Relativity and Habit Formation in Life Satisfaction*

Andreas Aristidou†, Arie Kapteyn‡

January 2019

Abstract

We examine the determinants of life satisfaction. While past literature has utilized reduced form cross-sectional and time-series models, we examine life satisfaction partly through a structural perspective. We propose a structural model of how an individual evaluates her life satisfaction as a linear combination of satisfaction with k domains of life, each of which is evaluated separately. Our main focus is on the modeling of satisfaction with income. In our model, individuals evaluate their income satisfaction as their ranking in a perceived income distribution. In this perceived income distribution, two critical components are the reference weights, representing the weight that the incomes of others play in one's perception of incomes, and a memory function, representing the weight that past perceived income distributions play in one's perceived income distribution. We apply this framework to an innovative dataset from the Gallup World Poll, which contains a number of vignettes describing life situations of hypothetical individuals. The model is used to explain how respondents rate the life satisfaction of the hypothetical individuals. As a result we are able to estimate the importance of relativity concerns and habit formation in determining the rating of life satisfaction. Simulations illustrate how life satisfactions in countries are affected by economic growth, income inequality and one's position in the income distribution.

*This research was supported by grant 1 R01 AG036784-01 from the National Institute on Aging. Zhishuo Han provided excellent research assistance.

†University of Southern California (aaristid@usc.edu)

‡University of Southern California, Center for Economic and Social Research (kapteyn@usc.edu)

1 Introduction

Life satisfaction measures have been used for decades to inform researchers and policy-makers on individuals' subjective evaluation of their own well-being. While plenty of measures of life satisfaction have been proposed and are used in the literature, the most common measure used throughout the literature is Cantrils Ladder after Hadley Cantril (Cantril 1965). The question is, what reference points do people use to evaluate their own or others' life satisfaction? We postulate that rather than one reference point, individuals use a frame of reference to evaluate their own or others' situation. In this paper, we concentrate on the role of income in contributing to life satisfaction, while controlling for other dimensions that have been found in the literature to be important. The simplest hypothesis regarding the frame of reference used to judge satisfaction with income is to assume that individuals assess incomes by their rank in a subjective income distribution. This subjective income distribution is a convex combination of income distributions an individual has perceived over her lifetime. Furthermore, the perceived income distribution at any given moment includes an individual's own income, so that this formulation incorporates both habit formation and notions of relativity in evaluations.

In this paper we use data from the Gallup World Poll (GWP), including a special module with anchoring vignettes. The vignettes represent brief descriptions of hypothetical individuals with respect to their health, family situation, job, and income. Every respondent sees six vignettes (out of a set of twelve) and is asked to rate the life satisfaction of the person described in each vignette. A key element of the vignette descriptions is that the incomes are either equal to half the median income in a country, or the median, or twice the median. This allows us to relate the vignette ratings by respondents to the income distribution in their country of residence, as well as to the respondent's own income. The GWP is a cross section, so we don't observe the individual income history of respondents. We make the strong assumption that all incomes in a country have grown at the same rate. As an estimate of the growth rate of incomes we then take observed GDP growth in the various countries in the dataset.

The remainder of the paper is structured as follows: The next section contains a brief discussion of the literature on life satisfaction. Section 3 lays out the formal model. Section 4 presents the data and the operationalization of a number of key concepts. Section 5 presents the results of the empirical analysis. Moreover, to provide intuition of the estimation results, we also provide some simulations, where we vary growth rates and income dispersion and show their impact on observed life satisfaction. Section 6 concludes.

2 Brief Literature Review

There is a vast literature on the relative nature of satisfaction with consumption or income. (Conti, Berndt et al. 2006). (Frank 2012) discusses evolutionary factors that would suggest that relative ranking should matter. For example, during a famine the strongest individuals would survive; the strongest individuals would be most likely to find a mate, etc. Primate studies have found that if they move up in the social hierarchy, their serotonin levels increase. Serotonin is related to feelings of happiness and well-being (van Vugt and Tybur 2015).

