to get "cleaner" products. Second, producers will be willing to lose surplus (during periods of market closure) in order to gain a price premium in other periods, *as long as* the premium is sufficient to compensate for the loss.

5.2 The behavioral model

The traditional approach to consumer theory has been to assume that goods are the directs object of utility; i.e., the arguments over which we have preference orderings. Instead, in the Lancaster (1966) framework it is the properties or characteristics of goods from which utility is derived. In other words, Lancaster assumes that consumption is an activity in which goods, singly or in combination, are inputs and in which the output is a collection of characteristics.

The appeal of this approach for the question under consideration is that we can view a particular product or collection of products (such as poultry or poultry products) as possessing the same (set of) characteristic(s); in this case, the "clean market" characteristic.[13] Accordingly, we can gauge consumers' willingness to pay for this attribute and producers' willingness to accept losses (during market closure) to supply it.

The relationship between the level of activity k, y_k, and the goods consumed in that activity are linear and objective, so that if x_j is the j th commodity we have

$$x_j = \sum_k a_{jk} y_k, \tag{1}$$

and the vector of total goods required for a given activity vector is given by

$$x = Ay. \tag{2}$$

[13] As for the terminology, we recognize that poultry supplied in formal markets have most of the time come from many distinct origins, typically via traders. So, we find the claim that "market closure" will lead to "safe poultry" too bold. However, it is not unreasonable to claim that poultry sold in a market subject to "market closure" has come from a cleaner market environment in the sense that such market contains cleaner storing and slaughter facilities. So, we call this the "clean market" or "market cleanliness" attribute of poultry as opposed to the "food safety" attribute.

Each consumption activity k produces a fixed vector of characteristics, z. The relationship between the level of activity k, y_k, and the ith characteristic z_i is linear

$$z_i = \sum_k b_{ik} y_k \tag{3}$$

or

$$z = By. \tag{4}$$

It is assumed that the coefficients a_{jk} and b_{ik} are objectively determined so that they hold for all individuals. In other words, these coefficients are determined by the intrinsic properties of the goods and characteristics themselves and possibly by the context of technological knowledge in society.

Consider this model in relation to the application under consideration. In this case, a particular consumption activity y_k could be "consuming from a traditional Indonesian market". The associated commodity could be "pounds of chicken" and the associated characteristic could be whether or not the market is subject to a market closure, i.e., "market cleanliness".

It is further assumed that the individual possesses an ordinal utility function $U(.)$ on characteristics z, $U(z)$ and that he will choose a situation that maximizes z. $U(z)$ is assumed to satisfy the standard properties of a neoclassical utility function (i.e., increasing, twice continuously differentiable and regular strictly quasiconcave).

Note that in this model, the relationship between the collections of characteristics available to the consumer, z, which are direct ingredients of preferences and welfare, and the collection of goods available, x, which represent the relationship with the rest of the economy, is indirect through the activity vector y.

Suppose that the relationship between goods and activities is one-to-one. Then, we can rewrite $x = Ay$ as $y = A^{-1}x$ such that $z = By$ can be expressed as $z = Cx$, where $C = BA^{-1}$. As a result, we can write the consumer problem in simpler form as follows:

$$\max_z U(z) \quad \text{s.t.} \quad px \le k, \quad \text{with} \quad z = Cx, \quad z, x \ge 0, \tag{5}$$

which can further be rewritten as

$$\max_z U(z) \quad \text{s.t.} \quad pC^{-1}z \le k, \quad \text{with} \quad z \ge 0. \tag{6}$$

For simplicity, suppose that $k \in \mathbb{R}^1$, $z, p \in \mathbb{R}^2$ with $z = (z_1, z_2)^T$ (where the T stands for transpose) and $p = (p_1, p_2)$ and $\mathbb{R}^{2\times 2} \ni C^{-1} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$.

Given these simplifying suppositions, the consumer choice problem becomes:

$$\max_{\{z_1, z_2\}} U(z_1, z_2) \quad \text{s.t.} \quad (p_1c_{11} + p_2c_{21})z_1 + (p_1c_{12} + p_2c_{22})z_2 \le k, \quad \text{with} \quad z_1, z_2 \ge 0, \tag{7}$$

with necessary (and sufficient by assumption of regular strict quasiconcavity) first-order conditions:

$$\frac{\partial L}{\partial z_1} = U_1 - \lambda(p_1 c_{11} + p_2 c_{21}) = 0 \quad (8)$$

$$\frac{\partial L}{\partial z_2} = U_2 - \lambda(p_1 c_{12} + p_2 c_{22}) = 0 \quad (9)$$

$$\frac{\partial L}{\partial \lambda} = k - (p_1 c_{11} + p_2 c_{21}) z_1 - (p_1 c_{12} + p_2 c_{22}) z_2 = 0, \quad (10)$$

where L stands for the Lagrangian, λ stands for the Lagrange multiplier and $U_i \equiv \frac{\partial U}{\partial z_i}$.

