Chapter 2

Concept and validity

A clear definition of the concept of trust is necessary in order to measure it. Building on a review of different theoretical approaches to the concept of trust, this chapter provides a working definition and a conceptual framework to underpin its measurement. The framework distinguishes between interpersonal and institutional trust and between the two main categories of interpersonal trust: trust in others, and trust in friends, family and neighbours. Four approaches to measuring trust are identified, which provide the organising framework for the question modules proposed in these Guidelines. The chapter also discusses the statistical quality of trust measures, with a focus on their relevance and accuracy. While measures of both interpersonal and institutional trust are of high relevance, the picture is mixed with respect to accuracy. In particular, while there is strong evidence supporting the reliability and validity of measures of trust in others, the evidence base is weaker with respect to trust in friends, family and neighbours, and, even more so, with respect to trust in institutions. For this reason, the chapter recommends that measures of trust in others should be included in official statistics, while measures of trust in friends, family and neighbours and of trust in institutions should be regarded as more experimental.

2.1. Introduction

This chapter focuses on the definition of trust and summarises the available evidence relating to the statistical quality of measures of trust. The first part of the chapter outlines a conceptual framework for measuring trust: before trust can be measured, it is essential to have a clear understanding of the concept(s) to be measured and of how this relates to different measurement approaches. Following this, the second part of the chapter considers issues of statistical quality, with a particular focus on the relevance and accuracy of trust measures. Relevance and accuracy are the focal aspects of statistical quality considered here because they are the elements of the *Quality Framework and Guidelines for OECD Statistical Activities* (OECD, 2008) that are most relevant to the issues covered by these Guidelines.

The conceptual framework developed here addresses the nature of trust and provides a model of the different types of trust that could be measured. This provides a framework for thinking about the most relevant measurement approaches and for identifying those types of trust that will be the focus for these Guidelines. Of the different types of trust identified in the conceptual framework, interpersonal trust and institutional trust are identified as falling within the scope of the Guidelines. Within these broad categories, trust in others (i.e. generalised trust) is identified as the most important element of interpersonal trust. A typology of approaches to measuring trust is also developed; this forms the basis for the different question modules proposed by the Guidelines, as described in Annex 2.

The largest part of this chapter focuses on issues of statistical quality, particularly on relevance and accuracy. Relevance concerns the degree to which measures of trust are sufficiently useful to justify collecting them in the context of the official statistical system. One section of the chapter deals with the degree to which measures of interpersonal and institutional trust are useful to inform policy making and the general public. Accuracy – and particularly validity – refers to the degree to which measures of trust capture the intended concept, which is a major concern for the measurement of trust. The chapter provides a brief overview of approaches to assessing validity, before considering both interpersonal and institutional trust in terms of the concepts of face validity, convergent validity and construct validity. Unlike most of the rest of the Guidelines, this section contains a substantial empirical component, which looks at the consistency of trust measures across different survey vehicles and over time.

2.2. A conceptual framework for measuring trust

To measure something, it is necessary to have a clear idea of the concept to be measured. Developing a sound conceptual framework is therefore an essential prerequisite to providing advice on the measurement of trust. A conceptual framework provides a working definition of the concept to be measured and sets out how the concept relates to other similar ones. In particular, a good conceptual framework should provide a clear definition of the concept to be measured and of any dimensions or components related to it.

This framework should be clear about what falls both inside and outside the scope for its measurement and relate the concept to specific measurement instruments.

In the case of trust, a sound conceptual framework is particularly important. Unlike some relatively straightforward concepts, such as age, gender or marital status, trust is inherently intangible. Although it is possible to observe trusting behaviour and to obtain self-reports from respondents about their stated levels of trust, it is not possible to directly observe trust as such. This raises the issue that respondents might not share a common view of what is meant with respect to trust in a survey question.

Theories of trust

There is an extensive literature on trust spanning a wide range of different disciplines within the social sciences, including political science, sociology, economics and psychology.[1]

In a review of what is known about measures of generalised trust, Nannestad (2008) identifies a wide range of definitions proposed by different authors who variously class the concept as a *general outlook on human nature*, an *affective attitude*, a *relationship*, a *decision* and an *action*. This highlights the degree to which there is not a single agreed definition of trust that can be adopted for these Guidelines. Nonetheless, several unifying themes cut across most measures of trust and provide a useful starting point. In particular, Hardin (2004) notes that trust is "in the cognitive category with knowledge and belief". This approach is echoed elsewhere: Morrone et al. (2009) note that expectations are central to most definitions of trust, and Uslaner (2008) identifies the central idea behind trust as a belief that most people share your moral values. A similar approach is shared by most other treatments of the subject, cutting across a wide range of different academic disciplines, from political science (e.g. Rothstein and Uslaner, 2005), to sociology (e.g. Delhey, Newton and Welzel, 2011) and economics (e.g. Zak and Knack, 2001).

A second common point shared by most theoretical approaches to trust is that they involve one person giving discretion to another party to affect the person's interests. There is thus an element of risk involved. This approach is spelled out explicitly by Hardin (2004), but is implicit in all accounts of the subject. Trust does not arise as an issue unless the person being trusted has the ability to affect the interests of the person doing the trusting materially or otherwise.

Beyond these points of commonality, however, there are significant differences in how trust is viewed. Nannestad (2008) identifies two main theories of trust: *rational trust* and *moralistic trust*. This distinction addresses the issue of whether trust should be thought of as a *belief* about other people's trustworthiness (rational trust) or as a *norm* regarding how to treat other people that is received from a person's parents, culture or environment (moralistic trust). In the latter case, it is assumed that trust is a belief about how a person should behave towards others (i.e. how much trust to extend to them) rather than a belief about how others are likely to behave towards you.

The strongest statement of rational trust is provided by Hardin (2004). Hardin defines trust in terms of a specific trusting relationship of the type **A** trusts **B** to do **X**, emphasising the strategic nature of trust. In particular, Hardin focuses on the notion of "encapsulated interest": in this formulation, trust occurs where the person or institution to be trusted has an incentive to be trustworthy because they internalise the interests of the person doing the trusting.[2] Their interest encapsulates the interest of the person trusting others. Because of the nature of this encapsulated interest – the person doing the trusting must

have good reason to believe that the person whom they trust has taken their interests on board – Hardin argues that trust is always "particularised". One must specify who is trusting, who is to be trusted, and what the person is to be trusted to do. In this sense, trust is strategic because it is conditional on a set of beliefs about the person being trusted and on their likely behaviour with respect to a given activity.

The main alternative to a rational notion of trust is the concept of moralistic trust. Uslaner (2008) argues that trust is inherited through socialisation rather than acquired. In this sense, trust is still an expectation about how others will behave, but it is not a strategic expectation. Trust is not grounded in the view that the person to be trusted encapsulates the interests of the person doing the trusting, but is rather a general attitude based on the life experiences and cultural background of the person doing the trusting. In Uslaner's formulation, trust is a "moral commandment to treat people as if they were trustworthy".

An important implication of the distinction between rational trust and moralistic trust lies in the role of reputation. In Hardin's formulation, trust becomes possible only in the existence of repeated interactions between the person doing the trusting and the person being trusted. Trust is rational in the sense that the trustor's expectation of good behaviour on the part of the person being trusted is based on either that person's reputation for good behaviour in the past or on a solid rational reason to expect that person's interests to align with those of the trustor. A restaurant owner, for example, might trust a supplier to provide the goods requested after being paid in advance because, if they are cheated by the supplier, then the supplier will lose all future business with the restaurant. The supplier's interests thus encapsulate the interests of the restaurant owner in this respect. This would appear to exclude an element of trust identified as important in much of the literature on the subject, i.e. the ability of trust to facilitate exchanges between people or to allow for people to co-ordinate actions in situations where they face each other in one-off situations in which their short-term incentives *do not* necessarily align. Many commercial and other exchanges between people are largely one-off and anonymous, effectively precluding trust as encapsulated interest, yet such exchanges still take place. In this perspective, Arrow (1972) stressed the importance of mutual confidence and trust in facilitating economic development.

In contrast, Uslaner's notion of moralistic trust allows trust to play a role in situations where the incentives of the trustor and of the person being trusted are not aligned. This approach is more in keeping with a view of trust that sees it as a mechanism for enabling strangers in a complex society to co-operate with each other. Fukuyama (1995) takes a similar view, emphasising the role of both repeated behaviour and shared social norms; in this approach, trust is understood as "the expectation that arises within a community of regular, honest and co-operative behaviour, based on commonly shared norms, on the part of the other members of that community."

Both rational and moralistic conceptions of trust define the concept as attitudinal, or in terms of the beliefs and expectations that people hold. An alternative definition of trust is provided by Fehr (2009), who argues for a behavioural definition of trust. Although Fehr notes that trust is tightly connected with preferences and beliefs, he argues that trust is best defined as a type of behaviour. In Fehr's approach, trust has two elements. The first element, which is unambiguously behavioural, is that the person who trusts places resources at the disposal of another party, without the means to guarantee that these will be returned. The second element is that the act of trust is associated with an expectation on the part of the person doing the trusting that the act of trust will be of direct benefit to them. This element

distinguishes trust from altruism and is associated with attitudes and beliefs rather than behaviour.

Conceptually, Fehr's behavioural notion of trust is compatible with either rational or moralistic trust. A behavioural view of trust abstracts entirely from the reasons why the person doing the trusting has an expectation of positive behaviour from the person who is trusted and focuses solely on the act of trusting itself. From the perspective of these Guidelines, this approach has the advantage that it does not pre-suppose a strong model of what creates trust. Instead, it treats trust as an empirical phenomenon that may be investigated independently of beliefs about how it is created.

Types of trust

All the theories of trust discussed above focus, in the first instance, on trust by people in other individuals. This is useful in establishing what is meant by trust conceptually. However, it is not sufficient to serve as the basis for measurement. There are obviously quite significant differences in asking about trust in another specific individual, trust in other people more generally or trust in a specific institution.

Table 2.1, below, provides one possible framework for thinking about how the different measures of trust fit together. This framework classifies measures of trust primarily in terms of the parties involved in the trusting relationship, and it has the advantage of capturing a very comprehensive range of situations. However, the framework also has some limitations in that it primarily focuses on distinguishing individual trust from the different elements of institutional and political trust.

Table 2.1. **A framework for multiple trust relationships**

By whom/on whom	Resident	Institution	Leaders
Resident	Interpersonal trust	Institutional trust	Political trust
Institutions	Civic trust	Inter-institutional trust	Political-administrative trust
Leaders	Political trust	Political-administrative trust	Multilateral trust

StatLink ⧉ *http://dx.doi.org/10.1787/888933584032*

Despite its limitations, the framework in Table 2.1 is useful for narrowing down the scope of these Guidelines. Some of the types of trust identified by Table 2.1 are not suitable for measurement in surveys (e.g. civic trust) or do not obviously fall within the remit of official statistics (e.g. political trust), or both (e.g. multilateral trust). Trust by institutions or organisations (as opposed to trust in institutions) is not suitable for measurement through surveys of the general population, and hence none of the measures of trust by institutions are covered by these Guidelines. Similarly, a household survey cannot specifically target political leaders as respondents, making this group out of scope even if there were no other reason to avoid collecting data of this sort. However, collecting information from political leaders is clearly a significant problem from the perspective of maintaining the independence of national statistical offices. For this reason, the scope of the Guidelines is kept relatively narrow, focusing on people's trust in individuals (i.e. **interpersonal trust**) and people's trust in institutions (i.e. **institutional trust**).

Some important distinctions highlighted by the literature on trust between different types of interpersonal trust are not well captured by Table 2.1. For example, Delhey et al. (2011) argue that, when framing survey questions, it is important to specify the relationship

between the person to be trusted and the person doing the trusting. For example, the level of people's trust in family or close friends might be expected to be different (narrow trust) compared to their trust in casual acquaintances or strangers (generalised trust). Delhey et al. also note that there are important differences across countries in the sort of relationship that people envisage when asked a generic question on trust in others. The degree to which these differences are important empirically will be explored later in the section of this paper on validity.

A similar set of issues to those raised by Delhey et al. (2011) can also be considered in the case of institutional trust. In particular, the specific institution that is being considered will have enormous implications for the degree of trust that people are ready to place in it.[3] While it is conceptually possible to consider institutional trust as a single construct, trust in specific institutions such as the police, courts, parliament or banks is likely to be of much greater practical relevance.

From a measurement perspective, the issue of people's trust in precisely which institutions or people should be measured is an important one. Obviously, users' needs provide an important guide as to which institutions or persons are relevant from a measurement perspective, but there are also empirical considerations. Respondents may not always clearly understand the differences between closely related institutions (e.g. between the legislative and the executive branch of government) or, even if they do, they may not have distinct views on all the institutions involved. Similarly, it matters a great deal whether responses to conceptually overlapping questions, such as those on trust in people you know personally as opposed to in neighbours or strangers, differ from each other. Evidence from the World Values Survey (WVS), which contains a wide range of questions on trust in different groups of people and institutions, suggests that, while there are differences in trust between different institutions and groups, there may be little practical value in making overly fine distinctions in trust questions (Box 2.1).

Box 2.1. **How many types of trust should be measured?**

The academic literature on trust includes a wide range of different measures capturing a diverse range of types of trust. Distinctions are made between different parts of government (e.g. parliament, the civil service, political parties, the current government) and between different types of persons (e.g. neighbours, personal acquaintances, family, people you meet for the first time). Many authors argue that such distinctions are fundamental to understanding trust meaningfully (e.g. Delhey, Newton and Welzel, 2011; Hardin, 2004). Such a view has obvious intuitive plausibility, as it is easy to imagine that trust in your family might be different from trust in someone you do not know personally. In practice, however, the main issue is whether the differences between these concepts are significant empirically. If respondents provide essentially the same answer to questions on trust in other family members and on trust in people they do not know personally, then there is little reason to collect information on the two concepts separately.

The World Values Survey provides a useful database for examining the degree to which different measures of trust capture different constructs. Wave 6 of the WVS contains questions on 18 different types of trust, with responses from 68 486 individuals, which allows an empirical analysis of how different responses group together. Table 2.2 shows the results of a principal component analysis of the WVS trust questions, aimed at identifying the degree to which the different questions collect fundamentally different information.

Box 2.1. **How many types of trust should be measured?** *(cont.)*

The table takes the 18 trust questions from wave 6 of the WVS and identifies the degree to which these reflect a smaller number of "latent" factors that best explain the variance in the dataset. For each trust measure, the higher the number in a particular column, the more strongly that measure is correlated with the latent factor associated with that column. Factor 2, for example, has high values for the government, political parties, parliament and the civil service, suggesting that the latent factor is related to political institutions. With a value of 0.4185 for factor 2, trust in the courts is clearly associated with trust in government institutions more broadly, to a much greater degree than trust in the armed forces (0.1342), but still not to the same degree as the core legislative and executive government institutions.

Table 2.2. **The dimensionality of trust measures in the World Values Survey**

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
The government (in your nation's capital)	0.1752	**0.7480**	0.0335	0.3271	0.0597
Political parties	0.2352	**0.8235**	0.0803	0.1339	0.0695
Parliament	0.2604	**0.8302**	0.0703	0.1655	0.0550
The civil service	0.3710	**0.6175**	0.0795	0.2346	0.0994
The armed forces	0.2348	0.1342	0.0597	**0.7278**	0.0983
The police	0.1687	0.3064	0.0662	**0.7708**	0.1152
The courts	0.2387	0.4185	0.0600	**0.6510**	0.0643
Major companies	**0.7160**	0.2356	0.0381	0.1722	0.0639
Banks	**0.6384**	0.2744	-0.0369	0.2069	0.0878
Universities	**0.6355**	0.2135	0.0764	0.2474	0.0470
Environmental organisations	**0.8024**	0.1589	0.0919	0.0753	0.0569
Women's organisations	**0.7783**	0.1101	0.1463	0.0629	0.0437
Your family	0.1084	0.0336	-0.0637	-0.0120	**0.7179**
Your neighbourhood	0.0415	0.1176	0.1914	0.1260	**0.7572**
People you know personally	0.0461	0.0336	0.3150	0.1434	**0.6838**
People you meet for the first time	0.0250	0.1079	**0.6301**	0.0688	0.3439
People of another religion	0.1019	0.0426	**0.8892**	0.0437	0.0549
People of another nationality	0.1230	0.0502	**0.8862**	0.0437	0.0610

Note: Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalisation. Total variance explained: 65.3%. Cronbach's Alpha: 0.8844.
Source: OECD calculations based on wave 6 of the World Values Survey (database) *www.worldvaluessurvey.org/WVSDocumentationWV6.jsp* as contained in the OECD Trust Database

StatLink 🔗 *http://dx.doi.org/10.1787/888933584051*

An analysis of Table 2.2 shows that 5 different factors account for 65% of the total variance (across individuals) of the 18 trust questions included in the WVS. These factors are: trust in non-governmental institutions (factor 1: major companies, banks, universities, environmental organisations, women's organisations); trust in political institutions (factor 2: government of the day, political parties, parliament, civil service); trust in law and order institutions (factor 4: armed forces, police, courts); trust in people known to the respondent (factor 5: family, neighbours, people you know); and trust in strangers (factor 3: people you meet for the first time, people of another nationality, people of another religion). These broad results are robust to different specifications such as looking only at institutional trust or only at interpersonal trust, and hold across different waves of the WVS.

Similar factor analyses have been performed by other researchers. Naef and Schupp (2009) focus on a three-factor model using the German Socio-Economic Panel and find that interpersonal trust has two distinct factors relating to people you know and to strangers,

> Box 2.1. **How many types of trust should be measured?** *(cont.)*
>
> respectively, while institutional trust forms a distinct third factor. Uslaner (2002) finds similar results, but adds the important detail that the standard Rosenberg question on generalised trust (Rosenberg, 1957; also see below) is associated with the factor relating to strangers rather than people you know. Schneider (2016), who undertakes a more focused analysis of 35 countries in Eastern Europe and the former Soviet Union to examine the dimensionality of institutional trust, finds a similar split to that identified in Table 2.2 between trust in political institutions and trust in law and order institutions, although she does not test for non-governmental organisations.
>
> This analysis confirms the salience of the conceptual distinction between interpersonal trust and institutional trust and identifies distinct sub-dimensions within each of these two main categories. However, it also highlights that many of the finer distinctions often made between different categories of trust are not very informative empirically. This does not necessarily imply that there is no value in asking more specific questions about trust: different trust questions may closely co-vary for reasons other than that they are measuring the same thing (e.g. they may have very similar drivers). If users have a sufficiently strong need, even relatively minor differences between closely related concepts may be important. However, it does suggest that, when deciding which trust measures are worth collecting, a relatively narrow range of measures will cover the most important aspects.