The most direct evidence that subjective well-being of humans is influenced by how well others are doing in an individuals reference group (roughly defined as: others the individual compares herself with) is provided in laboratory settings. For instance (Fliessbach, Weber et al. 2007) use functional magnetic resonance imaging (fMRI) to measure brain activity of subjects who have to perform an estimation task and are rewarded according to their performance. Subjects are not only informed about their own payments but also about the payments of other subjects. It is shown that neurophysiological activity responds strongly to relative payments. Outside the laboratory, establishing the effect of relative comparisons requires the definition of reference groups, i.e. groups of people one compares oneself to (Dahlin, Kapteyn et al. 2014). Reference groups have traditionally been defined a priori, e.g. by using individuals characteristics to define groups, for instance based on education, gender, or age (Van de Stadt, Kapteyn et al. 1985, McBride 2001, Ferrer-i-Carbonell 2005, Clark and Senik 2014). Similarly, coworkers or individuals in the same profession have been used as reference groups to explain the impact of individuals rank in the wage distribution on their satisfaction with the pay they receive (Brown et al., 2008).

The most commonly made assumption is that individuals mainly compare themselves to others in the same geographical area (Blanchflower and Oswald 2004, Luttmer 2004, Ferrer-i-Carbonell 2005, Barrington-Leigh and Helliwell 2008, Clark, Frijters et al. 2008). The definition of reference groups based solely on geography leads to complications, since for instance neighbors incomes are likely to be related to the quality of public and private goods in an area (access to parks, better stores, etc.). Using the Gallup Healthways Well-being Index survey, (Deaton and Stone 2013) regress SWB on log income (and average SWB on average income) at increasingly high levels of aggregation in the U.S. (individual, zip code, metro area, state). They note that if relative income is all that is important, then the coefficient on income should be rapidly declining as we move to higher levels of aggregation. Instead, they find that a regression of average SWB on average log-income at the zip-code level yields a higher coefficient than when using individual level data. When moving to larger geographic units the coefficient declines, but by only about one fifth compared to the individual level coefficient estimate. They control for age, sex, and race, but not societal goods. Also using the Gallup Healthways Well-being Index survey,(Ifcher, Zarghamee et al. 2016) find that neighbors incomes positively impact SWB in the U.S. at the local (zip code) level, but that at higher levels of aggregation (metro area), the effect is negative. This is true for several measures of well-being, including Cantrils ladder.

Yet, studies that try to elicit directly what reference groups respondents use find little evidence that geography is important. For instance, (Goerke and Pannenberg 2013) use pretest modules of the German Socio Economic Panel for the years 2008-2010, which contain questions about the importance of different groups for income comparisons. Their sample is restricted to employed respondents aged 17 to 65. They find that only colleagues at work, other people with the same occupation, and friends matter. (Dahlin, Kapteyn et al. 2014) also find very little support for the notion that reference groups would be primarily formed on the basis of geographical proximity.

It is safe to assume that for the vast majority of individuals their reference group will be limited to their own country. In the empirical analysis in this paper, we take advantage of that assumption by using data from a large number of countries.

3 Theory and Model

Life satisfaction is inherently multi-attribute and thus it is reasonable to expect that evaluation of life satisfaction is the combination of individual evaluations of the various components that make up life satisfaction. A simple way to represent this is in a linear equation taken from Kapteyn et. al., 2009:

$$W(x) = \beta_1 W_1(x_1) + \beta_2 W_2(x_2) + ... + \beta_K W_K(x_K) \tag{1}$$

Where x is a K x 1 vector of life dimensions, $W(\cdot)$ is the overall welfare evaluation function of life satisfaction and $k \in \{1, .., K\}$ indicates the K individual life domains. For instance we may interpret $W_k(x_k)$ as the evaluation of satisfaction with a value of x_k in life domain k. Without loss of generality we set $k = 1$ to represent income. For notational simplicity, from here on, we will replace x_1 with y. i.e. $W_1(y)$ would represent an individual's satisfaction with income y. We will refer to W_1 as a "welfare function of income (Van Praag 1971; Van Praag and Kapteyn, 1973). Finally the β parameters act as weights representing the relative importance of each individual welfare life domain evaluation on the aggregate welfare evaluation of life satisfaction.