By the Implicit Function Theorem, it can be shown that there exist Marshallian demand functions for the characteristics, which can be expressed as $z_1^* = z_1(p_1, p_2, k; c_{ij})$ and $z_2^* = z_2(p_1, p_2, k; c_{ij})$ for $i, j = 1,2$. Recall that c_{ij} relates the characteristics back to the goods in question.

So far, we have discussed the attributes approach for the consumer side. Similarly, we can think of producers supplying goods that have certain characteristics, say z_1 and z_2. Under such conditions we can think of a market where equilibrium price is determined based on demand for and supply of certain product characteristics; in this case, "market cleanliness".

5.3 Experimental design: treatments and protocol

We anticipate conducting two types of experiments in Indonesia as part of this project. The first set is a set of risk and time experiments as discussed in Andersen et al. (2008).[14] As discussed previously, we expect preferences for risk to play an important role in how economic agents react to an occurrence of HPAI. The second set of experiments—which constitute the main experiments of interest—will assess (i) whether there exists demand (i.e., willingness to pay) for "market cleanliness" and (ii) whether suppliers are willing to lose some surplus in order to supply "market cleanliness" and charge a higher price in the future. A comparison of these two measures will give an idea of the extent to which the proposed market closures might induce consumers and producers to engage in informal market activity. These main experiments will be valuation experiments in the sense that they will attempt to put a monetary value on agents' preferences for the attribute of interest; i.e., “market cleanliness”.[15]

There is a vast literature on (contingent) valuation and choice experiments, spanning several fields including economics. The typical issues encountered in this literature are: (i) how to minimize overestimation due to the hypothetical nature of such experiments (in particular, how to make them incentive compatible), (ii) how to frame such experiments and (iii) how to do ex post estimation and inference based on stated preference data once the experimental data have been collected. First, we discuss a potential way of conducting these experiments. Then, we use this to elaborate on the

[14] For a typical risk experiment, please see Andersen et al. (2008) and/or the references cited within.

[15] We also anticipate interacting these valuation treatments with other variables of interest such as information. Prior to stating/revealing preference, it is necessary to be well informed about the specific attribute of interest. So, varying degrees of information is an important aspect of study.

above mentioned issues and explain how we plan to mitigate them as part of our experimental design.

Consider the following question: to what extent do consumers exhibit a willingness to pay for "market cleanliness" when consuming poultry products? How can we possibly answer such a question? Well, one way is to conduct a willingness to pay (stated preference) experiment. Suppose we have a sample of 100 respondents that have been (randomly) selected for this particular treatment. One potential way of eliciting their willingness to pay is to proceed as follows:

1. Calibrate an initial value of a particular poultry product at P_0 according to the current market price or in the absence thereof, a randomly assigned initial price from a discrete uniform distribution with reasonable support.
2. Based on such initial price, ask the agent whether or not he or she is willing to pay P_0 for the product with the characteristic of interest.
3. If the answer is "Yes", then proceed to the next higher price, say $P_0 + j$, where $j = 1,2,..., J$.
4. Continue this process until the answer to the question is "No". Suppose the answer is "No" on the Kth iteration. Then, the value $P_0 + K$ becomes the consumers' willingness to pay.

The above process in itself is relatively simple. However, the issues mentioned above arise immediately, if this procedure is applied as is. First, the fact that the question is hypothetical (i.e., non-binding and not affecting payoffs), may not give the respondent sufficient incentive to answer truthfully. This is contrary to the case in which he or she were actually forced to purchase the product at the specified price. To deal with this issue, Lusk et al. (2008) and Alfnes et al. (2006)—among others—have attempted to make the consumers' choice binding by requiring the consumer to buy the product once his or her choice has been made. Furthermore, once the consumer choice is (possibly) binding, he or she has an incentive to lie about $P_0 + K$. So, these researchers have also spent considerable time trying to ensure that the ranking elicitation mechanism used is incentive compatible.