A *definition of trust*

To provide guidance on the measurement of trust, it is important to move from a description of different types of trust to a more specific definition that can provide a focus for the Guidelines. As discussed above, these Guidelines focus on trust **by individuals in other individuals** and **in institutions**. A broad definition of trust is adopted, one that is clear and intuitive and that can serve as the basis for breaking down the broader notion of trust into more specific categories. For the purposes of the Guidelines, trust is defined as:

> a person's belief that another person or institution will act consistently with their expectations of positive behaviour.

This definition has two key elements. First, it focuses on a person's beliefs or expectations about behaviour. This is consistent with all accounts of trust found in the academic literature. Second, it focuses on expectations of positive behaviour. In other words, trust implies not only that a person being trusted will act consistently with expectations, but that those expectations are of good behaviour. Although it might be possible to state, in casual usage of the term, that a known thief can be trusted to take advantage of an opportunity to steal something, this is clearly not the concept that the Guidelines are trying to measure.

Among the theories of trust outlined earlier, the definition used here is closest to that proposed by Fehr, in that it takes no position on the way in which trust is formed. The key difference between the definition adopted here and Fehr's purely behavioural definition is that this definition focuses on **belief** rather than **behaviour,** which allows for measurement in situations where behaviour cannot be observed. The definition can also be located in the framework set out in Table 2.1. It implies that, for the purposes of the Guidelines, trust is to be understood as trust **by individuals**, rather than by leaders or institutions. Conversely, the object of trust is wider, encompassing both other individuals and institutions. This reflects the fact that interpersonal trust and institutional trust are both policy-relevant.

In addition to distinguishing between interpersonal and institutional trust, it is also important to be precise about the specific categories of trust to be measured. Building on the factor analysis in Box 2.1, two categories of interpersonal trust as well as three types of institutional trust are identified. For interpersonal trust, these Guidelines distinguish between *generalised trust* and *limited trust*. Generalised trust refers to trust in people who are not known to the respondent or to trust in situations where the person being trusted is not specified. This is a concept of trust measured by the so-called Rosenberg question, first introduced in 1957, which asks, "generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" This can be contrasted with limited trust, which focuses on persons known to the respondent, including family, friends and people living in the respondent's immediate neighbourhood. Throughout the rest of these Guidelines, the term interpersonal trust will be used in situations that apply jointly to both generalised and limited trust, while the narrower terms will be used where there is a need to be more specific. As will be discussed later, generalised trust will be the primary focus for the Guidelines, due to its greater policy relevance and stronger evidence linking it to other outcomes of interest.

Institutional trust is used to refer to all types of trust in institutions, with trust in *political institutions*, *law and order institutions* and *non-governmental institutions* used to refer to a narrower concept. Although there is a high level of policy interest in trust in specific institutions, there is relatively little empirical evidence suggesting that respondents are able to make meaningful distinctions between these categories. For this reason, the Guidelines will make less use of these three categories than the distinction between generalised and limited trust.

In addition to the distinction between different types of institutions, there is a second important conceptual distinction to make with respect to institutional trust. A common theme in the academic literature on institutional trust (e.g. Nooteboom, 2007) is the distinction between "trust in competence" (i.e. whether the functioning of institutions matches people's expectations about the competencies of those steering them) and "trust in intentions" (which captures whether institutions act in a way that is perceived by people as ethical and fair). These distinctions are extended by Bouckaert (2012), who distinguishes between the "logic of consequences", where trust is derived causally from outcomes, and the "logic of appropriateness", where trust is based on the values of integrity and transparency. This distinction between the outcomes of an action and the intention that guided it forms the basis of the OECD trust framework, endorsed by the Organisation's Public Governance Committee. The framework identifies two key components of trust in institutions: *Competence* and *Values*. Within each component, relevant dimensions that are amenable to policy change are identified based on the common threads in the literature (Mcknight, Choudhury and Kacmar, 2002) and on the OECD update of this evidence (OECD, 2017a). The five dimensions identified are: *Responsiveness, Reliability, Integrity, Openness* and *Fairness*.

2.3. Approaches to measuring trust

Defining trust is a necessary precondition for measuring it, but it is not sufficient. To provide meaningful guidance on the best way of measuring trust, it is necessary to understand the various approaches to its measurement.

There is no single framework for classifying the different approaches to measuring trust in the academic literature. Nonetheless, drawing on a range of sources, it is possible to identify several distinct measurement approaches. At the most basic level, a long tradition of survey questions has directly asked people about their trust in others (e.g. Almond and

Verba, 1963) and institutions (e.g. WVS). Morrone, Tontoranelli and Ranuzzi (2009) make a distinction between traditional trust questions and attempts to measure trust through peoples' expectations about the behaviour of others (specifically, a survey question on whether a lost wallet is likely to be returned). Although the use of such "expectations questions", drawing on specific hypothetical scenarios is, so far, relatively limited, they set a distinctly different conceptual task for respondents than direct questions about trust, and thus might represent a source of additional information.

Beyond this, a large literature has compared actual trusting behaviour in experimental settings with survey questions on trust (Glaeser et al., 2000; Fehr et al., 2003; Gachter et al., 2004; Lazzarini et al., 2004; Naef and Schupp, 2009; Johnson and Mislin, 2012; Algan and Cahuc, 2013; Falk et al., 2016). Although experimental approaches to measuring trust might be regarded as beyond the scope of official statistics, these measures provide important insight into the validity of more conventional survey-based measures, and they have been used as a basis to develop better survey questions (Falk et al., 2015). Finally, it is important to note the existence of a wide suite of questions grounded in people's experiences that, while not focused directly on trust, can provide information on the subject. Taken together, these different approaches to measuring trust can be organised into four broad groups:

- evaluations
- expectations
- experiences
- experiments.

The different approaches to trust identified by these four groups do not represent substantively different concepts of trust, in the same way that, for example, *evaluative, affective* and *eudaimonic* measures capture substantially different constructs of subjective well-being (OECD, 2013a). Rather, they correspond to different approaches to measuring the same concept. The first three approaches are all survey-based and vary in the focus of the question asked to respondents. In particular, while evaluative questions focus on the respondent's views at the time of the interview, expectations questions have a future orientation, and questions on experiences focus on the past. In contrast, an experimental approach to measuring trust captures actual behaviour from respondents in a situation that is designed to elicit trusting (or non-trusting) behaviour.

Evaluations

Evaluative approaches to measuring trust focus on the respondent's response to questions on whether they have trust in an individual or institution. These questions ask the respondent to make an evaluation of their own feelings and/or beliefs at the current point in time, rather than recalling information about past experiences or speculating about the future. Most of the commonly used survey questions related to trust are evaluative. An example of an evaluative question is the standard generalised trust question by Rosenberg that is used in the WVS.

"Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?"

Expectations

An alternative to asking for people's evaluation of how they feel is to ask them about their expectations of what would or will happen in a given, hypothetical situation. These

expectations questions are occasionally described as situational. Because these questions focus on expectations of how people will behave, it is sometimes argued that they are more specific and behavioural than is the case for evaluative questions (Morrone, Tontoranelli and Ranuzzi, 2009). However, the hypothetical nature of the question is also a potential weakness, as respondents may have difficulty providing an accurate answer to a hypothetical situation. The *lost wallet* question used in the Gallup World Poll is an example of an expectations question:

> "If you lost a wallet or a purse that contained items of great value to you, and it was found by a stranger, do you think it would be returned with its contents, or not?"

Experiences

An alternative to asking about what people expect to happen in a given situation is to ask the respondent about past experiences. Such questions do not focus on what the respondent feels or believes but rather on concrete events that have already taken place. This approach is the one customarily used for victimisation surveys, which often include questions related to experiences that represent low levels of trust, such as how often the respondent is fearful in their local town or city centre (e.g. New Zealand Crime and Safety Survey, 2014). Similarly, in the context of institutional trust, questions relating to people's experiences of corruption or discrimination have been used in both research and official surveys. Experience questions have not traditionally been a major source of data on trust, but are included here as they have the potential to provide information on the degree to which respondents have experienced situations where their expectations of (positive) behaviour from others have or have not been met. An example of this is the *experience of discrimination* question, which reflects a violation of social norms relating to the expectation of fair treatment. For instance, the German Satisfaction with Government Services Survey (Zufriendenheitsbefragung, 2015) asks respondents to assess their experience with a range of individual public agencies, including whether the respondent trusts the agency and felt discriminated against during the interaction with the institution. Similarly, the New Zealand General Social Survey includes the following question on the experience of discrimination:

> "The next question is about discrimination in New Zealand. By discrimination I mean being treated unfairly or differently compared to other people. In the last 12 months have you been discriminated against?"

Experiments

An alternative to collecting information on trust through surveys is to use experimental techniques to measure trusting and trustworthy[4] behaviour by participants, either in the field or in controlled laboratory conditions. Pioneered by experimental and behavioural economists and more recently by psychologists and neuroscientists interested in the cognitive processes underlying decision making, games such as the Trust Game (Berg et al., 1995)[5] collect information on the respondent's actual behaviour in circumstances where genuine rewards – typically, but not always, relatively small (yet large enough to form genuine incentives) – are at stake (Box 2.2). Because they measure actual behaviour, experimental methods avoid many of the limitations of survey-based approaches to measuring trust, such as social desirability bias in responses, or different understandings of the survey question between respondents.

Box 2.2. **Experiments for measuring trust**

Experiments can either take place in the real world or in the controlled setting of a laboratory, which is typically a fixed physical or virtual infrastructure that participants visit to undertake carefully scripted decision-making tasks. One approach to experimentally measuring trust in the field is illustrated by the 1996 experiment by the *Reader's Digest* (Felte, 2001). A number of wallets, each containing USD 50 in local currency and a fictitious ID, were "lost" in public places in a selection of major cities around the world. The proportion of wallets returned was then used by *Reader's Digest* to construct an index of trustworthiness for the different cities involved. This experiment was recently repeated in 2001, with over 1 100 wallets lost in different cities (Felte, 2001). The results were largely consistent with the picture provided by survey data. A high proportion of wallets were returned in the Nordic countries, New Zealand, Korea and Japan, while relatively few wallets were returned in Italy, Argentina and Mexico. A similar experiment was run in 2007 but with lost cell phones replacing wallets (Shanahan, 2007).

While the *Reader's Digest* field experiments with wallets and cell phones are intuitively appealing, they have a number of significant limitations. In particular, the number of wallets/cell phones involved in the experiments is low (10 wallets per city), and it is difficult to be confident that the proportion of lost objects returned is not influenced by differences in the particular areas where the objects were abandoned in the different cities. Although the *Reader's Digest* experiments have been influential, leading the Gallup World Poll to use a lost wallet question (see above, under "Expectations"), the significant costs and methodological limitations associated with "losing" large numbers of wallets in different cities mean that this approach has never become a major source of trust data.

In the laboratory, a more systematic and rigorous approach to testing trust and trusting behaviour experimentally was developed by Berg, Dickhaut and McCabe (1995). The Trust Game has since become the standard approach to measuring trust experimentally in controlled settings, and it has formed the basis for a wide range of different variant games testing different aspects of trust and trusting behaviour. The game has now been repeated in dozens of countries involving tens of thousands of individuals, allowing for large-scale meta-analyses of the results (e.g. Johnson and Mislin, 2011; Johnson and Mislin, 2012; Zelmer, 2003). Although conducted under controlled conditions, and typically for rather small stakes, the main findings of the Trust Game have been replicated with relatively high stakes, and have been found not to be affected by social desirability bias or by the range of strategies allowed to participants (Naef and Schupp, 2009). There is today a growing body of literature examining the relationship between laboratory results from the Trust Game and outcomes in the real world.

The original Trust Game developed by Berg et al. involved two anonymous players in different rooms of a laboratory. Each player starts with an endowment of USD 10 of real money provided by the experimenters. Player 1 is then invited to give some (or all) of their endowment to player 2. Any amount given by player 1 is tripled by the experimenters so that if, for example, player 1 gives USD 5 to player 2, the second player will receive USD 15. After the first transaction, player 2 is then provided with the opportunity to return some (or all) of their money to player 1. This amount is not adjusted by the experimenters in any way. Given the structure of the experiment, the total payoff to both players is maximised when player 1 gives all his/her money to player 2, in which case the total payoff would be USD 40 (USD 10 from player 1, tripled to USD 30, plus the USD 10 original endowment of player 2); if player 2 is trustworthy, and sends half their money back to player 1 at the end of the game, this would allow each player to walk away with USD 20, as opposed to their original endowment of only USD 10 each. However, player 1 has no way to ensure that player 2 responds in this way so that, from the perspective of player 1, there is always a risk that any money given to player 2 is simply gone.

The Trust Game produces two measures of trust. First, the proportion of money given by player 1 to player 2 can be taken as a measure of trust by player 1 that player 2 will reciprocate any money sent to him or her and will respond in kind. Second, the proportion of money returned by player 2 to player 1 can be interpreted as a measure of trustworthiness

of player 2. Repeating the game over a representative sample of the population allows estimating the level of trust and trustworthiness among participants.

However, experimental games also have limitations and might be less relevant when the goal is to measure trust among the wider population. Lab games can be expensive to run, and most laboratory studies have so far been restricted to university students, thus basing their findings on small samples that, apart from being unrepresentative of the whole population, can be affected by self-selection bias.[6] This is a significant limit when it comes to establishing the external validity of the findings of many laboratory experiments. Algan and Cahuc (2013) review a range of studies indicating that experimental behaviour in a laboratory setting is more effective than survey responses in predicting a range of real-world outcomes. Further encouraging evidence on the external validity of lab experiments comes from Naef and Schupp's (2009) finding that both survey and experimental measures of trust correlate with related factors such as being an entrepreneur or a shareholder. Being an entrepreneur increases exposure to risk, particularly to the risks associated with informal agreements with other people (Guiso et al., 2006), while investing in the stock market implies trust in both the stock market system as a whole and in the executive management of the companies being invested in (Guiso et al., 2008). Furthermore, Karlan (2005), in a study of borrowers in a Peruvian microcredit programme, shows that behaviour in a Trust Game predicted repayment rates of the participants' loans one year later.[7]

While experimental techniques such as the Trust Game may not be suitable for direct inclusion in most large-scale sample surveys run by national statistical offices, they are relevant to these Guidelines in two ways. First, experimental data provide one of the best sources of information for testing the validity of different survey measures and for developing better survey questions (Fehr et al., 2003; Naef and Schupp, 2009; Falk et al., 2015). This is important given the challenges involved in testing the validity of measures of intangible concepts such as trust. Second, although national statistical offices may not be well positioned to collect experimental data in large sample surveys, this does not mean that such an approach cannot be pursued. The *Trustlab* project aims to do exactly this for a number of OECD countries: over time, the evidence gathered through this project will provide a rich base of information on trust collected on a comparable basis from a representative population sample across the OECD (Box 2.4).

2.4. The statistical quality of trust measures

For measures of trust to be taken seriously, statistical quality is crucial. Unless data accurately capture the concept being measured, there is little point in collecting it. This is particularly true for official statistics, which are expected to be of the highest quality. The United Nations Fundamental Principles of Official Statistics state that "official statistics provide an indispensable element in the information system of a society, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation" (United Nations, 2014). It is therefore crucial that these Guidelines are solidly grounded in a clear understanding of the quality of trust measures.

The OECD's approach to dealing with issues of statistical quality is set out in the *Quality Framework and Guidelines for OECD Statistical Activities* (OECD, 2008). The Framework defines quality as *fitness for use* in terms of user needs. The ultimate benchmark of the quality of statistics is whether they meet the needs of the user in terms of providing useful information. Because users must often make decisions about a course of action, whether or not statistical information is available, it is helpful to focus on the *fitness for purpose* of

measures rather than on whether the measure in question provides a "perfect" representation of the concept that it is intended to capture. This may involve accepting the use of data that are less than perfectly accurate, provided that they are of sufficient quality to improve on rather than detract from the quality of decision making.

The OECD framework identifies seven dimensions of statistical quality. These seven dimensions define the characteristics of high-quality data and provide a structured way of assessing the quality of a particular set of statistics. These dimensions are:

- *Relevance*, i.e. the degree to which data address the purposes for which they are sought by users.
- *Accuracy*, i.e. the degree to which data correctly describe the quantities or characteristics that they are designed to measure.
- *Credibility*, i.e. the confidence that users place in statistics based on the reputation of the data producer.
- *Timeliness*, i.e. the length of time between the availability of data and the phenomenon or event that the data describe.
- *Accessibility*, i.e. how readily the data can be located and retrieved by users.
- *Interpretability*, i.e. the ease with which users can understand and properly use and analyse the data.
- *Coherence*, i.e. the degree to which the data are mutually consistent with other similar measures and logically integrated into a system of statistics.

These seven criteria, along with the more general principle of cost effectiveness in producing/collecting such data, define the OECD's overall framework for assessing statistical quality. However, most of these criteria relate to **how** statistics are measured and collected rather than to **what** is collected. Credibility, for example, is a quality of the data producer, rather than of the measure. Similarly, timeliness, availability, interpretability and coherence largely relate to the broader context in which data are collected and made available to the public, rather than to the choice of measures used. For the purposes of these Guidelines, the concern of statistical quality is more narrowly focused on what should be collected and whether the proposed measurement instruments are fit for purpose. Thus, the main focus for assessing the quality of measures of trust will be the principles of relevance and accuracy.

Relevance

Data are relevant when they meet the needs of users. In the case of official statistics, where there may be a legal requirement for respondents to participate in data collection, it is of particular importance that information is collected only where there is a clear use for the data and where there are no obvious alternative information sources. For this reason, it is essential to identify the purposes of trust data before developing advice on measurement. Clarity about the relevance of trust data helps both to identify which trust measures are important to collect and to provide assurance that trust measures are worth collecting in the first place.

Modern societies are increasingly complex, and ill-informed policy choices may incur high costs and be difficult to reverse. Information about how key societal outcomes are changing, the drivers of those changes, and the impact of policy on these outcomes is therefore essential to good decision making. The ability of data to inform policy is thus central to relevance, particularly for national statistical offices. Beyond this, there is a broader public interest in accessing information on key social and economic outcomes.

This interest is distinct from that of policy making and is grounded in the value of an informed public. Trust data can therefore be considered to be relevant if they inform policy or the wider public.

Two core uses of trust data can be identified. These are:

- monitoring progress (current levels of people's well-being and sustainable development for the future)
- evaluating the effectiveness of government policies.