3.1 The welfare function of income

In what follows we will focus mainly on the functional form of $W_1(\cdot)$ in which we will incorporate the effects of habit formation and relativity, first proposed in Kapteyn, 1977. The basic idea is that individuals use the incomes of others and themselves, both currently and in all past periods, as a frame of reference for evaluating their satisfaction with an income y. This evaluation is done on a finite scale, which we normalize to $[0, 1]$, where a higher number corresponds to a higher evaluation of satisfaction with income. We model time to be discrete: $\tau \in \{-\infty, .., 0\}$, with $\tau = 0$ representing the most recent period. At each time τ, the individual assigns time-independent reference weights q_i to incomes $y_{i\tau}$ of individuals $i = \{0, 1, ..., n\}$ in her reference group.[1]. The weights $\{q_0, q_1, ..., q_n\}$ must sum to one. Note that while the reference weights have no time subscript since we are assuming are constant across periods, this can easily be generalized at the cost of added notational complexity. Also, the development is done in terms of individuals rather than households; this can be generalized to an analysis of households by using equivalence scales. Our interest and analysis will be individual 0 (the self). Obviously, q_0 then represents the weight that the individual assigns to her income, denoted by $y_{0\tau}$ at time τ.

[1]Without loss of generality we can assume that any individual not in one's reference group is assigned weight of $q_i = 0$

Next, we proceed with defining notation and the four distributions which will play an important role in the development of our model.

3.2 Definitions

1. **Inclusive Perceived Income Distribution at Time τ:**

As is perhaps suggested by the name, this represents an income distribution that individual $i = 0$ perceives at time $\tau = \{0, -1, ..., -\infty\}$. The adjective "perceived" signifies that this distribution is not necessarily the true income distribution but rather the distribution that individual $i = 0$ perceives as the true income distribution. The adjective "inclusive" signifies that an individual's own income is part of the perceived income distribution (we will make the distinction between own income and income of others in definition 2). To motivate this one can think of an extreme case where individual $i = 0$ is currently living in isolation and only observes the income she receives. It might then be reasonable to assume that individual $i = 0$'s perception of the income distribution at time τ in the society is degenerate and equal to her own income $y_{0\tau}$. In our model, such a case would be represented by the individual assigning weights of $q_0 = 1$, $q_i = 0, \forall i \neq 0$. The inclusive perceived income distribution at time τ is a weighted empirical CDF, which includes the individual's own income ($y_{0\tau}$) with weight q_0, and the incomes of others ($y_{i\tau}$) with weights $q_i, i = \{1, 2, .., n\}$:

$$
\begin{aligned}
H(x|\tau) &= q_0 \mathbf{1}[y_{0\tau} \leq x] + (1 - q_0) F(x|\tau) \\
&= q_0 \mathbf{1}[y_{0\tau} \leq x] + \sum_{i=1}^{n} q_i \mathbf{1}[y_{i\tau} \leq x] \\
&= q_0 \mathbf{1}[y_{0\tau} \leq x] + (1 - q_0) \sum_{i \neq 0 | y_{i\tau} \leq x} \frac{q_i}{(1 - q_0)} \\
&= \sum_{i=0}^{n} q_i \mathbf{1}[y_{i\tau} \leq x]
\end{aligned}
$$

We refer to q_i as the **reference weight** that individual $i = 0$ places on the income of individual $i \in \{0, 1, .., n\}$. These weights capture our concept of **relativity** in our setup. Obviously we have that $\sum_{i=0}^{n} q_i = 1$. For later purposes it will be useful to define how individual $i = 0$ perceives the distribution of incomes excluding her own. We do that next.

2. **Perceived Income Distribution at Time τ:**

When we remove the "inclusive" portion of definition 1 we have the perceived income distirbution at time τ. This represents the income distribution that individual $i = 0$ perceives at time $\tau \in \{0, -1, ..., -\infty\}$ of all individuals in her reference group excluding herself $i \in \{1, ..., n\}$. While this definition might be less intuitive, given that it's hard to imagine an individual having knowledge of others incomes except her own, this distribution will still be very useful as it will help us identify the wedge between the weight an individual places on her own income and the incomes of others. Like definition 1, this is an empirical CDF.

$$
\begin{aligned}
F(x|\tau) &= \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \mathbf{1}[y_{i\tau} \leq x] \\
&= \sum_{i \neq 0 | y_{i\tau} \leq x} \frac{q_i}{(1-q_0)}
\end{aligned}
$$

Since this empirical CDF does not include $i = 0$'s income, we have that $\sum_{i=1}^{n} q_i \leq 1$, but $\sum_{i=1}^{n} \frac{q_i}{(1-q_0)} = 1$.