Second, the fact that the question is often posed for products that may not necessarily exist at the time the study is being conducted, requires careful framing and elicitation of the respondent's preference statement. Train and Wilson (2007) (and the numerous references within) suggest an approach called "pivoting" that builds on an actual choice that a respondent recently made. An example would be the following: Suppose we ask the respondent to describe a recent purchase of poultry at a traditional market and how much he or she paid for it. Then, based on that, the experimenter would pivot and ask about a choice that is slightly different; namely, one in which the same product is purchased from a market that has been closed (i.e., regularly thoroughly cleaned). The main advantage of "pivoting" (i.e., iterating from a familiar choice) is that it increases the realism of the task by relating the (hypothetical) stated preference choice with an actual choice that the respondent made in a revealed preference environment. So, this in part takes care of framing as well.

Finally, if the above considerations are taken into account when eliciting participants' stated preferences, ex post, the experimenter—now in the role of econometrician—is faced with the task of estimating willingness to pay using stated preference data, possibly combined with revealed preference data. This gives rise to endogeneity, namely, a dependence between the stated

preference attributes and unobservables. Train and Wilson (2007) propose an estimation technique that deals with this issue of lack of independence.

Given these considerations, we can think of a revised willingness to pay experiment that consists of the standard model adapted for elements of Lusk et al. (2008) and Train and Wilson (2007). Suppose again that we have a sample of 100 respondents. Then, we could conduct an experiment by performing the following steps:

1. Ask the participant regarding a recent experience where he or she purchased a poultry product from a particular market.
2. From this, infer a benchmark for the participant.
3. Iterate from this benchmark (similarly to Train and Wilson, 2007) to present the subject with a new task that is somewhat familiar: purchasing a poultry product from a particular market that has been closed.
4. Now, apply conjoint ranking analysis (similarly to Lusk et al., 2008) to have the participant value and rank the poultry product from the closed versus the non-closed market. [16] The ranking can be done in terms of prices and involves the random event that the choice becomes binding. The prices (i.e., ranking) and probabilities of winning the lottery—in which case the choice becomes binding—are proportional such that they induce incentive compatibility; i.e., the highest ranked alternative has the higher probability of winning. The lottery can be implemented as a wheel similar to that used in game shows such as "Wheel of Fortune". Different pieces of the pie indicate different probabilities.
5. Implement the lottery and determine the respondent's willingness to pay.

These five steps now constitute a willingness to pay experiment, to be applied with consumers. Similarly, we can envision a willingness to accept/lose experiment with producers. We can then analyze the data (perhaps, using a method similar to Train and Wilson, 2007) to compare the responses to see if the net surplus justifies the proposed policy. It should be noted that while the above may seem detailed, it is just a brief discussion of a potential experiment. In particular, several additional details need to be sorted out prior to implementation of these experiments. Some of these are: (i) the different treatments to be implemented (e.g., we can vary the type of the poultry product, the information provided to respondents, the pricing and the ranking mechanism), (ii) the location and frequency of the experiments (i.e., which regions, which markets, which agents, during what timeframes), (iii) the sample sizes per treatment in order to achieve sufficient power (e.g., we can envision 100 respondents per treatment) and (iv) the best way (i.e., level and method) to reward subjects for their participation in the experiments. We expect these details to be sorted out soon based on the pending discussions with stakeholders.

[16] We can also envision a non-dichotomous choice in which we vary the frequency of the market closure; e.g., consider a market that was closed for t hours/week, $t+i$ hours/week, where i=1,2,….

6. Conclusion

This paper discusses the concept of a field experiment, the motivation for conducting a field experiment as well as an application of a proposed field experiment to be conducted as part of the DFID-funded project on pro-poor policy options for HPAI Prevention and Control in Indonesia. The designs of these field experiments are under construction, since consultations with several relevant stakeholders are still in progress. Any questions and/or comments should be directed to a.viceisza@cgiar.org; especially, if one would like to cite this work.

7. References

Alfnes, F., A. G. Guttormsen, G. Steine, and K. Kolstad (2006). Consumers' willingness to pay for the color of salmon: A choice experiment with real economic incentives. American Journal of Agricultural Economics 88 (4), 1050–1061.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2008). Eliciting risk and time preferences. Econometrica 76 (3), 583–618.

Banerjee, A. (2005). New development economics and the challenge to theory. Economic and Political Weekly (October 1, 2005), 4340–4344.

Banerjee, A., E. Duflo, R. Glennerster, and M. Kremer (2006). Randomization in development economics: A toolkit. In Handbook of Development Economics, Volume 4.

Bardhan, P. (1993). Economics of development and the development of economics. Journal of Economic Perspectives 7 (2), 129–142.

Basu, K. (2005). New empirical development economics–remarks on its philosophical foundations. Economic and Political Weekly (October 1, 2005), 4336–4339.