These two core uses of trust data are of relevance both from the perspective of policy making and from that of the general public. Policy makers have a clear interest in being able to evaluate whether their decisions are resulting in the desired outcomes and in understanding how effective different government institutions are at implementing policy. However, assessing sustainability, well-being and the effectiveness of government are also of interest to the general public from the perspective of democratic accountability.

Monitoring progress (current levels of people's well-being and sustainable development for the future)

The past decade has witnessed an increasing consensus on the need to measure the outcomes that are the ultimate goals of policy (well-being) and to account for all of the capital stocks that contribute to the sustainability of those outcomes over time. This two-part model, focusing on flows of population well-being underpinned by stocks of natural, economic, human and social capital, is at the core of the modern consensus on the measurement of well-being and of sustainable development (Stiglitz, Sen, and Fitoussi, 2009; OECD, 2011; UNECE, 2014). Central to this approach is a multidimensional view of people's well-being, seen as encompassing all those different outcomes – objective and subjective – that are valued by people in and of themselves or that are fundamental to allowing people to pursue the sorts of lives they have reason to value. This perspective is complemented by the recognition that measuring changes in the capital stocks used to produce well-being is essential to understanding sustainability. This perspective is reflected in reports by international organisations such as the OECD's *How's Life?* (2011, 2013c, 2015) and, most recently, in the United Nations Sustainable Development Goals (SDGs) (UN, 2015; OECD, 2017b). National initiatives to measure well-being and sustainable development also reflect this general framework, including those pursued in the United Kingdom (ONS, 2016), Israel (Ministry for Environmental Protection, 2014), Italy (ISTAT, 2013) and many other countries. Both of the two dimensions of trust covered by these Guidelines – interpersonal and institutional trust – are of vital importance to measuring well-being and sustainability.

The United Nations Sustainable Development Goals (SDGs)

A major task for national statistical offices over the next decade will be measuring progress towards the SDGs. These goals set out an ambitious plan of action for *people, planet, prosperity* and *peace*, to be achieved through *partnership*, and with the overarching objective of leaving no one behind. At the core of this plan are 17 SDGs comprising 169 targets. The 2030 Agenda set out by the United Nations (UN, 2015) allows countries to take their own approach to measuring progress towards the SDGs, but "encourage[s] member states to conduct regular and inclusive reviews of progress at the national and sub-national levels". While this resolution allows for some discretion in how to measure progress, it pushes nations very strongly towards measuring the SDGs in some kind of way.

Trust – both institutional and generalised trust – is central to monitoring progress towards the SDGs. In particular, Goal 16 of the SDGs ("Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable, and inclusive institutions at all levels") focuses explicitly on governance, institutional quality and political participation. These themes have clear links to trust, and this is recognised in the work of the Praia Group on Governance Statistics. The explicit aim of the Praia Group is to improve the measurement of governance and encourage the production of internationally comparable statistics on governance with a view to informing Goal 16 of the SDGs.[8] This will involve a work programme covering different elements of governance, including citizens' trust in public institutions, their experiences of the performance of public institutions, and measures of generalised trust.

The Praia Group's deliberation of institutional trust as a potential indicator for Goal 16 of the SDGs is based primarily on its importance as a direct measure of institutional quality (Klijn, Edelenbos and Steijn, 2010). Institutional trust is identified in the OECD's *How's Life?* report (OECD, 2011) as an important aspect of civic engagement and governance. In discussing the statistical agenda ahead for this dimension, the OECD (2011) identified a number of actions aimed at improving measurement in this area:

> …better measures are required to assess how people perceive the quality of the democratic institutions in the country where they live. While many unofficial surveys contain questions for measuring civic engagement and governance, they typically have a narrow geographic coverage, small sample size and inadequate sampling procedures. Steps should be taken to include questions on civic participation and trust in institutions in large-scale official surveys, through a combination of a few recurrent questions in regular surveys and more detailed questions in (less frequent) dedicated modules, as is already done in some OECD countries (e.g. special modules of the Current Population Survey on voting and civic engagement in the United States).

The OECD decision to use measures of institutional trust among the indicators of the civic engagement and governance dimension of well-being is supported by analyses of the drivers of subjective well-being. Boarini et al. (2012) find that confidence in the judicial system has a significant positive impact both on people's satisfaction with their own life and on their affect balance, after controlling for a wide range of other potential drivers of well-being.[9] With respect to the effect on life satisfaction, the magnitude of this impact is reported as 1.1 times the impact that a doubling of individual income would have. Similarly, Helliwell and Putnam (2004) find a robust association between trust in the police and life satisfaction, with similar results replicated with different datasets in Helliwell and Wang (2010).

Generalised trust is also highly relevant as a measure of the quality of governance (OECD, 2011). This is because there are feedback loops from institutional quality to levels of generalised trust in society. Rothstein and Uslaner (2005) argue that high levels of generalised trust depend crucially on the effective functioning of governance and justice institutions. It is only when these institutions are performing both fairly and competently that it is possible for a person to extend trust to strangers without putting themselves at risk. Furthermore, there is also a causal pathway running in the opposite direction, i.e. from high levels of generalised trust to fair and effective public institutions. In particular, public institutions are likely to be fairer and more effective in societies where individual civil servants have reason to believe that other officials and citizens are trustworthy, and where the cost of corrupt or unfair behaviour is likely to be high and the risk of being

caught out in such behaviour is more probable. For this reason, even though generalised trust itself is conceptually distinct from institutional trust and from governance more generally, it is nonetheless a powerful indicator of the quality of governance.

Measuring social capital for sustainable development

In addition to contributing to monitoring Goal 16 of the SDGs, measures of trust are of importance for monitoring the stocks of capital that underpin the production of well-being. The OECD and the UNECE (UNECE, 2014) define sustainability as a situation in which current levels of well-being occur with stable or increasing capital stocks. This implies the need for adequate measures of the capital stocks underpinning sustainable development – economic capital, natural capital, human capital and social capital (OECD, 2011). Social capital is used here in a narrower sense than in some of the wider literature. The term social capital has been applied to a number of different concepts, including the networks and values that enable elites to transmit social advantage across generations (Bourdieu, 1984), the resources inherent in networks that allow people to pursue individual outcomes (Coleman, 1988), and socially useful norms and values (Putnam, 1993). But it is only the latter concept that is relevant to the measurement of sustainable development. The OECD uses a variant of this definition when it describes social capital as the "shared norms, values and understandings that facilitate co-operation within or among groups" (OECD, 2001).

Although social capital is important, as an intangible capital stock it is difficult to measure. Residual-based estimates of the importance of intangible capitals (i.e. human and social capital) suggest that the magnitude of the effect of these intangible stocks on other outcomes is large (World Bank, 2006; Hamilton and Liu, 2013). In order to move beyond the general observation that social capital is important in producing other valued outcomes to useful policy advice, it is essential to have direct measures of the size of the social capital stock and of how it changes over time. Generalised trust is one of the best available proxy measures of social capital, and better measures of generalised trust would make a very significant contribution to understanding social capital, its drivers and consequences (Scrivens and Smith, 2013). The role of generalised trust in measuring social capital is also stressed by other authors (e.g. Knack and Keefer, 1997; Helliwell and Putnam, 2004; Morrone, Tontoranelli and Ranuzzi, 2009; Sturgis and Smith, 2010; Algan and Cahuc, 2013). However, for generalised trust to inform policy, adequate information must be available. Existing analyses are largely static and cross-sectional in nature, and provide only limited information on trends in stocks of social capital or on the impact of policy changes on this stock. Key limitations of existing data are the relatively small sample size of existing non-official surveys such as the WVS and the lack of sufficiently regular data to enable meaningful analysis of changes in interpersonal trust over time.

Indicators of well-being

Interpersonal trust also has an obvious relevance as an indicator of well-being. It is unpleasant to live in a situation where a person feels that he/she cannot trust the people around them, and low levels of trust in others place important limits on how a person lives their life. If well-being is viewed in terms of the capabilities a person has to live the sort of life they have reason to value (Sen, 1992 and 2009), then trust in other members of society is in itself an important element of well-being. In particular, the level of trust a person has in his/her fellow citizens shapes the sort of choices that a person might feel comfortable making (Alsop et al., 2006). With this rationale in mind, the OECD has chosen generalised

trust as one indicator of its "social connections" dimension in *How's Life?* (OECD, 2011). Similarly, from a more utilitarian perspective on well-being, interpersonal trust is one of the main drivers of subjective well-being.

The importance of interpersonal trust for assessing well-being is highlighted by the *World Happiness Report* (Helliwell et al., 2016), which presents an analysis of the main drivers of country differences in subjective well-being across the globe. Although interpersonal trust itself is not included in the analysis, as it is not part of the dataset on which the *World Happiness Report* is based (the Gallup World Poll), access to social support – which is closely correlated with trust in others – is one of the six main determinants of cross-country differences in life satisfaction identified in the report.

More directly, using data from the Gallup World Poll and the Canadian General Social Survey, Helliwell and Wang (2010) investigate the relationship between generalised trust and subjective well-being in more depth. After controlling for income and a range of social and demographic factors, the authors find a robust relationship between generalised trust and people's life satisfaction; this positive relation holds regardless of whether trust is measured via the traditional (Rosenberg) question on generalised trust or via a question about whether a lost wallet would be returned. Boarini et al. (2012) use data from the Gallup World Poll to test whether the dimensions of well-being used in the OECD Better Life Initiative correlate with overall life evaluations: they find that the average level of generalised trust has a significant positive impact on life satisfaction, after controlling for other dimensions of well-being. Fleche, Smith and Sorsa (2012) report similar results in a cross-country analysis using data from the WVS.

However, it is not only generalised trust that matters directly for well-being. There is also reason to believe that institutional trust plays an important role in well-being over and above its instrumental value in achieving other outcomes. While institutional trust is certainly important in this regard, the justification for measuring institutional trust here is its intrinsic importance to people's well-being. This is explored by Frey, Benz and Stutzer (2004) and Frey and Stutzer (2005, 2006), who highlight the importance to well-being of "procedural utility", i.e. the process through which important decisions are made. This approach suggests that procedural utility has an important impact on people's well-being, independently of the resulting outcome. In this perspective, the welfare gains from a policy decision that increases total income can be partially or totally offset by losses resulting from a decision-making process perceived by people as unfair or illegitimate. This approach strongly underscores the importance of institutional structures for people's well-being. The argument about the importance to people of being part of the process through which collective decisions are made and enforced is echoed by Alkire (2002), who surveys a wide range of international initiatives to define the key dimensions of "human flourishing", concluding that items relating to political participation and processes feature in nearly all of these.

Evaluating government effectiveness

While generalised trust is of primary importance for measuring social capital, institutional trust is most relevant to evaluating the effectiveness of government policies and programmes. Institutional trust is fundamental to the effective functioning of institutions and networks (e.g. Klijn, Edelenbos and Steijn, 2010). When people have a high level of trust in institutions, compliance is less of a problem, and it is easier to implement policies that may involve trade-offs between the short and long-term, or between different parts of society (Marien and Hooghe, 2011; OECD, 2013b). Institutional trust is especially important to

government activities that address market failures (e.g. public health or environmental protection) or where long-term gains require short-term sacrifices (e.g. education and pensions). In most of these cases, individuals cannot obtain an immediate, exclusive benefit from policy actions, and there is a high risk of free-riding. In these cases, trust in institutions may encourage people to accept some personal sacrifice in the short run for the sake of a larger benefit in the future, for themselves or for the community at large. Trust in institutions may be important also in cases, such as obtaining a business license, where decisions are influenced by the reliability of the rules that frame decisions or by the reliability of the institutions that make these rules. Thus, while high levels of institutional trust reduce government failures and the costs of enforcement and transaction, a lack of institutional trust provides incentives to opportunism, informality and free-riding, compromising the effectiveness of public policies.

A significant body of research has looked at the relationship between trust in institutions and government effectiveness. This research is based largely on unofficial data, but identifies some of the key elements in the relationship between institutional trust and government performance. One key theme here is the view that institutional trust depends on the congruence between people's (and businesses') preferences (i.e. their interpretation of what is right and fair) and their perceptions of how the government functions (Bouckaert and van de Walle, 2003). While a core theme cutting across this literature is the value of institutional trust as a measure of government effectiveness, this research also highlights the need for more detailed information. In particular, better measurement of institutional trust requires not only narrowing down the scope of the trust relationship to be measured, but also using a more operational notion of trust that can be deconstructed into meaningful policy questions.

Accuracy

Accuracy is of fundamental importance to any statistical measure. If a proposed measure does not correctly reflect the underlying concept that it is intended to capture, then it fails the basic test of providing useful information. Accuracy is typically assessed by looking at two dimensions of the measure, i.e. its **reliability** and **validity**. The reliability captures the degree to which a measure produces consistent information over time and across different measurement vehicles: a measure that takes different numerical values when it is repeated over short time periods is clearly unreliable, and fails a basic test of accuracy. Reliability is thus concerned with the *variance* of the measure. Validity is concerned not with variance of the measure but rather with its *central tendency*, i.e. does the measure actually reflect the underlying concept to be measured? A measure can be valid but not reliable if it produces the correct result on average, but with a wide variance. Similarly, a measure can be reliable but invalid, when it captures items unrelated to the variable of interest.

Validity is usually analysed in terms of face validity (whether the measure makes sense intuitively), convergent validity (whether the measure correlates well with other proxy measures of the same concept) and construct validity (whether the measure behaves as theory and common sense dictate). Where validity cannot be directly assessed, a measure can be considered valid if it performs well in terms of all three aspects of validity outlined above.[10] Ideally, this would be assessed through a large and well-developed scientific literature that covers all the main types of validity, and which is sufficiently mature for a consensus to have emerged.

In some cases, a large and developed body of literature for a measurement concept will already exist. For example, the *OECD Guidelines on Subjective Well-being* were based on an extensive body of research by experimental psychologists on the validity of different types of subjective well-being measures. This reflected more than a decade of intense interest in the measurement of subjective well-being from a large number of highly regarded economists, psychologists and sociologists. Unfortunately, the same is not true with respect to the measurement of trust. While some excellent academic work on trust exists, taken as a whole, this body of work lacks the size, rigour and methodological focus that characterised work on the validity of subjective well-being measures. Because of this, a substantial focus for these Guidelines is to assess the validity of trust measures; this involves an empirical component to the work that was not present in the *OECD Guidelines on Measuring Subjective Well-being*. This empirical work, which is summarised in this chapter, is based largely on the OECD Trust Database (Box 2.3).

Box 2.3. **The OECD Trust Database**

A key challenge in building a better understanding of the drivers of trust and its impact on other outcomes is the limitations of the available data. This limitation has also had an important impact on the ability to assess the validity and reliability of trust data rigorously.

For a long time, the only substantial source of cross-country trust data was the World Values Survey (WVS). Recently, a wider range of non-official data have become available. The OECD Trust Database is an effort to map existing sources of (so far non-official, apart from EU-SILC) data and compile them into a single repository of information. Table 2.3 displays the different surveys included in the OECD Trust Database, which has been constructed as a cross-country panel dataset. Its coverage goes beyond OECD member states and includes up to 124 countries depending on the survey considered, spanning the period between 2002 and 2015 or the latest year available. The year 2002 was chosen as the database's starting point, since the number of surveys regularly collecting trust data doubled with the appearance of the European Social Survey (ESS) in that year.

Table 2.3. **Surveys included in the OECD Trust Database and their key characteristics**

Survey	Inception	Frequency	Number of countries in the OECD Trust Dataset	Coverage of the OECD Trust Dataset
EU Statistics on Income and Living Conditions (EU-SILC)	2003	2013 ad-hoc module	33	2013
Gallup World Poll (GWP)	2006	Yearly	115	2006-2015
World Value Survey (WVS)	1981	Every 5 years	6 (Wave 4) 46 (Wave 5) 45 (Wave 6)	1999-2004 (Wave 4) 2005-2009 (Wave 5) 2010-2014 (Wave 6)
European Social Survey (ESS)	2002	Every 2 years	22 (Round 1) 25 (Round 2) 23 (Round 3) 28 (Round 4) 27 (Round 5) 29 (Round 6)	2002 (Round 1) 2004 (Round 2) 2006 (Round 3) 2008 (Round 4) 2010 (Round 5) 2012 (Round 6)
European Quality of Life Survey (EQLS)	2003	Every 3 years	31 (Round 2) 34 (Round 3)	2007-2008 (Round 2) 2011-12 (Round 3)
Eurobarometer	1973	Yearly	34	2003-2015
Latinobarometer	1995	Yearly	19	2002-2015

StatLink ⬚🔗 *http://dx.doi.org/10.1787/888933584070*

Box 2.3. **The OECD Trust Database** (cont.)

Different surveys have different geographical coverages and collection frequencies. In the case of the Gallup World Poll, data are collected annually for countries in all regions of the world, while for other surveys – including the WVS, ESS and the European Quality of Life Survey (EQLS) – the coverage is limited to a smaller set of countries. Data collection takes place every 2 years for the ESS, every 3 years for the EQLS and roughly every 5 years for the WVS.

Both institutional and interpersonal trust feature in the OECD Trust Database. In the case of institutional trust, questions are traditionally formulated through a common heading (e.g. *do you have confidence in your…*) followed by a list of primarily public institutions (e.g. government, congress, etc.) and less commonly private (e.g. major companies). Survey wording varies considerably, both in terms of the general construction of the question and in the use of the term trust or one of its various synonyms (e.g. confidence). Also, different surveys, or even different questions within the same survey, make use of different response scales. For example, while the GWP Poll relies primarily on a *yes/no/don't know* response format, other surveys such as the ESS and EQLS use longer numeric scales (0-10 and 1-10 in the case of ESS and EQLS, respectively). In the case of the WVS, questions are usually answered using a 4-point Likert scale (i.e. *a great deal*, *quite a lot*, *not very much* and *none at all*). In the OECD Trust Database, different questions have been re-scaled to a binary *yes/no* format that allows comparability across surveys.

Further observation of the available data on institutional trust reveals that questions sometimes refer to similar concepts while using quite different descriptions. For instance, some surveys refer simply to *the courts* while others ask about the *judicial system*. While in most cases the interpretation of these concepts is straightforward, in others the lack of clarity may have more significant implications. For example, although most surveys ask about trust in government, the ESS asks about trust in politicians, and EU-SILC addresses trust in the political system. In turn, the Latinobarometer has included questions about trust in government, the state and public administration, concepts traditionally related to each other but not strictly synonymous. The EQLS includes a few questions worded in a more comprehensive way: instead of asking about a particular institution, the EQLS has extended the question to a system (e.g. state pension system or social benefit system) that includes a set of institutions.