3. Inclusive Perceived Income Distribution:

This is a memory weighted convex combination of all inclusive perceived income distributions that individual $i = 0$ has perceived over her lifetime.

$$
H(x) = \sum_{\tau=-\infty}^{0} a(\tau) H(x|\tau)
$$

where $a(\tau)$ denotes the **memory function**, which represents the weight that individual $i = 0$ places on time period $\tau = \{0, -1, ..., -\infty\}$. The memory function generalizes the notion of **habit formation**, by not just including the effect of own past incomes, but also the incomes of all others who receive a non-zero reference weight. We impose $a(\tau) \geq 0, \forall \tau$ and $\sum_{\tau=-\infty}^{0} a(\tau) = 1$. A simple specification of the memory function, which we will use in the empirical analysis, is geometric decay of memory: $a(\tau) = (1-a)a^{-\tau}$, where $0 \leq a \leq 1$. In that case we call the parameter "a" the memory parameter. The higher its value the more past perceived income distributions matter in the perception of the current income distribution. At the one extreme of a tending to zero, the individual ignores all past perceived income distributions at $\tau = \{-1, -2, .., -\infty\}$ and thus $H(x) = H(x|0)$. At the other extreme of $a = 1$ the individual weights every period in her life equally.

4. Perceived Income Distribution:

We define the perceived income distribution (the analog of the inclusive perceived income distribution) as the memory weighted convex combination of all perceived income distributions that individual $i = 0$ has perceived over time of all individuals excluding herself.

$$F(x) = \sum_{\tau=-\infty}^{0} a(\tau)F(x|\tau)$$

3.3 Assumptions

To capture both of our concepts of habit formation and relativity we integrate the following hypothesis:

$$W_1(y) \equiv H(y) \tag{2}$$

This is the theory of preference formation developed in Kapteyn (1977). An individual's satisfaction with income (or her welfare function of income) is given by the position of her income in her inclusive perceived income distribution. In other words, an individual evaluates her income satisfaction by where it ranks in her inclusive perceived income distribution. The setup combines both habit formation (all own past incomes are represented in $H(y)$) and relative income considerations.

For later purposes it will be useful to put a parametric structure on $H(y)$. In line with the descriptive literature on the shape of income distributions, we will assume that the inclusive perceived income distribution is lognormal:

$$H(y) = \Phi\left(\frac{\ln(y) - \mu}{\sigma}\right) \tag{3}$$

Where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

3.4 Log-moments

The above formulation allows for an exploration of the effect on income satisfaction of both income growth and of one's position in the income distribution. To further spell out the implications of the model, it is useful to concentrate on the first two log-moments of the aforementioned distributions. We define those moments as follows: μ_τ [σ_τ^2] as the log-mean [log-variance] of the inclusive perceived income distribution at time τ, $H(x|\tau)$; m_τ [s_τ^2] as the log-mean [log-variance] of the perceived income distribution at time τ, $F(x|\tau)$; μ [σ^2] as the log-mean [log-variance] of the inclusive perceived income distribution, $H(x)$; and m [s^2] as the log-mean [log-variance] of the perceived income distribution, $F(x)$.

We first examine the above log-means of the perceived income distribution at τ, (m_τ) & inclusive perceived income distribution at τ, (μ_τ):

$$
\begin{aligned}
m_\tau &= \int_{-\infty}^{\infty} \ln(x) dF(x|\tau) \\
&= \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \ln(y_{i\tau})
\end{aligned}
$$

$$
\begin{aligned}
\mu_\tau &= \int_{-\infty}^{\infty} \ln(x) dH(x|\tau) \\
&= q_0 \ln(y_{0\tau}) + (1-q_0) m_\tau \\
&= \sum_{i=0}^{n} q_i \ln(y_{i\tau})
\end{aligned}
$$

and the log-means of the perceived income distribution, (m) & the inclusive perceived income distribution, (μ):

$$
\begin{aligned}
m &= \int_{-\infty}^{\infty} \ln(x) dF(x) \\
&= \sum_{\tau=-\infty}^{0} a(\tau) \int_{-\infty}^{\infty} \ln(x) dF(x|\tau) \\
&= \sum_{\tau=-\infty}^{0} a(\tau) m_\tau
\end{aligned}
$$

$$
\begin{aligned}
\mu &= \int_{-\infty}^{\infty} \ln(x) dH(x) \\
&= \sum_{\tau=-\infty}^{0} a(\tau) \int_{-\infty}^{\infty} \ln(x) dH(x|\tau) \\
&= \sum_{\tau=-\infty}^{0} a(\tau) \mu_\tau
\end{aligned}
$$