Bennett, R. and K. Willis (2007). The value of badger populations and control of tuberculosis in cattle in England and Wales: A note. Journal of Agricultural Economics 58 (1), 152-156.

Bertrand, M., S. Mullainathan and E. Shafir (2004). A Behavioral-Economics View of Poverty. American Economic Review Papers and Proceedings, 419-423.

Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural India. American Journal of Agricultural Economics 62 (3), 395.

Birol, E., P. Koundouri and Y. Kountouris (2008). Applications of the Choice Experiment Method in Europe: A Review. In Birol, E. and P. Koundouri (Eds.) Choice Experiments Informing Environmental Policy: A European Perspective. Edward-Elgar Publishing, Wally Oates and Henk Folmer's 'New Horizons in Environmental Economics' Series, Cheltenham, UK.

Brookshire, D. S., D. L. Coursey, and W. D. Schulze (1987). The external validity of experimental economics techniques: Analysis of demand behavior. Economic Inquiry 25 (2), 239–250.

Camerer, C. (2003). Behavioral Game Theory. Princeton, NJ: Princeton University Press.

Caplin, A. and A. Schotter (Eds.) (2008). The Foundations of Positive and Normative Economics, A Handbook. The Handbooks in Economic Methodologies Series. New York, NY: Oxford University Press.

Cox, J. C. (2004). How to identify trust and reciprocity. Games and Economic Behavior 46 (2), 260–281.

Duflo, E. (2006). Field experiments in development economics. In R. Blundell, W. Newey, and T. Persson (Eds.), Advancements in Economic Theory and Econometrics, Volume 2. Cambridge, UK: Cambridge University Press.

Gul, F. and W. Pesendorfer (2008). The case for mindless economics. In A. Caplin and A. Schotter (Eds.), The Foundations of Positive and Normative Economics, A Handbook, The Handbooks in Economic Methodologies Series. New York, NY: Oxford University Press.

Haavelmo, T. (1958). The role of the econometrician in the advancement of economic theory. Econometrica 26 (3), 351–357.

Harrison, G. W. and J. A. List (2004). Field experiments. Journal of Economic Literature 42 (4), 1009–1055.

Kőszegi, B. and M. Rabin (2008). Revealed preferences and revealed mistakes. In A. Caplin and A. Schotter (Eds.), The Foundations of Positive and Normative Economics, A Handbook, The Handbooks in Economic Methodologies Series. New York, NY: Oxford University Press.

Lancaster, K. J. (1966). A new approach to consumer theory. The Journal of Political Economy 74 (2), 132–157.

Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? The Journal of Economic Perspectives 21, 153–174.

Levitt, S. D. and J. A. List (2008). Field Experiments in Economics: The Past, The Present, and The Future. NBER Working Paper 14356.

Lusk, J. L., D. Fields, and W. Prevatt (2008). An incentive compatible conjoint ranking mechanism. American Journal of Agricultural Economics 90 (2), 487–498.

Mas-Colell, A., M. Whinston, and J. Green (1995). Microeconomic Theory. Oxford University Press.

Mookherjee, D. (2005). Is there too little theory in development economics today? Economic and Political Weekly (October 1, 2005), 4328–4333.

Mullainathan, S. and R. Thaler (2001). Behavioral economics. In International Encyclopedia of Social Sciences (1st ed.)., pp. 1094–1100. Pergamon Press.

Ray, I. and L. Zhou (2001). Game theory via revealed preferences. Games and Economic Behavior 37 (2), 415–424. 13

Robbins, L. (1945). An Essay on the Nature and Significance of Economic Science. London: Macmillan.

Samuelson, L. (2005). Economic theory and experimental economics. Journal of Economic Literature 43 (1), 65–107.

Samuelson, P. A. (1948). Consumption theory in terms of revealed preference. Economica 15 (60), 243–253.

Smith, V. L. (1982). Microeconomic systems as an experimental science. The American Economic Review 72 (5), 923–955.

Smith, V. L. (1987). Experimental methods in economics. In J. Eatwell, M. Milgate, and P. Newman (Eds.), The New Palgrave: A Dictionary of Economics.

Train, K. and W. W. Wilson (2008). Estimation on stated-preference experiments constructed from revealed-preference choices. Transportation Research, Part B: Methodology 40, 191–203.

Weibull, J. (2004). Testing game theory. In S. Huck (Ed.), Advances in Understanding Strategic Behaviour: Game Theory, Experiments and Bounded Rationality. Essay in Honour of Werner Güth., pp. 85–104. Palgrave MacMillan.

Wooldridge, J. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: The MIT Press.