Generally, non-official household surveys provide less information on interpersonal trust compared to institutional trust and, where interpersonal trust is included, the focus is limited to generalised trust. The most common question (asked by four of the surveys under study), which is very similar to the version introduced by Rosenberg in 1957, is the following:

"Generally speaking, would you say that most people can be trusted or you can't be too careful in dealing with other people?"

In the case of the WVS, in addition to *most people*, the questions inquire about additional parties to be trusted (e.g. family, neighbourhood, known people, people met for the first time, etc.). Only the EQLS asks a direct question about trusting people in general.

The surveys included in the OECD Trust Database vary in terms of data quality. While all of the surveys have different strengths and weaknesses, some of them place a greater emphasis on methodological rigour and consistency than do others. Both the ESS and EQLS are directly funded by the European Commission and aspire to very high standards of data quality. A great deal of attention is paid to consistency across countries, and changes between waves

> ### Box 2.3. **The OECD Trust Database** (cont.)
>
> are carefully managed. Covering a much greater range of countries, the Gallup World Poll varies more in survey content from wave to wave, but retains a strong focus on methodological consistency and minimises the impact of questionnaire changes by having a fixed core questionnaire. The WVS has evolved over time, and data quality is higher in more recent waves than in earlier waves. In contrast, both the Eurobarometer and Latinobarometer put more emphasis on responsiveness to policy issues, and thus the questionnaires change more frequently and the response rates are generally lower. The OECD Trust Database is available on-line as an electronic annex of Gonzalez and Smith (2017).

Reliability

As discussed above, the reliability of a measure captures the degree to which it produces consistent results when measured at different times. Ideally, a good measurement instrument should produce the same estimate of the concept measured whenever it is applied in similar circumstances, and should produce a different result only when there has been a substantive change in the object of measurement. In survey research, the standard measure of reliability is test-retest reliability, where the same measurement item is administered to the same person after a delay of some period: this may be later on in the same survey, or it may involve the respondent being re-surveyed after a fixed period of time. Unfortunately, there are relatively few studies that investigate the test-retest reliability of trust measures. One of the few studies that does is Naef and Schupp (2009), who examine the reliability of measures of generalised trust: the authors find a relatively low correlation coefficient between the two time periods at which trust was measured (0.48) but adequate levels of reliability when considering three measures of generalised trust jointly. Taken on its own, this appears to point towards a need for multi-item measures of trust in order to obtain sufficient levels of reliability. However, an alternative strategy for assessing the reliability of trust measures is available. The OECD Trust Database provides repeated measures of trust for a wide number of countries at different points in time, from different surveys, and using different measurement instruments. This provides a strong basis for assessing the reliability of trust measures at the country-average level. In particular, if different surveys produce consistent results for different countries, despite differences in timing and methodology, then it can be said that trust measures display a high degree of reliability at the cross-country level. Evidence of reliability based on analysis of data in the OECD Trust Database is presented below separately for institutional and interpersonal trust.

Institutional Trust.   The wide range of questions on institutional trust available in the OECD Trust Database, and the fact that all surveys include at least some of them, makes it possible to conduct a fairly thorough analysis of the reliability of these measures. This includes looking at whether the reliability of questions on institutional trust is different for different types of institution. Table 2.4 reports the correlation coefficients between measures of trust in the national government for seven different surveys. In each case, the comparison is between country-average levels of trust in the national government for the same country and year, but as measured through different surveys. Because the surveys may be carried out at different periods within the same year, and because each survey has a unique sample of respondents, some difference in reported levels of trust may be expected. Still, a high correlation coefficient provides strong evidence that the measures produce reliable information.

Table 2.4. **Correlation in trust in the national government across surveys**

Survey	GWP	EB	ESS	WVS	EQLS	LB	EU-SILC
GWP							
EB	0.84						
ESS	0.81	0.74					
WVS	0.84	0.91	..				
EQLS	0.89	0.91	0.89	..			
LB	0.77	..	..	0.59	..		
EU-SILC	0.81	0.85	0.87	..	0.86	..	

Note: The Pearson correlation coefficient is a test for the strength of a relationship between two variables or datasets. Under this method, it is assumed that the data are normally distributed and that the expected relationship between them is linear. If the correlation between the two items is high, this suggests that the two measures capture the same underlying concept. GWP = Gallup World Poll; EB = Eurobarometer; ESS = European Social Survey; WVS = World Values Survey; EQLS = European Quality of Life Survey; EU-SILC= European Union Statistics on Income and Living Conditions; LB = Latinobarometer. The ESS question refers to trust in politicians rather than the government.
Source: Calculations based on the OECD Trust Database.

StatLink *http://dx.doi.org/10.1787/888933584089*

Table 2.4 shows correlation coefficients between surveys of between 0.59 (Latinobarometer and the WVS) and 0.91 (Eurobarometer and both the WVS and the EQLS). With the exception of the Latinobarometer/WVS result, all the correlation coefficients are above 0.74, and the majority are well above 0.80. It is worth noting that all correlations below 0.80 involve either the Latinobarometer or the Eurobarometer, which are the two surveys included in the analysis with the lowest levels of methodological rigour. In particular, both surveys involve relatively frequent changes both to the questionnaire and to the time when the survey is in the field, and response rates are relatively low. It is also of interest that the European Social Survey, which asks about trust in politicians rather than in national governments, produces a lower correlation than other surveys.

Table 2.5 reports a similar analysis that focuses on trust in the judicial system rather than the national government. Again, the results obtained from different surveys correlate with each other very strongly, ranging from 0.68 (Latinobarometer/GWP) to 0.93 (European Social Survey with the GWP and with Eurobarometer). Ignoring the Latinobarometer, the range of correlations is even narrower (0.82 to 0.96) than in the case of trust in the national government, reflecting the fact that all of the surveys reviewed here include relatively similar

Table 2.5. **Correlation in trust in the judicial system across surveys**

Survey	GWP	EB	ESS	WVS	EQLS	LB	EU-SILC
GWP							
EB	0.90						
ESS	0.93	0.93					
WVS	0.82	0.82	..				
EQLS	0.92	0.84	0.90	..			
LB	0.68	..	..	..	..		
EU-SILC	0.92	0.96	0.95	..	0.91	..	

Note: The Pearson correlation coefficient is a test for the strength of a relationship between two variables or datasets. Under this method, it is assumed that the data are normally distributed and that the expected relationship between them is linear. If the correlation between the two items is high, this suggests that the two measures capture the same underlying concept. GWP = Gallup World Poll; EB = Eurobarometer; ESS = European Social Survey; WVS = World Values Survey; EQLS = European Quality of Life Survey; LB = Latinobarometer, EU-SILC= European Union Statistics on Income and Living Conditions.
Source: Calculations based on the OECD Trust Database.

StatLink *http://dx.doi.org/10.1787/888933584108*

questions on trust in the judicial system. Similarly, when the same analysis is performed on trust in the police, the correlation coefficients range from 0.75 to 0.90.

Taken together, Tables 2.4 and 2.5 support the view that measures of institutional trust are reliable at the country-average level. For all instances where the question is directly comparable and the surveys are of relatively high quality, cross-country correlations are in excess of 0.80 and typically above 0.90 for those surveys with the highest data quality (the GWP, the ESS and the EQLS). Even where the data quality is lower, the correlation between country averages is still respectable.

Interpersonal Trust.   Evidence from the OECD Trust Database on the reliability of interpersonal trust measures is sparser than in the case for institutional trust. This is because the Gallup World Poll does not include measures of interpersonal trust in the core questionnaire, while the WVS country coverage only partially overlaps with that of the ESS and the EQLS for the post-2002 period covered by the OECD Trust database. Nonetheless, for those comparisons that are possible, the picture of reliability that emerges is fairly good. Figure 2.1 shows two comparisons for generalised trust that are possible using the OECD Trust Database. The ESS and the EQLS produce very consistent results, with a correlation coefficient of 0.82. In addition, the absolute levels of generalised trust for each country are very similar across the two surveys (Panel A). The correlation between levels of generalised trust in the WVS and the Eurobarometer is a little lower at 0.75, but still relatively high (Panel B).

Although lack of data in many of the surveys included in the OECD Trust Database limits the number of comparisons that can be made between surveys, a second strategy for investigating reliability is possible using the EQLS and ESS. This involves looking at correlations between waves of the same survey. If the measure used is reliable, then correlations between successive waves should be relatively high, and the correlation should decline between waves that are more separated in time. The WVS is not suitable for this approach due both to the long period between waves and to the changes in country coverage between waves. However, both the EQLS and ESS have smaller gaps between waves (two and three years respectively), and cover a consistent panel of countries.

Table 2.6 below shows the correlation between waves of the ESS. It can be seen that the correlations are very high (between 0.96 and 0.98 for successive waves). The correlation falls, as would be expected, with waves that are more separated in time. Nonetheless, these correlations also remain high, and the combination of high inter-wave correlations falling with time is consistent with a high level of reliability.

Generalised trust measures are available for fewer waves of the EQLS than the ESS. Nonetheless, it is possible to look at the correlation between waves 2 and 3 of the EQLS. Figure 2.2 illustrates this and shows a high correlation of 0.93. This is a little below that found for the ESS, even allowing for the slightly longer gap between waves in the EQLS when compared to the ESS. However, this level of correlation is still high in absolute terms and is consistent with a high degree of reliability in the measure.

Face validity

Face validity addresses the degree to which a measure is intuitively plausible. This is important as it can impact on the quality and uptake of data. Poor face validity affects the quality of data because respondents may be unwilling to provide a well-considered answer

Figure 2.1. **Correlation between measures of generalised trust across surveys**

European Social Survey and European Quality of Life Survey

r=0.82

World Values Survey and Latinobarometer

r=0.75

Note: Both the European Social Survey (ESS) and the European Quality of Life Survey (EQLS) responses have been recoded so that a response from 5-10 on the 0-10 response scale (ESS) and a response from 6-10 on the 1-10 response scale (EQLS) counts as indicating trust in other people.
Source: Calculations based on the OECD Trust Database.

StatLink 📊 *http://dx.doi.org/10.1787/888933583576*

Table 2.6. **Intra-wave correlation for generalised trust, based on the European Social Survey**

Wave	1	2	3	4	5	6
1						
2	0.98					
3	0.98	0.97				
4	0.97	0.97	0.98			
5	0.94	0.94	0.98	0.96		
6	0.94	0.95	0.96	0.95	0.98	

Note: The correlation is based in the standard generalised trust question: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?"
Source: OECD calculations based on the European Social Survey (database) *www.europeansocialsurvey.org/data/* as contained in the OECD Trust Database.

StatLink 📊 *http://dx.doi.org/10.1787/888933584127*

Figure 2.2. **Intra-wave correlation for generalised trust based on the European Quality of Life Survey**

Note: The correlation is based on the standard generalised trust question: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" Responses have been recoded so that a response from 6-10 on the 1-10 response scale counts as indicating trust in other people.

Source: OECD calculations based on Eurofund (2016): European Quality of Life Survey (database) *https://discover.ukdataservice.ac.uk/catalogue/?sn=7316&type=Data%20catalogue* as contained in the OECD Trust Database.

StatLink ᘒᒲᔿ *http://dx.doi.org/10.1787/888933583595*

to a question that seems frivolous or lacking in meaning. Similarly, data users may be unwilling to place faith in evidence that looks implausible or difficult to interpret. Hence, although face validity is a relatively "loose" and imprecise concept that is difficult to quantify, it is nonetheless of fundamental importance in establishing whether data are fit for purpose. There exists no obvious metric for assessing the face validity of a survey question, so discussion of face validity must either use qualitative or anecdotal evidence or make inferences from the behaviour of respondents. Both approaches to assessing face validity can be applied to measures of trust.

On the qualitative front, the key issue is whether the measures of trust in use seem intuitively plausible. This is a subjective judgement, and not one for which a definitive answer is possible. A number of objections to the wording of the standard trust questions have been raised over the years, notably by Hardin (2004) and Delhey et al. (2011). In both cases, it is argued that the standard Rosenberg question on generalised trust is too general to allow for a meaningful answer. Hardin also raises similar issues with respect to questions on trust in institutions, arguing that, in order for respondents to provide a meaningful answer, such questions should be specific not only about the institution considered but also in terms of what activity is expected of it.

While these issues are important, the standard questions have been considered as meaningful by the large number of authors who work with them or who have opted to use them in surveys. Knack (2001) notes that countries with high measured levels of interpersonal trust using these questions are "the Nordic nations, where citizens commonly leave unlocked bicycles and unattended strollers in public areas". While these considerations, in and of themselves, do not allow conclusive judgement on the face validity of trust measures, a more quantitative approach is to look at how difficult respondents find it to answer questions on trust. A simple way of evaluating this is to look at the item-specific
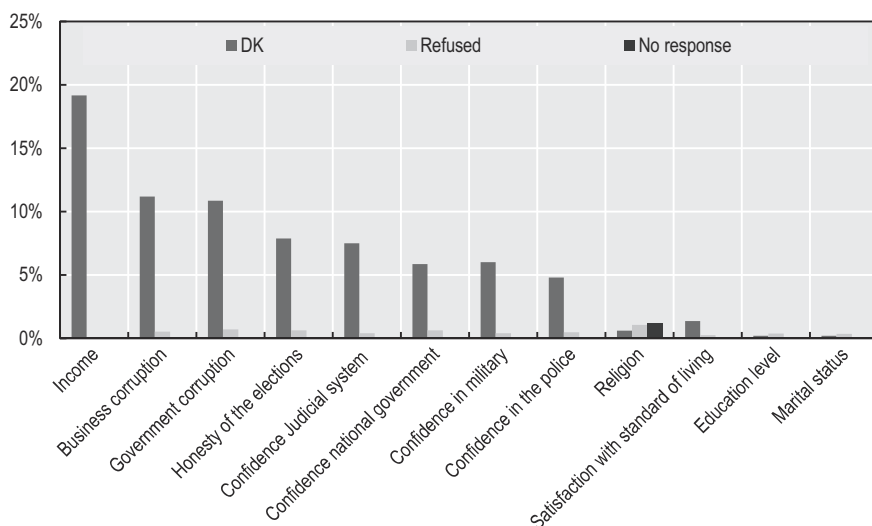
non-response rates for measures of trust compared to other widely used survey items. If respondents find the trust questions confusing or difficult to answer, then a higher item-specific non-response rate should be anticipated for these items.

Figure 2.3 shows the item-specific non-response rates for a range of different survey questions included in the European Social Survey. This includes questions on institutional trust (e.g. trust in the European Parliament, the United Nations, national parliament, the legal system and the police), questions related to generalised trust (e.g. most people take advantage, most of the time people are helpful, most people can be trusted) and a number of questions on topics unrelated to trust that are commonly asked in household surveys (e.g. income, marital status, gender). Income is, by a large margin, the most difficult topic for respondents to answer, as it requires combining information from very different income sources paid to different members of the same household; over 20% of respondents either refuse to answer or reply *don't know*. By way of contrast, only 8% of respondents did not answer questions on marital status, while gender has a response rate of close to 100%.

Figure 2.3. **Item specific non-response in the European Social Survey, 2002-14**

Source: OECD calculations based on the European Social Survey (database) *www.europeansocialsurvey.org/data/* as contained in the OECD Trust Database

StatLink ⧉ *http://dx.doi.org/10.1787/888933583614*

In comparison to gender and marital status, questions on institutional trust did not fare particularly well. More than 10% of respondents failed to provide an answer to questions about the European Parliament or United Nations, and even the more familiar national institutions had non-response rates of between 1% and 3%. These rates are much lower than

in the case of income, but well above those for gender. Questions on generalised trust performed better, with non-response rates of 1% or less for the three questions included here.

Figure 2.4 provides further evidence on the face validity of institutional trust by extending the analysis of non-response rates to the Gallup World Poll. This survey has a significantly larger sample than the European Social Survey, and includes a much broader cross-section of countries. The picture that emerges from the GWP analysis is very similar to that from the ESS. Income is the item with the highest item-specific non-response rate, at 19%, while marital status and education both have very low non-response rates (at around 0.50%), and religion has a non-response rate of 3%. All of the institutional trust measures in the GWP have high item-specific non-response rates, ranging from 5% for confidence in the police to 8% for confidence in the judiciary; questions on corruption and the honesty of elections have even higher item-specific non-response rates.

Figure 2.4. **Item-specific non-response rates in the Gallup World Poll, 2006-16**

Note: The response option "No response" for the question on religion was only given to respondents before 2011. Income data refers to Gallup World Poll waves 1 to 9. Missing responses were coded as "Don't know" in the case of income.
Source: OECD calculations based on the Gallup World Poll (database) *www.gallup.com/services/170945/world-poll.aspx* as contained in the OECD Trust Database.

StatLink ⟦⟧ *http://dx.doi.org/10.1787/888933583633*

The overall picture on face validity that emerges from an analysis of non-response rates is, hence, mixed. While all trust questions perform better than income – which is commonly collected in almost all household surveys – the item-specific non-response rates are higher for the bulk of the trust questions than for questions on marital status, education or gender. Even religion, which might be considered a relatively sensitive question, had a non-response rate that was less than half of most institutional trust questions (although still a substantially higher non-response rate than is the case for generalised trust in the ESS). An analysis of non-response rates does not, in itself, provide conclusive evidence on face validity. High non-response rates for measures of institutional trust could simply reflect that, although the question is reasonable, many people lack adequate information to provide a meaningful answer. The very high non-response rates for questions on less familiar organisations such as the United Nations and the European Parliament, when compared to national institutions, provide some support for this view.

Convergent validity

A measure is said to have convergent validity if it correlates well with other measures of the same construct. For example, a self-report measure of individual health status might be considered to have good convergent validity if it is correlated with the use of health services (higher usage being a proxy of poor health) and with a third-party assessment of the subject's health status. Convergent validity is easier to assess for areas where a range of plausible proxy measures of the same underlying construct exists, but for which no ideal measure exists (in which case, contingent validity would not be relevant). In the case of trust, no unambiguous direct measure is obviously available, but a number of proxies exist. This is the case for measures of interpersonal trust and particularly generalised trust, where a range of both experimental and non-experimental measures of trust can be used to test convergent validity. For institutional trust, however, the available evidence on convergent validity is sparser.