Next we investigate the log-variances of the perceived income distribution at τ, (s_τ^2) & inclusive perceived income distribution at τ, (σ_τ^2):

$$\begin{aligned}
s_\tau^2 &= \int_{-\infty}^{\infty} \big(\ln(x) - m_\tau\big)^2 dF(x|\tau) \\
&= \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \big(\ln(y_{i\tau}) - m_\tau\big)^2
\end{aligned}$$

$$\begin{aligned}
\sigma_\tau^2 &= \int_{-\infty}^{\infty} \big(\ln(x) - \mu_\tau\big)^2 dH(x|\tau) \\
&= \sum_{i=0}^{n} q_i \big(\ln(y_{i\tau}) - \mu_\tau\big)^2 \\
&= q_0 \big(\ln(y_{0\tau}) - \mu_\tau\big)^2 + (1-q_0) \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \big(\ln(y_{i\tau}) - \mu_\tau\big)^2 \\
&= q_0 \big(\ln(y_{o\tau}) - q_0 \ln(y_{o\tau}) - (1-q_0) m_\tau\big)^2 \\
&\qquad + (1-q_0) \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \big[\ln(y_{i\tau}) - q_0 \ln(y_{0\tau}) - (1-q_0) m_\tau\big]^2 \\
&= q_0 \big((1-q_0)(\ln y_{0\tau} - m_\tau)\big)^2 \\
&\qquad + (1-q_0) \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \big[(\ln(y_{i\tau}) - m_\tau) - q_0(\ln(y_{0\tau}) - m_\tau)\big]^2 \\
&= (1-q_0) \Big\{ q_0(1-q_0)(\ln y_{0\tau} - m_\tau)^2 + \underbrace{\sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \big(\ln(y_{i\tau}) - m_\tau\big)^2}_{=s_\tau^2} \\
&\qquad - 2q_0 \big(\ln(y_{o\tau}) - m_\tau\big) \underbrace{\sum_{i=1}^{n} \big(\ln(y_{i\tau}) - m_\tau\big)}_{=0} + q_0^2 \big(\ln(y_{0\tau}) - m_\tau\big)^2 \underbrace{\sum_{i=1}^{n} \frac{q_i}{(1-q_0)}}_{=1} \Big\} \\
&= (1-q_0)\big\{q_0(\ln(y_{o\tau}) - m_\tau)^2 + s_\tau^2\big\}
\end{aligned}$$

and the log-variances of the perceived income distribution, (s_2) & inclusive perceived income distribution, (σ^2):

$$
\begin{aligned}
s^2 &= \int_{-\infty}^{\infty} \big(\ln(x) - m\big)^2 dF(x) \\
&= \sum_{\tau=-\infty}^{0} a(\tau) \int_{-\infty}^{\infty} \big(\ln(x) - m\big)^2 dF(x|\tau)
\end{aligned}
$$

$$
\begin{aligned}
\sigma^2 &= \int_{-\infty}^{\infty} \big(\ln(x) - \mu\big)^2 dH(x) \\
&= \sum_{\tau=-\infty}^{0} a(\tau) \int_{-\infty}^{\infty} \big(\ln(x) - \mu\big)^2 dH(x|\tau) \\
&= \sum_{\tau=-\infty}^{0} a(\tau)\mathbf{E}_\tau\big[\ln(x) - \mu + \mu_\tau - \mu_\tau\big]^2 \\
&= \sum_{\tau=-\infty}^{0} a(\tau)\big[\mathbf{E}_\tau(\ln(x) - \mu_\tau)^2 + \mathbf{E}_\tau(\mu_\tau - \mu)^2 + 2\,\underbrace{\mathbf{E}_\tau((\ln(x) - \mu_\tau)}_{=0}(\mu_\tau - \mu))\big] \\
&= \sum_{\tau=-\infty}^{0} a(\tau)\big[\sigma_\tau^2 + (\mu_\tau - \mu)^2\big]
\end{aligned}
$$

Obviously, μ will be higher if past incomes were higher or if people in an individual's reference group have (or had) higher incomes. σ^2 will be higher if incomes in the individual's reference group have been more dispersed, if her income is further removed from the median and if incomes have varied a lot over time (so that the term $(\mu_\tau - \mu)^2$ tends to be large).