Interpersonal Trust. There are two primary sources of data on interpersonal trust that can be used to assess convergent validity. The first of these involves looking at the correlation between different survey questions on trust. Section 2.3 identified several approaches to asking questions on trust (evaluation, expectation and experience), and the correlation in responses to these different types of question provides one potentially valuable source of information. It is also possible to look at how responses from different sources vary with respect to trust. The second main source of information on convergent validity, and potentially the strongest source of evidence, lies in the results of experimental studies. There is now a large body of experimental data that can be used to validate survey questions on trust. Finally, although more limited in scope, there is a small body of evidence on the bio-physical correlates of trust that is also relevant to assess convergent validity. Almost all of this evidence focuses on generalised trust, both because data on generalised trust are more widely available than is the case for limited trust and because generalised trust is of greater relevance than limited trust for most policy and analytical purposes.

Knack (2001) provides a good overview of the validity of generalised trust measures from the perspective of convergent validity. He notes, in particular, that data from the *Reader's Digest* lost wallet experiments (Box 2.2) correlate well with survey measures of generalised trust. Despite the relatively low sample size for the lost wallet experiment, the proportion of wallets returned correlates with country values of the WVS measure of generalised trust at 0.65 (p < 0.01). When per capita income is controlled for, the correlation is even stronger. This finding is reinforced by Helliwell and Wang (2010), who note that the proportion of lost wallets returned in the *Reader's Digest* experiment correlates well with data from an expectations question on whether a lost wallet would be returned, which in turn correlates well with the WVS generalised trust question. This is also true for other generalised trust questions. For example, Falk et al. (2015) note that their preferred measure of trust ("I assume that people have only the best intentions") is correlated with the WVS measure of interpersonal trust at 0.53 (p < 0.01).

Looking at a wider range of questions, Knack (2001) notes that responses to the WVS question on generalised trust are strongly correlated with items from the same survey relating to respondents' attitudes towards taking advantage of others (e.g. cheating on taxes or not reporting damage to a parked vehicle). These correlations focus on people's assessment of their own trustworthiness rather than on whether other people can be trusted, so the measures are sufficiently different to add substantial information. Knack and Keefer (1997) note that the relationship between generalised trust and attitudes towards

taking advantage of others, like that between trust and the proportion of lost wallets returned, is stronger after controlling for per capita income. Naef and Schupp (2009) look at the relationship between people's past trusting behaviour (e.g. lending personal possessions, lending money, leaving the door unlocked) and measures of generalised trust. Using data from the German Socio-economic Panel, they find a robust relationship between generalised trust and past instances of trusting behaviour.

Another source of evidence of convergent validity is the correlation between country-average levels of generalised trust and evaluations by foreigners of how trustworthy people from different countries are. Knack reports that, using Eurobarometer data on *how much you would trust people from different countries*, there is a 0.45 correlation (p = 0.056) with generalised trust measured in the WVS.

Looking beyond survey data, there is now a large body of experimental evidence on the validity of measures of generalised trust. This rests largely on lab experiments using one or more variants of the Trust Game (Berg and McCabe, 1995), as described in Box 2.2. Glaeser et al. (2000) provided the first systematic use of laboratory experiments to validate survey measures of trust. Using a sample of 189 Harvard students, the authors found no significant relationship between the standard WVS measure of generalised trust[11] and trust as measured in the Trust Game (the proportion of the starting endowment sent by player 1 to player 2). While this might seem strong evidence against the validity of trust measures, Glaeser et al. also find that the survey-based measures of generalised trust are a strong and significant predictor of *trustworthy* behaviour in the Trust Game (the proportion of player 2's net endowment returned to player 1, see Box 2.2). In other words, the experimental evidence suggests that the standard generalised trust questions do collect valid data about a person's trustworthiness, but not about an individual person's willingness to trust *ex ante*. Lazzarini et al. (2004) repeat the same experiment in Brazil, while also investigating the impact of a face-to-face set-up, as opposed to an anonymous set-up for the Trust Game. Their results confirm Glaeser et al.'s findings that survey-based measures of generalised trust correlate with individual trustworthiness in the Trust Game, but not with the individual's level of trust. Cox (2004) and Capra et al. (2007) explore the relationship between survey-based measures of trust and experimental results in more detail. By including measures of other-regarding preferences – both experimental and survey-based – these authors show that trusting behaviour in the Trust Game is predicted well by the standard WVS question once altruism is controlled for. Finally, Gachter et al. (2004) report that the standard generalised trust question is associated with co-operation in the public goods game.

All these studies share one limitation: they use a small and largely unrepresentative sample of participants in the experimental games. This raises the issue of whether the results can be extended to the population as a whole. While there are currently no cross-country experimental studies involving large-scale nationally representative samples, an alternative approach is to look at the results of a systematic meta-analysis of experimental studies. Johnson and Mislin (2011, 2012) undertake a thorough meta-analysis of experimental studies, involving the Trust Game, covering 162 replications of the Berg and McCabe trust experiment, 35 countries and over 23 000 respondents. Although most of these studies are small (the average sample size is 148), they cover a wide range of countries, both developing countries (e.g. Cameroon and Uganda) and developed countries (e.g. the United States and Sweden). Contrary to earlier experimental studies, Johnson and Mislin find a significant positive correlation between the WVS measure of generalised trust and trusting behaviour in experimental games, but no relationship with trustworthiness. One

explanation for this apparent contradiction is that the authors consider the relationship between country-average levels of trust in both survey responses and experimental results, while the studies cited earlier look at individual-level correlations.

A final source of information to assess the convergent validity of generalised trust measures is provided by Fehr (2009), who discusses a series of experiments analysing the effect of oxytocin (a neuro-transmitter highly associated with pro-social behaviour in mammals). In an experimental set-up, players of the Trust Game who received a nasal spray containing oxytocin immediately before the game showed significantly higher levels of trusting behaviour than those who received a placebo spray. Fehr argues convincingly that, in this experimental design, one can effectively rule out the possibility that oxytocin affected trust via affecting players' general altruism or their risk preferences. This study hence suggests instead that the measures of trust produced by the Trust Game capture genuine trusting behaviour, and are strongly grounded in a neurophysical mechanism.

Institutional Trust. While there are more survey data available on trust in institutions than is the case for interpersonal trust – thanks to the inclusion of institutional trust measures in the Gallup World Poll, Eurobarometer and Latinobarometer – there is much less information on the convergent validity of institutional trust measures. This is particularly the case with respect to experimental analyses, as the standard Trust Game does not lend itself to the analysis of trust in institutions. Another complicating factor is that, while for interpersonal trust there is a clear consensus on using generalised trust as the preferred measure of interpersonal trust, institutional trust covers several dimensions, all of which are of interest (cf. Box 2.1).

Bouckaert and Van de Walle (2003) review the public administration literature with respect to trust and argue that measures of institutional trust do not adequately capture good governance, although their analysis is more discursive than empirical. Davidov and Coromina (2013) take a more empirical approach and use data from four waves of the European Social Survey to test measurement invariance across European countries, in both interpersonal and political trust. While they find limited support for scale invariance for measures of interpersonal trust (similar to the findings of Reeskens and Hooghe, 2008, who used metric equivalence tests on the trust item in the European Social Survey 2002 and 2004 waves), the evidence provides less support with respect to measures of political trust. In particular, their model rejects strong convergence between measures of trust in politicians and trust in parliament or the legal system. Cook and Gronke (2005) find similar ambiguity with different measures of trust in institutions, reporting correlations ranging from 0.22 to 0.36 between different questions. While Cook and Gronke view these correlations as acceptable, given that they capture slightly different concepts, they are nonetheless low when compared to the correlations between different measures of interpersonal trust. There is preliminary evidence that compares survey and experimental measures of institutional trust, relying on psychometric implicit association testing (see Box 2.4 for a description of the OECD's *Trustlab* project, or Intawan and Nicholson, 2017). The initial results, for trust in the national government, point towards a low but significant correlation between both types of measures, with implicit trust consistently scoring higher than self-reported trust.

New approaches to validating measures institutional trust. Although the existing evidence on the convergent validity of measures of institutional trust is limited, this does not mean that it is impossible to do better. In fact, there are a number of empirical strategies open

> ### Box 2.4. **Trustlab**
>
> A key limitation of the existing experimental literature on Trust Games is that almost all the studies draw on small, non-representative samples. Often, respondents are drawn from convenience samples such as undergraduate or graduate students in a particular course, raising concerns about the extent to which any study results would hold in a more general setting. Despite this, experimental techniques are attractive, both because they seem to predict behaviours in the real world rather well (Karlan, 2005; Benz and Meier, 2006; Algan and Cahuc, 2013) and because a different set of methodological biases is associated with them compared to those associated with the standard survey questions. This means that the two approaches are well suited to be used for validating each other. While meta-analyses such as that of Johnson and Mislin (2011) go some way to addressing these issues, and some authors (Fehr et al., 2003; Falk et al., 2016) have conducted representative surveys based on validation from small experimental samples of students, the ideal approach would involve a cross-country study involving large nationally representative samples using the same questionnaire and experimental design in each country.
>
> *Trustlab* is a joint initiative between the OECD, the Paris Institute of Political Studies (Sciences Po) and a range of additional academic and governmental partners to carry out a series of comparable experimental studies of trust across OECD countries. Each study is based on a nationally representative sample of 1 000 individuals stratified by age, gender and income. In the current set-up, respondents log onto a custom-made online platform to participate in three experimental games to provide behavioural measures of social norms and values. These are: the Dictator Game, which provides information on altruism; the Trust Game, which provides information on trust and trustworthiness; and the Public Good Game, which provides information on willingness to co-operate and contribute to public goods. In all three cases, respondents are matched with other respondents for the games, which are played with approximately USD 10 in actual money at stake. In some countries, Conditional Trust Games that assess trust in specific population groups are included. After the games, respondents complete a series of implicit association tests on attitudes towards a range of institutions (namely, the government, the judicial system and the media). Implicit association tests are a psychometric technique used to test respondent attitudes where issues of social desirability may make them unwilling to respond honestly, or in areas that are difficult to measure through explicit self-reporting due to lack of awareness (Greenwald et al., 2002). These tests have been applied successfully to measure perceptions, stereotypes and attitudes towards commonly stigmatised social groups such as Black people, women and the elderly (Dasgupta and Asgari, 2004; Aberson et al., 2004). The final part of *Trustlab* involves respondents completing a survey questionnaire containing a battery of trust questions. These cover a range of different approaches to measuring both interpersonal and institutional trust, self-reported items on other social norms such as altruism and reciprocity, as well as a range of questions on the potential policy drivers of trust, along with basic demographic and socio-economic information. *Trustlab* has so far been run on nationally representative samples in France, Korea, the US, Germany, Italy and Slovenia, with additional countries lined up to join. The first evidence is expected to become available at the beginning of 2018.

for investigating the convergent validity of these measures that might yield useful information. Most of these involve analysing existing data on institutional trust in the context of actual behaviour and evaluating whether the level of trust revealed by people's behaviour is consistent with survey measures of institutional trust. The fact that these have not to date

been the subject of any in-depth analysis represents a significant misallocation of scholarly resources. In particular, although there is a large political science literature analysing aspects of institutional trust, the systematic testing of the validity of institutional trust measures against actual behaviour is a major research gap. Although it is beyond the scope of these Guidelines to provide a detailed breakdown of all possible research strategies for testing the validity of measures of institutional trust, some proposals for research are set out below in the interests of ensuring that scholarly effort is directed towards issues of high importance.

Trust in the police, and in the legal system more broadly, are two areas where it should be possible to obtain information on convergent validity. Measures of trust both in the police and in the legal system are available from a number of surveys with good cross-country coverage (e.g. the ESS, EQLS, WVS): this provides the survey data for analysis. The behavioural data against which to test the validity of survey measures could come from victimisation surveys and police data. In particular, high trust in the police and the legal system should be reflected in a higher proportion of incidents of criminal victimisation being reported to the police. Victimisation surveys provide information on the base rate of criminal victimisation, while police statistics record the rate at which incidents of victimisation are reported to the police. It should therefore be possible to examine whether the ratio of incidents of criminal victimisation reported to the police as a proportion of total incidents of victimisation is closely associated with the values of trust in police across countries.

Other similar analyses are also possible to provide information on other aspects of institutional trust. For example, the per capita consumption of bottled water could be used as a proxy measure of trust in the institutions associated with the water supply and tested across different jurisdictions within a country. Clearly in such a case, controlling for the income of different areas would be essential, but this is not a major challenge. Alternatively, variance across a country in the proportion of the population with private medical insurance might be used as a proxy indicator of trust in public medical services. Other similar proxies could be identified for areas such as education or disaster preparedness. While analyses of this sort do not provide much information on overall trust in public institutions (possibly reflecting inherent difficulties in the concept), they would provide a much stronger grounding for measures of trust in specific institutions. For this reason, research into the convergent validity of institutional trust using behavioural data should be a major research priority with respect to institutional trust.
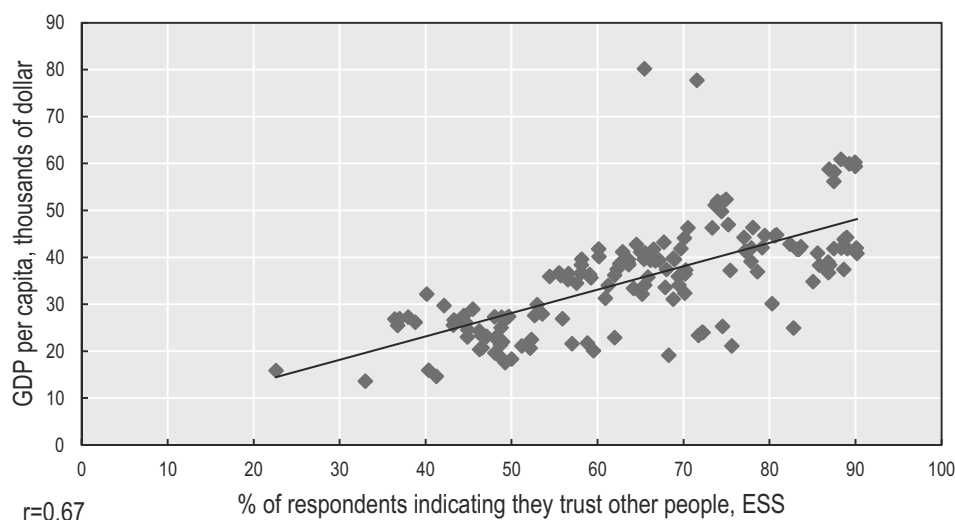
Construct validity

In addition to evidence that measures of trust drawn from different sources provide the same picture of the variation in trust across countries, it is also useful to look at whether trust measures behave in a way that is consistent with expectations. For example, if trust in strangers (i.e. the standard WVS question on generalised trust) is an important part of social capital (as argued by Scrivens and Smith, 2013), then one would expect to see higher levels of it associated with better economic and social outcomes. Similarly, higher levels of institutional trust should be associated with greater willingness on the part of people to co-operate with government agencies in the pursuit of common goals, or with higher measures of government trustworthiness.

Generally speaking, measures of both interpersonal and institutional trust are well supported in terms of construct validity. While only a few papers have analysed trust from the perspective of convergent validity (particularly for measures of institutional trust), a large literature has looked at the determinants and correlates of trust – both interpersonal and

institutional. This reflects the high degree of interest in measures of trust within economics, political science, sociology and public management. Indeed, the fact that trust measures perform so well in terms of construct validity helps to explain why there has not been a stronger emphasis on more formally testing the other types of validity of such measures.

Interpersonal Trust. For interpersonal trust, within the limits of existing unofficial data, convergent validity is well supported, in particular for generalised trust. Countries with high levels of generalised trust tend to have higher incomes (Fukuyama, 1995; Whitely, 1997; Knack and Keefer, 1997; La Porta et al., 1997; Algan and Cahuc, 2013), which is consistent with the view that generalised trust is a core element of social capital and is important for lowering transaction costs between people who do not know each other well. Figure 2.5 below illustrates the relationship between income (GDP per capita) and generalised trust in the ESS between 2002 and 2014, showing a clear correlation between the two measures (r of 0.67) for European countries.
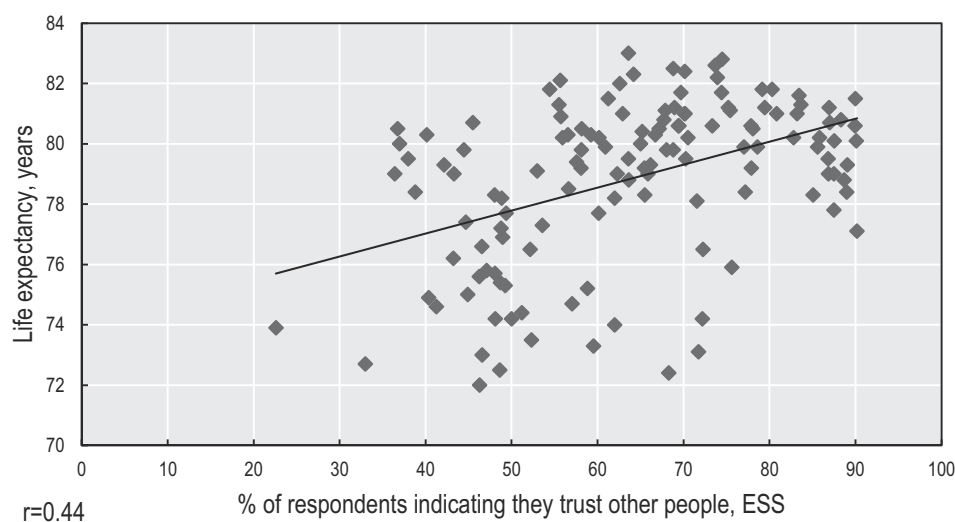
Figure 2.5. **Generalised trust and GDP per capita in European countries, 2002-14**

Note: The correlation is based on the standard generalised trust question: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" Responses have been recoded so that a response from 5-10 on the 0-10 response scale counts as indicating trust in other people.
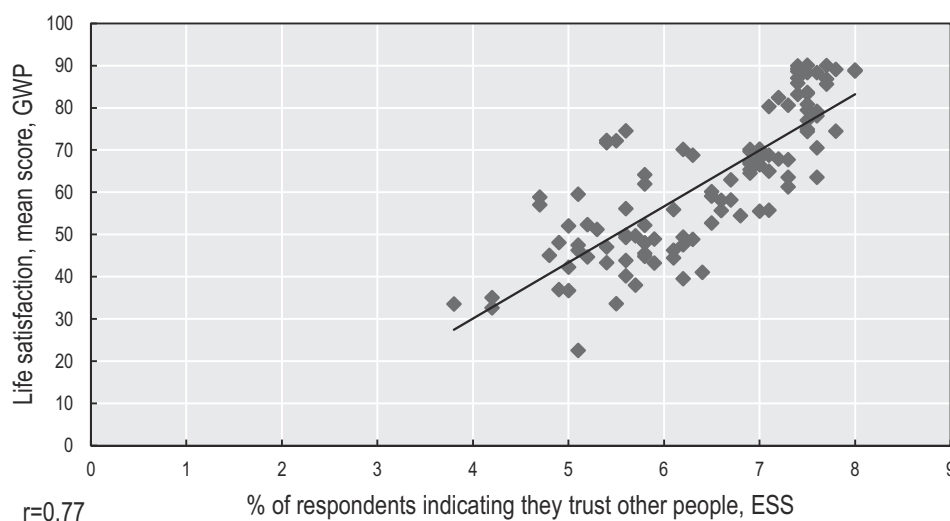*Source:* OECD calculations based on the European Social Survey (database) *www.europeansocialsurvey.org/data/* and OECD Stat (database) *http://stats.oecd.org/index.aspx?DataSetCode=PDB_LV* as contained in the OECD Trust Database.
StatLink ᵃᵐˢᵖ *http://dx.doi.org/10.1787/888933583652*

In addition to levels of income, trust is also correlated with the short-term performance of the economy. Figure 2.6 shows the correlation between generalised trust in the ESS and unemployment rates in European countries between 2002 and 2014. Although the relationship is weaker than for income, generalised trust is significantly negatively correlated with unemployment rates across European countries.

There is also evidence that generalised trust correlates positively with better health outcomes for individuals (Kawachi and Berkman, 2000; Boreham et al., 2004; Ginn and Arber, 2004). Figure 2.7 shows the relationship between generalised trust and life expectancy using the same dataset as Figures 2.5 and 2.6. Although the relationship is weaker than is the case for income, life expectancy is still strongly correlated with generalised trust (r of 0.44).

Figure 2.6. **Generalised trust and unemployment in European countries, 2002-14**

r=-0.38

Note: The correlation is based on the standard generalised trust question: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" Responses have been recoded so that a response from 5-10 on the 0-10 response scale counts as indicating trust in other people.
Source: OECD calculations based on the European Social Survey (database) *www.europeansocialsurvey.org/data/* and OECD Stat (database) *http://stats.oecd.org/index.aspx?queryid=36324* as contained in the OECD Trust Database.

StatLink ᵃᵐˢˡ *http://dx.doi.org/10.1787/888933583671*

Figure 2.7. **Generalised trust and life expectancy in European countries, 2002-14**
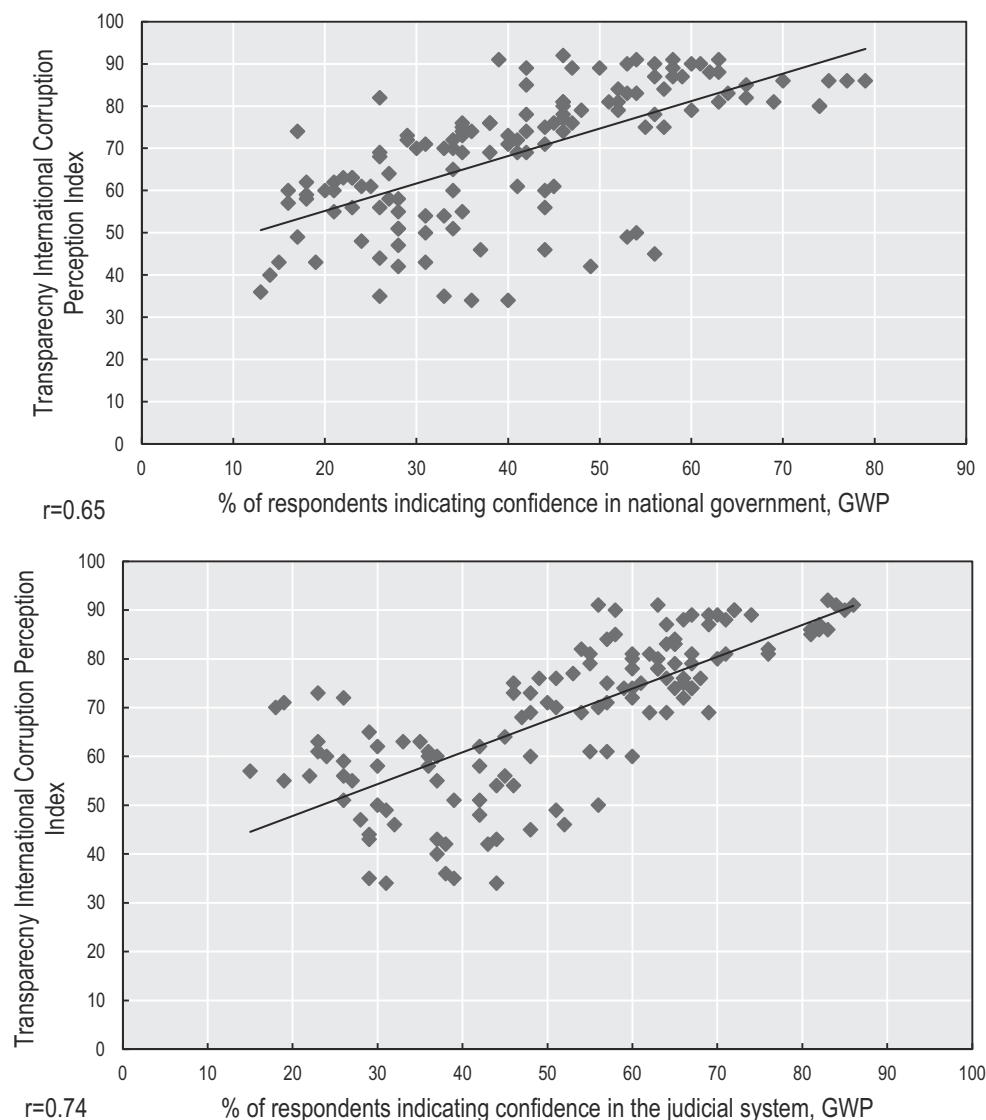

r=0.44

Note: The correlation is based on the standard generalised trust question: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" Responses have been recoded so that a response from 5-10 on the 0-10 response scale counts as indicating trust in other people.
Source: OECD calculations based on the European Social Survey (database) *www.europeansocialsurvey.org/data/* and OECD Stat (database) *https://stats.oecd.org/index.aspx?queryid=24879* as contained in the OECD Trust Database.

StatLink ᵃᵐˢˡ *http://dx.doi.org/10.1787/888933583690*

There is also evidence of a robust relationship between generalised trust and other important social outcomes. As with incomes, this is consistent with the social capital literature, suggesting that generalised trust is important in allowing strangers to co-ordinate actions to achieve a wide variety of different goals. Sampson (2012) finds a negative

correlation between generalised trust and levels of crime, while there is a large literature on the relationship between generalised trust and measures of subjective well-being (e.g. Algan and Cahuc, 2013; Helliwell and Wang, 2010). Figure 2.8 illustrates the correlation between generalised trust and life satisfaction at the country level, showing that variation in generalised trust co-varies strongly with variation in life satisfaction across European countries (r of 0.77). Boarini et al. (2012) take this analysis further and show that average levels of interpersonal trust at the country level are strongly correlated with the life satisfaction of the inhabitants of these countries, independently of the individual's own trust, and after controlling for demographic and economic variables.

Figure 2.8. **Generalised trust and life satisfaction in European countries, 2002-14**

r=0.77

Note: The correlation is based on the standard generalised trust question: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" Responses have been recoded so that a response from 5-10 on the 0-10 response scale counts as indicating trust in other people. Life satisfaction scores refer to a country's average score on a response scale from 0-10 for the question: "Overall, how satisfied are you with life as a whole these days?" where zero means "not at all satisfied" and ten means "completely satisfied".
Source: OECD calculations based on the European Social Survey (database) *www.europeansocialsurvey.org/data/* and the Gallup World Poll (database) *www.gallup.com/services/170945/world-poll.aspx* as contained in the OECD Trust Database.

StatLink ᵃ§ᵇ *http://dx.doi.org/10.1787/888933583709*

Institutional Trust.   There is good evidence of a positive relationship between institutional trust and citizen support for government policy. In a cross-country analysis, Zhao and Kim (2011) highlight a positive correlation between institutional trust and levels of the Foreign Direct Investment that a country receives. Murphy (2004) and Murphy et al. (2009) find a significant positive relationship between trust in regulators and voluntary compliance in the area regulated; similarly, Daude et al. (2012) have documented a strong link between institutional trust and willingness to pay taxes. Furthermore, Knack and Keefer (1997) analysed responses to the WVS across about 30 countries, finding a positive correlation between measures of citizens' confidence in government and objective indicators of bureaucratic efficiency. Moreover, the German Satisfaction with Government Services Survey (Zufriendenheitsbefragung, 2015) points to a positive correlation between satisfaction with services, on one side, and measures of competences of a specific agency and trust in that office, on the other. There is also a robust cross-country correlation between trust in institutions and perceptions of corruption (OECD, 2013b). Figure 2.9 illustrates this relationship

Figure 2.9. **Institutional trust and perceptions of corruption in OECD countries, 2006-15**
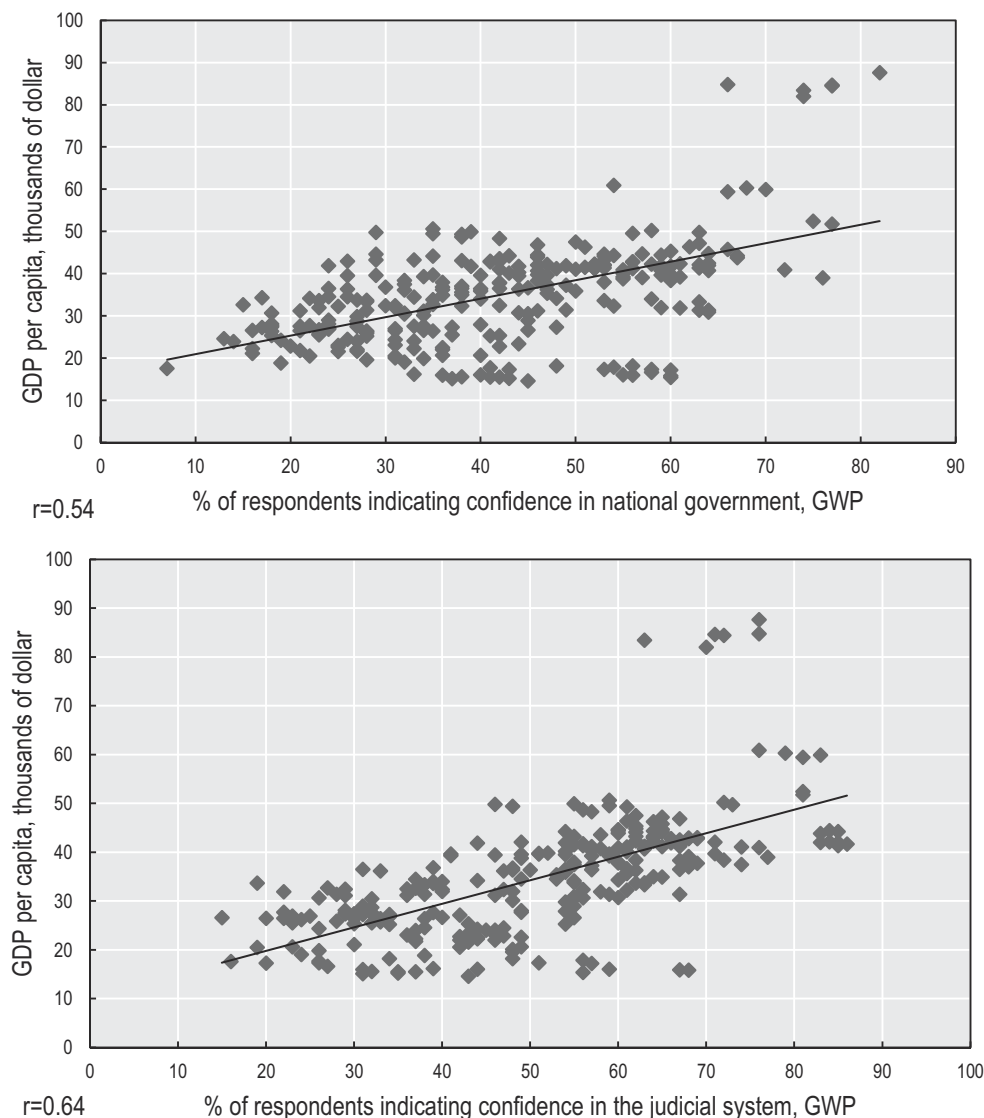

r=0.65

% of respondents indicating confidence in national government, GWP

r=0.74

% of respondents indicating confidence in the judicial system, GWP

Note: The Gallup World Poll asks the following two questions: "In this country, do you have confidence in each of the following, or not? How about national go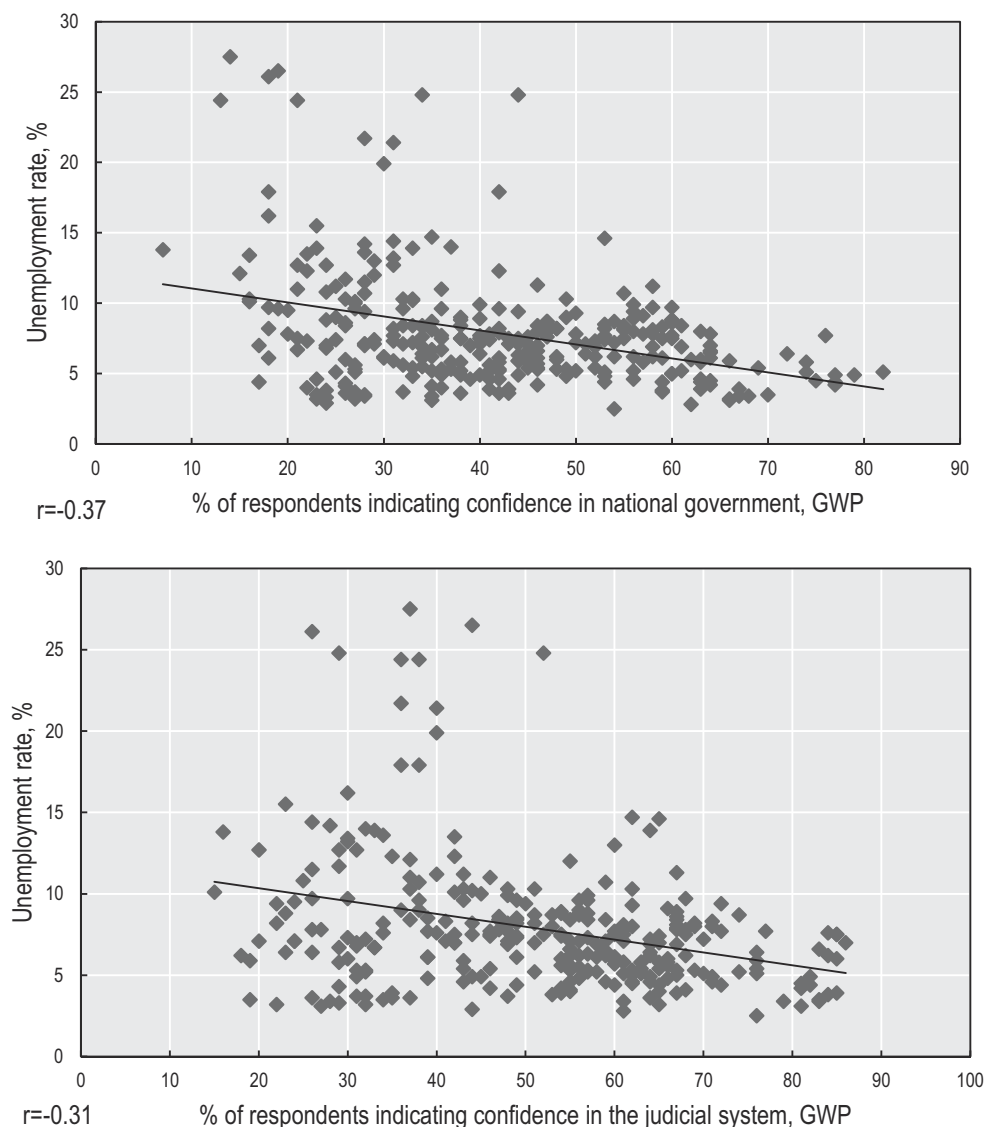vernment? How about judicial systems and courts?" Transparency International's Corruption Perception Index scores countries from 0-100. The higher the score, the more transparent/less corrupt a country is.

Source: OECD calculations based on the Gallup World Poll (database) *www.gallup.com/services/170945/world-poll.aspx* and Transparency International (2016): Corruption Perception Index (database) *www.transparency.org/news/feature/corruption_perceptions_index_2016* as contained in the OECD Trust Database.

StatLink ⎯⎯ *http://dx.doi.org/10.1787/888933583728*

for OECD countries, showing that perceived levels of corruption are correlated with both trust in government (r = 0.65) and trust in the judicial system (r = 0.74).

As with interpersonal trust, there is also evidence that institutional trust is linked to economic outcomes. Figure 2.10 shows the relationship between trust in two institutions – government and the judiciary – and GDP per capita. In both cases the correlation is strongly positive, slightly more so in the case of trust in the judiciary. This makes intuitive sense, as many of the key levers through which institutions affect economic outcomes, such as

Figure 2.10. **Institutional trust and GDP per capita in OECD countries, 2006-15**

r=0.54

r=0.64

Note: The Gallup World Poll asks the following two questions: "In this country, do you have confidence in each of the following, or not? How about national government? How about judicial systems and courts?"
Source: OECD calculations based on the Gallup World Poll (database) *www.gallup.com/services/170945/world-poll.aspx* and OECD Stat (database) *http://stats.oecd.org/index.aspx?DataSetCode=PDB_LV* as contained in the OECD Trust Database

StatLink 🔗 *http://dx.doi.org/10.1787/888933583747*

contract enforcement or regulation of the market place, have a more direct link to the judicial system than to the government more generally.