3.5 Examples

We can use this formulation to shed further light on one of the debates in the literature: does economic growth increase life satisfaction? Easterlin and coauthors in various papers [(Easterlin 1974) , (Easterlin, McVey et al. 2010), (Easterlin 2015)] have argued there is no long-term effect of economic growth on life satisfaction, although there are short-term business cycle effects. Our model is very much in the spirit of Easterlin's work and we will illustrate how the model generates stylized facts consistent with the work by Easterlin and his collaborators.

3.5.1 One-off income growth

The model suggested here certainly implies short run effects of changes in average income in society. To simplify the example assume that all incomes in society have been constant for a sufficiently long time so that we may assume that:

$$\mu = \mu_\tau;\ m = m_\tau;\ \sigma^2 = \sigma^2_\tau;\ s^2 = s^2_\tau$$

Further, we will impose a geometric decay of memory $a(\tau) = (1-a)a^{-\tau}$, $0 \geq a \geq 1$. Now consider an identical increase in income for all individuals in society by a factor $\exp(\alpha)$. It is then straightforward to derive satisfaction with income in the next period. The numerator in (3) becomes $\ln(y) - \mu + a\alpha$. This is intuitively plausible: a determines the speed of adjustment. The smaller a is, the quicker the welfare gain resulting from an income increase will dissipate. For the denominator in (3) we obtain the square root of $\sigma^2 + (1-a)\alpha^2$, so that the new level of satisfaction with income becomes:

$$W_1(y\exp(\alpha)) = \Phi\left(\frac{\ln(y) - \mu + a\alpha}{\sqrt{\sigma^2 + (1-a)\alpha^2}}\right) \tag{4}$$

The right hand side of equation (4) will be larger than equation (3) for positive α and smaller for negative α.[2] This illustrates that economic growth, defined as an equal proportional increase of all incomes will increase well-being in the short-run. However, the gain will dissipate over time as both numerator and denominator converge to their previous values.

3.5.2 Sustained income growth

The one time increase in income can be contrasted with the case were all incomes grow permanently at a rate α. Once again, to keep the discussion manageable it is useful to consider a special case. Let the log-mean and log-variance of the societal income distribution at time τ be m_τ and s^2_τ (which are also the moments of the perceived income distribution at time τ). For simplicity of notation denote $m \equiv m_0$, $s^2 = s^2_0$. We assume that all incomes grow at the same rate α. This additionally implies that:

$$\ln(y_{i\tau}) = \ln(y_{i\tau-1}) + \alpha;\ m_\tau = m_{\tau-1} + \alpha;\ s^2_\tau = s^2, \forall\tau$$

Which we can express in terms of current values:

$$\ln(y_{i\tau}) = \ln(y_{i0}) + \alpha\tau;\ m_\tau = m + \alpha\tau,\ \forall\tau$$

Using the above implications (and denoting current income with y, i.e. ommitting a subscript) we get that:

[2] This statement holds for values of of α that are not "too large". For extremely large values of α the increase in the value of the denominator may dominate the effect of the increase in the numerator.

$$\begin{aligned}
\mu &= \sum_{\tau=-\infty}^{0} a(\tau)\mu_\tau \\
&= \sum_{\tau=-\infty}^{0} (1-a)a^{-\tau}\big[q_0\ln(y) + (1-q_0)m_\tau\big] \\
&= \sum_{\tau=-\infty}^{0} (1-a)a^{-\tau}\big[q_0\ln(y) + (1-q_0)m + \alpha\tau\big] \\
&= (1-a)\Big\{\big[q_0\ln(y) + (1-q_0)m\big]\underbrace{\sum_{\tau=0}^{\infty} a^{\tau}}_{(1-a)^{-1}} + \alpha\underbrace{\sum_{\tau=0}^{\infty}(-\tau)a^{\tau}}_{a(1-a)^{-2}}\Big\} \\
&= q_0\ln(y) + (1-q_0)m + \frac{a\alpha}{1-a}
\end{aligned}$$

For σ^2, since we have that all incomes grow at the same rate we have that:

$$\sigma_\tau^2 = (1-q_0)\{q_0(\ln(y)-m)^2 + s^2\} \equiv \psi$$

Which in turn implies:

$$\begin{aligned}
\sigma^2 &= (1-a)\sum_{\tau=-\infty}^{0} a^{-\tau}\big\{\psi + (\mu_\tau - \mu)^2\big\} \\
&= \psi + (1-a)\alpha^2 \sum_{\tau-\infty}^{0} a^{-\tau}\big(\tau - \frac{a}{(1-a)}\big)^2 \\
&= \psi + \alpha^2\frac{a}{(1-a)^2}
\end{aligned}$$

Thus we obtain the following satisfaction with income:

$$W_1(y) = \Phi\left(\frac{(1-q_0)\ln(y) - \frac{\alpha a}{1-a}}{\sqrt{\psi + \alpha^2\frac{a}{(1-a)^2}}}\right) \tag{5}$$

As before, the numerator shows that income satisfaction will be larger for higher growth rates, but the effect is tempered by the fact that higher growth rates also increase the denominator. It is worth pointing out that in this scenario $W_1(y)$ is constant. In other words the positive growth rate does not imply a positive growth rate in life satisfaction. It only affects the level of satisfaction.

4 Empirical Analysis

4.1 Data description

We are using data collected by Gallup World Poll (GWP). Since 2005, the Gallup World Poll continually surveys residents in over 150 countries, interviewing about 1,000 randomly sampled individuals in each country. World Poll questions measure opinions about national institutions, corruption, youth development, community basics, diversity, optimism, violence, religiosity, and other topics. The World Poll questionnaire is translated into major languages of each country. The translation process starts with an English, French, or Spanish version, depending on the region. A translator proficient in both original and target languages translates the survey into the target language. A second translator reviews the language version against the original version and recommends refinements.

With some exceptions, all samples are probability-based and nationally representative of the resident population aged 15 and older. The coverage area is the entire country including rural areas, and the sampling frame represents the entire civilian, non-institutionalized, aged 15 and older population of each country. Exceptions include areas where safety of interviewing staff is threatened, scarcely populated islands, and areas interviewers can reach only by foot, animal, or small boat. Specifically, sampling in the Central African Republic, Democratic Republic of the Congo, Lebanon, Pakistan, India, Syria, Azerbaijan, Georgia, Morocco, Myanmar (Burma), Chad, Madagascar, Moldova, and Sudan was affected by security; some of these as well as Canada, China, Laos, and small parts of Japan had non-representative sampling of some geographic regions. In Arab countries (Bahrain, Kuwait, Saudi Arabia), sampling was of citizens (including Arab expatriates) and those who could complete the survey in Arabic or English; in the United Arab Emirates, all non-Arabs were excluded, i.e. more than half of the population. In the Philippines, urban areas were over-sampled. Israel excludes East Jerusalem (Gallup reports Palestinian Territories separately). We conduct various robustness specification tests excluding various non-representative sample countries, still our results are largely robust to these exclusions.

Telephone surveys are used in countries where telephone coverage represents at least 80% of the population or is the customary survey methodology. In Central and Eastern Europe and most of the developing world, an area frame design is used for face-to-face interviewing. In some countries, over-samples are collected in major cities or areas of special interest. In some large countries, such as China and Russia, samples of at least 2,000 are collected.

Gallup has created a worldwide data set with standardized income and education data. Similarly annual household income in international dollars is calculated using the Individual Consumption Expenditure corrected for the Household PPP ratio from the World Bank. These PPP-corrected values correlate strongly (r=0.94) with the World Bank estimate of per-capita GDP (PPP-corrected). The result is a household income measure that is comparable across all respondents, countries, and local and global regions.

Response rates are calculated according to AAPOR Standard Definitions (Callegaro and Disogra, 2008), and reported figures include completed and partial interviews, refusals, noncontacts, and unknown households. Gallup World Poll response rates vary by mode of survey and region. Response rates in Sub-Saharan Africa are higher than other world regions, ranging from a high of 96% in Sierra Leone to a low of 54% in Nigeria, with an average response rate of 80%. Average response rates for the Middle East, Asia, South America and former Soviet Union countries are 63%, 56%, 43%, and 50%, respectively.

As part of a National Institute on Aging supported project, Gallup added a module with anchoring vignettes to surveys conducted in 109 countries conducted during 2011-2014. These countries provide the sample for the current paper. Eighteen countries were interviewed in 2011, 39 in 2012, 26 in 2013, and 26 in 2014. Most countries have approximately 1,000 observations, with the exceptions of Russia (1,500), India (5,000), China (4,500), Germany (3,000), United Kingdom (3,000) and Haiti (500), for a total of about 120,000 observations.