Roth (2015) finds a strong negative correlation between interpersonal trust and unemployment for 12 countries in the euro area from 1999 to 2014. Figure 2.11 confirms this relationship, using a different set of data sources, and also provides information on the relative strength of the relationship for different forms of institutional trust. While there is a relationship between trust in the judicial system and the unemployment rate (r =- 0.31), the relationship is stronger for trust in the government (r = -0.37). This is consistent with a view that responses to questions on trust in institutions reflect the respondent's trust both

Figure 2.11. **Institutional trust and unemployment rates
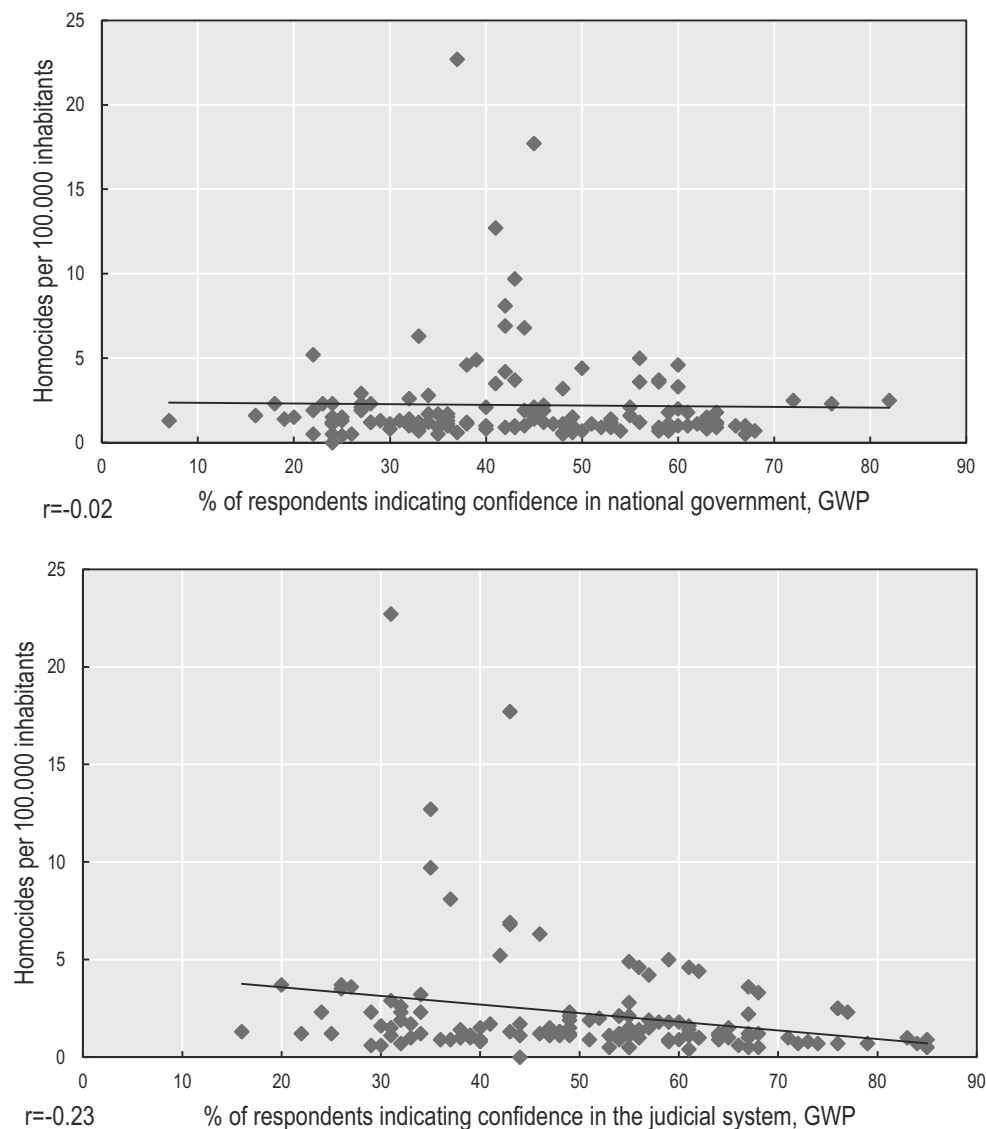in OECD countries, 2006-15**

r=-0.37   % of respondents indicating confidence in national government, GWP

r=-0.31   % of respondents indicating confidence in the judicial system, GWP

Note: The Gallup World Poll asks the following two questions: "In this country, do you have confidence in each of the following, or not? How about national government? How about judicial systems and courts?"
Source: OECD calculations based on the Gallup World Poll (database) *www.gallup.com/services/170945/world-poll.aspx* and OECD Stat (database) *https://stats.oecd.org/index.aspx?queryid=36324* as contained in the OECD Trust Database.

StatLink ⟪ᘯᔑ⟫ *http://dx.doi.org/10.1787/888933583766*

in institutions generally as well as in the specific institution that is the focus of the question. In this case, the correlation is higher for government than for the judicial system, reflecting the judicial system's weaker relevance to employment policy.

One area where one would expect to see a much stronger relationship with trust in the judicial system than with trust in the government more generally is crime. If measures of institutional trust are valid, then variation in the crime rate should be more strongly linked to trust in the judicial system than to trust in government as a whole. This is exactly what is shown in Figure 2.12. Using homicide rates per 100 000 as a proxy for the prevalence of

Figure 2.12. **Institutional trust and homicide rates in OECD countries, 2006-15**

Homocides per 100.000 inhabitants

r=-0.02   % of respondents indicating confidence in national government, GWP

Homocides per 100.000 inhabitants

r=-0.23   % of respondents indicating confidence in the judicial system, GWP

Note: The Gallup World Poll asks the following two questions: "In this country, do you have confidence in each of the following, or not? How about national government? How about judicial systems and courts?"
Source: OECD calculations based on the Gallup World Poll (database) *www.gallup.com/services/170945/world-poll.aspx* and UNDOC Homicide Statistics (database) *www.unodc.org/gsh/en/data.html* as contained in the OECD Trust Database.

StatLink ⟐⟐ *http://dx.doi.org/10.1787/888933583785*

crime, the figure shows that there is essentially no correlation between crime rates and trust in government, but a significant negative correlation with trust in the judicial system (r = -0.23).[12]

A final check on the convergent validity of measures of institutional trust is to look at the relationship between these measures and overall life satisfaction. A positive correlation between trust in institutions and life satisfaction could be expected both because trusted institutions function better and are therefore more likely to be associated with other outcomes that are important to people's life satisfaction and because trustworthy government is of direct importance for how people value their lives (Frey and

Stutzer, 2005, 2006). Figure 2.13 shows a strong positive relationship between overall life satisfaction and both trust in government and trust in the judicial system. Although this relationship is a little stronger for trust in the judicial system (r = 0.57) than for trust in government (r = 0.43), it is lower for both than is the case for the correlation between generalised trust and life satisfaction, illustrated in Figure 2.8 (r = 0.77).

Figure 2.13. **Institutional trust and life satisfaction in OECD countries, 2006-15**

r=0.43   % of respondents indicating trust in national government, EB

r=0.57   % of respondents indicating trust in the judicial system, EB

Note: The Eurobarometer asks the following two questions: "I would like to ask you a question about how much trust you have in certain institutions. For each of the following institutions, please tell me if you tend to trust it or tend not to trust it? National government. The legal system. "Life satisfaction scores refer to a country's average score on a response scale from 0-10 for the question": Overall, how satisfied are you with life as a whole these days?" where zero means "not at all satisfied" and ten means "completely satisfied".
Source: OECD calculations based on the European Commission: Eurobarometer (database) *http://ec.europa.eu/ commfrontoffice/publicopinion/index.cfm#p=1&instruments=STANDARD* and the Gallup World Poll (database) *www.gallup.com/services/170945/world-poll.aspx* as contained in the OECD Trust Database.

StatLink ⌗ᴸᴵˢⁿ *http://dx.doi.org/10.1787/888933583804*

Key messages

Validity of measures of interpersonal trust. In general, evidence for the validity of measures of generalised trust is strong, while there is relatively little evidence focusing specifically on the validity of measures of limited trust. Measures of generalised trust perform well in terms of face validity, construct validity and convergent validity. This holds whether measures are assessed at the cross-country level or at the level of individual responses. While there is some question over the test-retest reliability of some measures of generalised trust at the individual level, country-level results are highly reliable across different data sources and over time. Although the evidence base on the validity of generalised trust measures is less developed than would be ideal, it is sufficient to conclude that such measures are fit for purpose. In particular, measures of generalised trust can provide meaningful information on population well-being and social capital and on the drivers of other social and economic outcomes. With respect to limited trust, the lack of evidence is not perhaps quite such a significant problem as it may seem at first sight. While there is little direct evidence on limited trust, most limited trust questions are conceptually very similar to generalised trust questions, suggesting that more detailed analysis of validity might yield similar positive results.

Validity of measures of institutional trust. The evidence on the validity of measures of institutional trust is more mixed than is the case of interpersonal trust. While institutional trust measures generally perform well in terms of construct validity, the picture is much less clear with respect to face and convergent validity. The available evidence on non-response rates suggests that while still performing better than measures of income (which are themselves regularly collected as part of official statistics), measures of institutional trust have higher item-specific non-response rates than measures of interpersonal trust or typical demographic questions. Non-response rates are higher for institutions with which people have little day-to-day contact such as the United Nations or the European Union, indicating that a likely cause of low response rates is lack of knowledge on the part of respondents. In the case of convergent validity, the issue is more a complete lack of evidence rather than any specific evidence pointing to a lack of validity on the part of measures of institutional trust.

The limits of validity. Most of the discussion of validity has been couched in terms of evidence in favour of the validity of trust measures. To a large degree, this reflects the balance of the academic research on the subject, as the general picture is largely positive. However, some findings do suggest significant limitations with respect to the validity of measures of trust. One potential source of concern is the finding reported by Glaeser et al. (2000) and Lazzarini et al. (2004) that survey measures of generalised trust do not correlate with the standard measure of trust in the Trust Game, but only with trustworthiness. As discussed above, however, this result is overturned when other-regarding preferences (i.e. preferences over another individual's material payoffs, in addition to one's own) are controlled for (Cox, 2004; Capra et al., 2007) or when looking at cross-country averages as opposed to individual results (Johnson and Mislin, 2012).

A more significant issue raised by Delhey, Newton and Welzel (2011) is that there are significant cross-country differences in terms of how people interpret the Rosenberg generalised trust question. In particular, Delhey et al. find that, in the majority of countries,

the "most people" element of the Rosenberg generalised trust question refers to people whom the respondent does not know, while in a significant minority of countries respondents frame their answer in terms of family and people whom they know well; this is the case in particular of poorer countries and those with a Confucian culture. While the significance of the Delhey et al. findings is disputed (Uslaner 2002; Algan and Cahuc, 2013), a non-random cross-country pattern in how the question is interpreted is a significant cause for concern; not only does a systematic difference in how the generalised trust question is interpreted increase measurement errors, it would also imply a systematic bias in results that would make cross-country or cross-group comparisons difficult to interpret.

While, compared with measures of interpersonal trust, institutional trust does not face the issues associated with trust in strangers vs. trust in those you know, Schneider (2016) raises another important problem. In an analysis of institutional trust in Eastern European and former Soviet bloc countries, Schneider notes that the error term associated with measures of institutional trust is correlated across countries with what it is attempting to measure; in other terms, there is evidence that reported trust in institutions is "U-shaped" with respect to the trustworthiness of institutions. This implies that, initially, as institutional quality falls, people will report lower levels of trust in institutions; however, if government institutions become more corrupt, respondents become unwilling to report a lack of trust in government, presumably reflecting a lack of trust in the institution collecting the information. In contrast to the issue raised with generalised trust by Delhey et al., where the actual impact on the performance of the measure appears to be limited, the problems raised by Schneider with respect to institutional trust are significant.

Over and above the three issues already discussed, there is one more significant point that needs to be stressed, i.e. the lack of evidence with respect to many of the key points relating to the validity of trust measures. With respect to generalised trust, the evidence base is generally fairly strong, and gaps are limited to a relatively narrow range of issues. In particular, more evidence on the test-retest reliability of interpersonal trust measures would be valuable, as would a more systematic examination of the non-experimental aspects of convergent validity. While Knack (2001) and Naef and Schupp (2009) provide a useful overview of the available evidence, it would be valuable to see a more systematic quantitative investigation of convergent validity issues featuring data from surveys explicitly designed to test questions related to validity. To some degree, this is already happening (Naef and Schupp, 2009; Falk et al., 2015; *Trustlab*), and there is cause to be relatively optimistic here.

In the case of institutional trust, the gap in the available evidence on validity is more severe. There is very little evidence on the convergent validity of institutional trust measures. Given that the evidence from face validity and construct validity is ambiguous, this is a major limitation. While the balance of evidence that measures of institutional trust *do not* work is not strong, studies that could provide decisive support for the validity of institutional trust measures are still lacking. More importantly, there is little to suggest that improvements in this area are currently a major area of focus. Most of the recent developments in experimental measures focus on interpersonal trust rather than institutional trust, and those experimental studies that do exist generally have limited wider applicability (e.g. Penczynski and Santana, 2016). Until this gap is filled, a lack of evidence remains a barrier to higher confidence in the validity of measures of institutional trust.

2.5. Conclusion

Key points made in this chapter are as follows:

● Interest in levels of trust in society has never been higher, and this creates a strong incentive to find ways to measure trust: both trust in other people and trust in institutions are key ingredients for societies to function. From the perspective of national statistical agencies, the key question is whether such interest should translate into a systematic effort to measure trust. Relevance and accuracy are the two key criteria for national statistical agencies to inform this decision, and it is important that any proposed official statistical measures meet both criteria. Relevance is important because it reflects whether the data collected are important to users and whether it can contribute to informing policy makers and the public more generally, while accuracy is important because data can usefully contribute to decision making only if it provides a true picture of the issues it is intended to help assess.

● The relevance of measures of trust is not in doubt. Measures of interpersonal trust – particularly generalised trust – are of fundamental importance to assessing the well-being of societies, to measuring social capital, and to understanding the drivers of other social and economic outcomes. This is reflected both in the large and growing literature on the drivers and consequences of generalised trust and in the wide range of national and international initiatives that include generalised trust as an outcome of interest, ranging from national initiatives to measure well-being such as those of the United Kingdom and Israel through to the UN Sustainable Development Goals. Similarly, there is a high level of interest in institutional trust, which is reflected both in measurement initiatives and in a large and varied academic literature. In addition to being important for measuring well-being, understanding institutional trust is essential to understanding government effectiveness and the functioning of democratic systems of government.

● The accuracy of trust measures is less clear. In general, the evidence for the validity of measures of generalised trust is strong, while there is relatively little evidence focusing on the validity of limited trust measures. Measures of generalised trust perform well in terms of face validity, construct validity and convergent validity. This holds whether the measures are assessed at the cross-country level or at the level of individual responses. While there is some question over the test-retest reliability of some measures of generalised trust at the individual level, country-level results are highly reliable across different data sources and over time. As with all intangible concepts, measuring generalised trust raises a number of issues about respondents' interpretation of the question in front of them and their subjective judgement, but these issues are not intractable. There is good evidence that, despite these issues, questions on generalised trust produce valid data, and there is extensive research that is providing new insights into the remaining measurement issues. In fact, it is likely that more is understood about the validity of measures of generalised trust than is the case for many more traditional elements of official statistics. From this perspective, measures of generalised trust can be considered as fit for purpose and should be measured in official surveys where relevant.

● The picture for measures of institutional trust is more mixed than is the case for measures of interpersonal trust. While institutional trust measures generally perform relatively well in terms of construct validity, the situation is less clear with respect to face validity and convergent validity. There is thus good reason to believe that such measures might be biased in some circumstances, and for a number of key aspects of validity there is simply

no evidence one way or the other. For this reason, despite the relevance of measures of institutional trust, it is less easy to state that they are fit for purpose within the context of official statistics. This does not mean that such measures have no place in the official statistical system, however. Rather, it suggests that such measures should be considered more experimental and should be implemented in contexts where their experimental status is clear; this is particularly important for national statistical offices. On the other hand, the relevance of such measures suggests that they should be a high priority for further research, both in the academic community and within national statistical agencies. Many of the key methodological questions regarding the accuracy of measures of institutional trust will require the sample size and response rates that only national statistical offices are able to deliver. Progress is already being made on some of these issues, which will be the focus of Chapter 3.

Notes

1. A search of the Econlit database for the terms trust or trustworthiness between 2000 and 2016 returns 5 242 records. A similar search on Google Scholar for trustworthiness returns 89 200 results over the same period. These findings illustrate both the size of the literature on trust and the degree to which looking at only a single discipline – economics – will miss much of the available evidence and understanding. Not surprisingly, a literature this diverse produces many definitions of trust.

2. Hardin (2004) does not consider the possibility of an institution being the trusting party rather than the party being trusted, but his framing could logically be extended in this way.

3. Evidence from the German Satisfaction with Government Services Survey (Zufriedenheitsbefragung 2015), which found large differences in trust between individual public agencies, supports this argument.

4. Here, trusting behaviour refers to the person engaging in trust by taking a risk on the expectation of another person's behaviour, while trustworthy behaviour consists of responding in a way consistent with expectations.

5. A range of other interactive games have been used to measure preferences and social norms other than trust; these include altruism, through the Dictator Game; co-operation and pro-social behaviour, through the public goods game; reciprocity, through the ultimatum game; and risk aversion, through lotteries. For an overview of games used in experimental economics, see Smith (2008).

6. In psychology, for instance, a survey by Arnett (2008) found that 96% of subjects in studies published in top journals were from WEIRD (Western, Educated, Industrialized, Rich and Democratic) backgrounds. Researchers – often implicitly – assume that either there is little variation in experimental results across populations, or that these WEIRD standard subjects are as representative of the human species as any other population. This is not the case, as Heinrich et al. (2010) conclude in their review of studies conducted across the behavioural sciences that WEIRD subjects are among the *least* representative populations one could find for generalising about humans, and that there is substantial variability of results across countries.

7. See also Benz and Meier (2006) for a discussion of the altruistic behaviour of Swiss students in lab experiments as compared to in the real world.

8. *Trust in different public institutions* has been suggested as a possible indicator that could be used in the future for the monitoring of Target 16.6 ("Develop effective, transparent and accountable institutions at all levels") of the UN 2030 Agenda (*https://unstats.un.org/unsd/statcom/48th-session/documents/2017-2-IAEG-SDGs-E.pdf*).

9. Measures of affect can be thought of as measures of particular feelings or emotional states such as *joy*, *contentment*, *anger*, or *worry*, and are often collected with reference to a particular point in time. In Boarini et al. (2012), affect balance is calculated as the number of positive emotions reported by a respondent on the previous day less the number of negative emotions, averaged across all respondents.

10. Contingent validity is a fourth criterion of validity that applies where validity can be directly assessed. For example, a measure of taxes paid has contingent validity if it is compiled from the complete set of tax payments done by a person. This measure is necessarily valid *contingent* on using a data source that directly measures the concept of interest.

11. Using an index derived from the different questions on interpersonal trust – the Rosenberg question, whether people would take advantage of you if they got the chance, and would you say that most of the time people try to be helpful – Glaeser et al. (2000) find an insignificant negative relationship between survey-based measures of trust and experimental trust.

12. Although this correlation is low in absolute terms, this may reflect the limits of the homicide rate as a proxy for crime more generally: homicide is both a relatively rare crime (with rates that can vary a lot from year to year with only small changes in the actual number of homicides) and not necessarily a strong proxy for non-violent crime.

References

Aberson, C.L., C. Shoemaker and C. Tomolillo (2004), "Implicit bias and contact: The role of interethnic friendships", *The Journal of Social Psychology*, Vol. 144, No. 3.

Algan, Y. and P. Cahuc (2013), "Trust, growth and well-being: New evidence and policy implications", *IZA Discussion Paper,* No. 7464.

Algan, Y. and P. Cahuc (2010), "Inherited trust and growth", *The American Economic Review,* Vol. 100, No. 5, pp. 2060-2092.

Alkire, S. (2002), "Dimensions of Human Development", *World Development,* Vol. 30, No. 2, pp. 181-205.

Almond, G. and S. Verba (1963), *The Civic Culture: Political Attitudes and Democracy in Five Nations*, Sage Publications, London.

Alsop, R., M. Bertelsen and J. Holland (2006), "Empowerment in practice: From analysis to implementation", *Directions in Development*, World Bank, Washington, DC.

Arnett, J.J. (2008), "The neglected 95%: Why American psychology needs to become less American", *American Psychologist,* Vol. 63, No. 7, pp. 602-614.

Arrow, K. (1972), "Gifts and exchanges", *Philosophy & Public Affairs,* Vol. 1, No. 4, pp. 343-362.

Benz, M. and S. Meier (2006), "Do people behave in experiments as in the field? Evidence from donations", *Federal Reserve Bank of Boston Working Papers,* No. 06-8.

Berg, J., J. Dickhaut and K. McCabe (1995), "Trust, reciprocity, and social history", *Games and Economic Behavior,* Vol. 10, pp. 122-142.

Boarini, R. et al. (2012), "What makes for a better life? The determinants of subjective well-being in OECD countries: Evidence from the Gallup World Poll", *OECD Statistics Working Papers,* No. 2012/03, OECD Publishing, Paris, *http://dx.doi.org/10.1787/5k9b9ltjm937-en*.

Boreham, R. et al. (2004), "Neighbourhood social cohesion and health: Investigating associations and possible mechanisms", *Social Capital for Health: Issues of definition, Measurement and Links to Health*, A. Morgan and C. Swann (eds.), Health Development Agency, London.

Bouckaert, G. (2012), "Trust and public administration", *Administration*, Vol. 60, No. 1, pp. 91-115.

Bouckaert, G. and S. Van de Walle (2003), "Comparing measures of citizen trust and user satisfaction as indicators of 'good governance': Difficulties in linking trust and satisfaction indicators", *International Review of Administrative Sciences*, Vol. 69, No. 3, pp. 329-343.

Bourdieu, P. (1984), *Distinction: A Social Critique of the Judgement of Taste*, Routledge & Keegan Paul (eds), London.

Capra, C.M., K. Lanier and S. Meer (2007), "Attitudinal and behavioral measures of trust: A new comparison", *Working Paper*, Department of Economics, Emory University, Atlanta.

Coleman, J. (1988), "Social capital in the creation of human capital", *American Journal of Sociology*, Vol. 94, Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure, pp. S95-S120.

Cook, T.E. and P. Gronke (2005), "The skeptical American: Revisiting the meanings of trust in government and confidence in institutions", *Journal of Politics, Vol.* 67, No. 3, pp. 784-803.

Cox, J. (2004), "How to identify trust and reciprocity", *Games and Economic Behaviour,* Vol. 46, pp. 260-281.

Dasgupta, N. and S. Asgari (2004), "Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping", *Journal of Experimental Social Psychology*, Vol. 40, No. 5, pp. 642-658.

Daude, C., H. Gutiérrez and A. Melguizo (2012), "What drives tax morale?", *OECD Development Centre Working Papers*, No. 315, OECD Publishing, Paris, *http://dx.doi.org/10.1787/5k8zk8m61kzq-en*.

Davidov, E. and L. Coromina (2013), "Evaluating measurement invariance for social and political trust in Western Europe over four measurement time points (2002-2008)", *ASK. Research and Methods*, Vol. 22, pp. 37-54.

Delhey, J., K. Newton and C. Welzel (2011), "How general is trust in 'most people'? Solving the radius of trust problem", *American Sociological Review*, Vol. 76, No. 5, pp. 786-807.

Eurobarometer (2016), *Eurobarometer Data Service*,*www.gesis.org/eurobarometer-data-service/search-data-access/data-access* (accessed 06 February 2017).

Eurofund (2016), *European Quality of Life Survey*, *www.eurofound.europa.eu/surveys/european-quality-of-life-surveys* (accessed 06 February 2017).

European Social Survey (2016), *European Social Survey*, *www.europeansocialsurvey.org/data/* (accessed 06 February 2017).

Eurostat (2016), *European Union Statistics on Income and Living Conditions (EU SILC)*, *http://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions* (accessed 23 March 2017).

Falk, A. et al. (2016), "The preference survey module: A validated instrument for measuring risk, time, and social preferences", *Working Paper*, University of Bonn.

Falk, A. et al. (2015), "The nature and predictive power of preferences: Global evidence", *IZA Discussion Paper*, No. 9504.

Fehr, E. (2009), "On the economics and biology of trust", *Journal of the European Economic Association*, Vol. 7, No. 2-3, pp. 235-266.

Fehr, E. et al. (2003), "A nation-wide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative surveys", *CESifo Working Paper Series*, No. 866; *IZA Discussion Paper*, No. 715.

Felte, E. (2001), "Finders keepers", *Reader's Digest*, 2 April.

Fleche, S., C. Smith and P. Sorsa (2012), "Exploring determinants of subjective well-being in OECD countries: Evidence from the WVS", *OECD Statistics Working Papers*, No. 2012/01, OECD Publishing, Paris, *http://dx.doi.org/10.1787/5k9ffc6p1rvb-en*.

Frey, B.S. and A. Stutzer (2006), "Political participation and procedural utility: An empirical study", *European Journal of Political Research*, Vol. 45, No. 3, pp. 391-418.

Frey, B.S. and A. Stutzer (2005), "Beyond outcomes: measuring procedural utility", *Oxford Economic Papers*, No. 57, No. 1, pp. 90-111.

Frey, B.S., M. Benz and A. Stutzer (2004), "Introducing procedural utility: Not only what, but also how matters", *Journal of Institutional and Theoretical Economics*, Vol. 160, No. 3, pp. 377-401.

Fukuyama, F. (1995), *Trust: The Social Virtues and the Creation of Prosperity*, Hamish Hamilton, London.

Gachter, S., B. Herrmann and C. Thoni (2004) "Trust, voluntary cooperation, and socioeconomic background: Survey and experimental evidence", *Journal of Economic Behavior & Organization*, Vol. 55, No. 4, pp. 505–531.

Gallup (2016), *Gallup World Poll*, *www.gallup.com/services/170945/world-poll.aspx* (accessed 06 February 2017).

Ginn, J. and S. Arber (2004), "Gender and the relationship between social capital and health", *Social Capital for Health: Issues of definition, Measurement and Links to Health*, A. Morgan and C. Swann (eds.) Health Development Agency, London.

Glaeser, E. et al. (2000), "Measuring trust", *The Quarterly Journal of Economics*, Vol. 115, No. 3, pp. 811-846.

Gonzalez, S. and C. Smith (2017), "The Accuracy of Measures of Institutional Trust in Household Surveys: Evidence from the OECD Trust Dataset", *OECD Statistics Working Papers*, OECD Publishing, Paris.

Greenwald, A.G., B.A. Nosek and M.R. Banaji (2002), "Understanding and using the implicit association test: I. An improved scoring algorithm", *Journal of Personality and Social Psychology*, Vol. 85, No. 2, pp. 197-216.

Guiso, L., P. Sapienza and L. Zingales (2008), "Trusting the stock market", *The Journal of Finance*, Vol. 63, No. 6, pp. 2557-2600.

Guiso, L., P. Sapienza and L. Zingales (2006), "Does culture affect economic outcomes?", *The Journal of Economic Perspectives,* Vol. 20, No. 2, pp. 23-48.

Hamilton, K. and G. Liu (2013), "Human capital, tangible wealth, and the intangible capital residual", *OECD Statistics Working Papers,* No. 2013/02, OECD Publishing, Paris, *http://dx.doi.org/10.1787/5k4840h633f7-en.*

Hardin, R. (2004), *Trust and Trustworthiness*, Russel Sage Foundation, New York.

Helliwell, J., H. Huang and S. Wang (2016), "The distribution of world happiness", *World Happiness Report 2016, Update (Vol. I)*, Helliwell, J., R. Layard and J. Sachs (eds), Sustainable Development Solutions Network, New York.

Helliwell, J. and S. Wang (2010), "Trust and well-being", *NBER Working Papers,* No. 15911, National Bureau of Economic Research.

Helliwell, J. and R. Putnam (2004), "The social context of well-being", *Philosophical Transactions of the Royal Society,* Vol. 359, pp. 14.

Henrich, J., S. Heine and A. Norenzayan (2010), "Most people are not WEIRD", *Nature,* Vol. 446, p. 29.

Intawan, C. and S.P. Nicholson (2017), "My trust in government is implicit: Automatic trust in government and system support", unpublished manuscript.

ISTAT (2013), *Benessere Equo e Sostenibile,* Rome.

Johnson, N. and A. Mislin (2012), "How much should we trust the WVS trust question?", *Economics Letters,* Vol. 116, pp. 210-212.

Johnson, N. and A. Mislin (2011), "Trust Games: A meta-analysis", *Journal of Economic Psychology,* Vol. 32, No. 5, pp. 865-889.

Karlan, Dean, S. (2005), "Using experimental economics to measure social capital and predict financial decisions", *American Economic Review*, Vol. 95, No. 5, pp. 1688-1699.

Kawachi, I. and L. Berkman (eds.), (2000), *Social Epidemiology*, Oxford University Press, New York.

Klijn, E.H., J. Edelenbos and B. Steijn (2010), "Trust in governance networks: Its impact and outcomes", *Administration and Society*, Vol. 42, No. 2, pp. 193-221.

Knack, S. (2001), "Trust, associational life, and economic performance", *MPRA Paper,* No. 27247, University of Munich.

Knack, S. and P. Keefer (1997), "Does social capital have an economic payoff? A cross-country investigation", *Quarterly Journal of Economics,* Vol. 112, No. 4, pp. 1251-88.

La Porta, R. et al. (1997), "Trust in large organisations", *American Economic Review, Papers and Proceedings,* Vol. 87, pp. 333-38.

Lazzarini, S.G. et al. (2004), "Measuring trust: An experiment in Brazil", *Insper Working Papers,* Vol. 049/2004.

Marien, S. and M. Hooghe (2011), "Does political trust matter? An empirical investigation into the relation between political trust and support for law compliance", *European Journal of Political Research,* Vol. 50, No. 2, pp. 267-291.

McKnight, D.H., V. Choudhury and C. Kacmar (2002), "The impact of initial consumer trust on intentions to transact with a web site: A trust-building model", *The Journal of Strategic Information Systems*, Vol. 11, No. 3, pp. 297-323.

Ministry for Environmental Protection, Israel (2014), *Well-being Indicators for Israel*, State of Israel, Jerusalem.

Morrone, A., N. Tontoranelli and G. Ranuzzi (2009), "How good is trust? Measuring trust and its role for the progress of societies", *OECD Statistics Working Papers,* No. 2009/03, OECD Publishing, Paris, *http://dx.doi.org/10.1787/220633873086.*

Murphy, K. (2004), "The role of trust in nurturing compliance: A study of accused tax avoiders", *Law and Human Behaviour,* Vol. 28, No. 2, pp. 187-209.

Murphy, K., T.R. Tyler and A. Curtis (2009), "Nurturing regulatory compliance: Is procedural justice effective when people question the legitimacy of the law?", *Regulation and Governance,* Vol. 3, No. 1, pp. 1-26.

Naef, M. and J. Schupp (2009), "Measuring trust: Experiments and surveys in contrast and combination", *IZA Discussion Paper,* No. 4087.

Nannestad, P. (2008), "What have we learned about generalized trust, if anything?", *Annual Review of Political Science,* Vol. 11, pp. 413-436.

New Zealand Ministry of Justice (2014), *New Zealand Crime and Safety Survey, www.justice.govt.nz/justice-sector-policy/research-data/nzcass/* (accessed 21 April 2017).

Nooteboom, B. (2007), "Social capital, institutions and trust", *Review of Social Economy,* Vol. 65, No. 1, pp. 29-53.

OECD (2017a), *Trust and Public Policy: How Better Governance Can Help Rebuild Public Trust*, OECD Publishing, Paris, *http://dx.doi.org/10.1787/9789264268920-en.*

OECD (2017b), *Measuring Distance to the SDGs Targets: An Assessment of where OECD Countries Stand*, OECD Publishing, Paris, *www.oecd.org/std/OECD-Measuring-Distance-to-SDG-Targets.pdf.*

OECDStat (2016), *OECD Statistics, https://stats.oecd.org/* (accessed 22 April 2017).

OECD (2015), *How's Life? 2015: Measuring Well-being*, OECD Publishing, Paris, *http://dx.doi.org/10.1787/how_life-2015-en.*

OECD (2013a), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing, Paris, *http://dx.doi.org/10.1787/9789264191655-en.*

OECD (2013b), *Government at a Glance*, OECD Publishing, Paris, *http://dx.doi.org/10.1787/gov_glance-2013-en.*

OECD (2013c), *How's Life? 2013: Measuring Well-being*, OECD Publishing, Paris, *http://dx.doi.org/10.1787/9789264201392-en.*

OECD (2011), *How's Life*, OECD Publishing, Paris, *http://dx.doi.org/10.1787/9789264121164-en.*

OECD (2008), "Quality Framework and Guidelines for OECD Statistical Activities", OECD Publishing, Paris.

OECD (2001), *The Well-being of Nations*, OECD Publishing, Paris, *http://dx.doi.org/10.1787/9789264189515-en.*

ONS (2016), *Measuring National Well-being in the UK: Domains and Measures*, Office for National Statistics, UK, London.

Penczynski, S. and M. Santana (2016), *Measuring Trust in Institutions: A Lab-in-the-Field Study Using Time Preference Elicitation*, University of Mannheim.

Putnam, R. (1993), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press, New Jersey.

Reeskens, T. and M. Hooghe (2008), "Cross-cultural measurement equivalence of generalized trust. Evidence from the European Social Survey, 2002 and 2004", *Social Indicators Research,* Vol. 85, pp. 515-532.

Rosenberg, M. (1957), "Misanthropy and attitudes toward international affairs", *Journal of Conflict Resolution*, pp. 340-345.

Roth, F. (2015), "Political economy of EMU. Rebuilding systemic trust in the euro area in times of crisis", *EC Discussion Papers, No. 16*, Brussels.

Rothstein, B. and E. Uslaner (2005), "All for all: Equality, corruption and social trust", *World Politics,* Vol. 58, No. 3, pp. 41-72.

Sampson, R. (2012), *Great American City: Chicago and the Enduring Neighbourhood Effect*, The University of Chicago Press, Chicago and London.

Schneider, I. (2016), "Can we trust measures of political trust? Assessing measurement equivalence in diverse regime types", *Social Indicators Research, http://dx.doi.org/10.1007/s11205-016-1400-8.*

Scrivens, K. and C. Smith (2013), "Four interpretations of social capital: An agenda for measurement", *OECD Statistics Working Papers,* No. 2013/06, OECD Publishing, Paris, *http://dx.doi.org/10.1787/5jzbcx010wmt-en.*

Sen, A.K. (2009), *The Idea of Justice,* Harvard University Press, Cambridge.

Sen, A.K. (1992), *Inequality Re-examined,* Harvard University Press, Cambridge.

Shanahan, E. (2007), "Excuse me, is this your phone?", *Reader's Digest,* July.

Smith, V.L. (2008), "Experimental economics", *The New Palgrave Dictionary of Economics*, 2nd Edition.

Stiglitz, J.E., A. Sen and J.-P. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, *www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.*

Sturgis, P. and P. Smith (2010), "Assessing the validity and reliability of generalized trust questions: What kind of trust are we measuring?", *International Journal of Public Opinion Research,* Vol. 22, No. 1.

Transparency International (2016), *Corruptions Perception Index, www.transparency.org/research/cpi/overview* (accessed 06 February 2017).

UNECE (2014), *Conference of European Statisticians Recommendations on Measuring Sustainable Development*, United Nations, New York and Geneva.

United Nations (2014), *Fundamental Principles of Official Statistics*, No. A/RES/68/261, United Nations General Assembly, New York.

United Nations (2015), "Transforming our world: The 2030 Agenda for Sustainable Development", No. A/RES/70/1, United Nations General Assembly, New York.

UNODC (2016), *Crime and Criminal Justice,* United Nations Office on Drugs and Crime, *www.unodc.org/ unodc/en/data-and-analysis/* (accessed 06 February 2017).

Uslaner, E. (2008), "Trust as a moral value", *The Handbook of Social Capital*, D. Castiglione, J. Van Deth and G. Wolleb (eds.), Oxford University Press, Oxford, pp. 101-121.

Uslaner, E. (2002), *The Moral Foundations of Trust*, Cambridge University Press.

Whitely, P. (1997), *Economic Growth and Social Capital*, Political Economy Research Centre, Sheffield.

World Bank (2006), *Where is the Wealth of Nations? Measuring Capital for the 21st Century*, The World Bank, Washington, DC.

Zak, P. and S. Knack (2001), "Trust and growth", *The Economic Journal,* Vol. 11, pp. 295-321.

Zelmer, J. (2003), "Linear public goods experiments: A meta-analysis", *Experimental Economics,* Vol. 6, pp. 299-310.

Zhao, H. and S. Kim (2011), "An exploratory examination of the social capital and FDI linkage and the moderating role of regulatory quality: A cross-country study", *Thunderbird International Business Review,* Vol. 53, No. 5, pp. 629-646.

From:
OECD Guidelines on Measuring Trust