While we are interested in many of the measures collected by the GWP, our primary interest is respondents' evaluation of life satisfaction. Gallup collects various such data like experiential well-being questions. However in this paper we focus on responses to the following question:[3]

"*Please imagine a ladder with steps numbered from zero at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?*"

4.1.1 Vignettes

The innovation of this dataset comes from the fact that in addition to responses to the above question, respondents were asked to rate the life satisfaction of six hypothetical people (vignettes) who varied in gender, income, age and life situation. While each respondent was presented with six vignettes, we have data on twelve vignettes in two groups of six (set A and set B). Each respondent was asked to rate either the six vignettes in set A or the six vignettes in set B. Table 1 summarizes the 12 vignettes, breaking down the description of each in six dimensions, *Income, Sex, Age, Health, Family, Job.*[4]. Within each country, roughly half of the respondents rated the set A vignettes and about half answered the set B vignettes. For each of the six vignettes, each respondent was asked to answer the following question:

"*Imagine again a ladder with steps numbered from zero at the bottom to 10 at the top, where zero is the worst possible life and 10 is the best possible life. I will now read you some descriptions of several peoples lives. Please tell me on which step of the ladder you think each person stands.*"

[3]The question is often referred to as the "Cantril's ladder" (Cantril, 1965) and has been used extensively to evaluate life satisfaction in the literature.

[4]For a complete description of the vignettes see the Appendix.

Table 1: Summary of vignette descriptions.

ID	Income	Sex	Age	Health	Family	Job
A1	Median	Female	40	Back pain	Good	-
A2	Twice median	Male	50	Good	Divorced; Good rel. with daughter	Secure
A3	Half median	Male	25	Good; Some stress	Single; Many friends	Worries about losing job
A4	Median	Female	35	-	Married; No children	Dull; Secure
A5	Half median	Female	70	Back pain	Widow; Many friends	-
A6	Twice median	Male	60	Very active	Single; Many friends	No work but happy with it
B1	Twice median	Male	40	Severe back pain	Happy; Good	Likes job
B2	Half median	Female	65	Heart problems	Widow; Misses husb. Good rel. w/ family	No job bec. of health
B3	Half median	Male	35	-	Married; No children	Dull; Secure
B4	Twice median	Female	60	Trouble sleeping	Divorced; Little contact with children	Interesting
B5	Median	Female	70	Overweight; Trouble walking	Married but lives separately from family	-
B6	Median	Male	50	Rarely excercise; Obese; Trouble with knees	Married but spend little time together	Secure

See the Appendix for the complete vignette descriptions.

The use of anchoring vignettes like the ones used here is not new in the social sciences and the medical field. Primarily researchers have utilized vignettes to homogenize heterogeneous use of response scales by individuals.[5] However, we take a different approach. We will utilize responses to vignette questions in order to estimate a model of the determinants of life satisfaction.

4.2 Adaptation to the data

While we utilize a rich and innovative dataset, in order to estimate our model we are forced to make a number of behavioral assumptions in order for our most important parameters of interest to be fully identified.

[5] For an excellent introduction to the field and a review of existing literature visit https://gking.harvard.edu/vign.

1. An important part of our model is relativity. Absent information about reference groups, we will assume that an individual places equal weight on the incomes of everyone else in her country: $q_i = \frac{1}{n}, \forall i \neq 0$. We will thus utilize m and s^2 (the log-moments of the (non-inclusive) perceived income distribution) as the log-moments of the observed income distribution (in our sample). To sum up, an individual only cares about incomes of people in her own country. Given this, she correctly perceives the non-inclusive income distribution. q_0 then becomes an object of estimation as the wedge between an individual's perceived income distribution and the observed income distribution in her country.

2. Another important part of our model is habit formation. However our data is cross-sectional, so we have no information on individual income histories. Our best proxy for the evolution of individual incomes is country-level GDP growth. Hence, we will assume that every individual's income has been growing at the rate of their respective country's GDP growth rate over the past ten years. This introduces a simplification, as it implies that the relative position of all incomes remains constant over time. Using this information and assumptions we can identify a, the memory decay parameter.

The vignettes have incomes that are $\kappa \in \{0.5, 1, 2\}$ times the median of the observed income distribution in a country, which we assume follows a lognormal distribution. Since we have assumed that our individuals have the correct perception of the (non-inclusive) income distribution in their country, that median equals (exp(m) and we have $\ln(\kappa * \exp(m)) = \ln(k) + m$. Finally, the framework developed in section 3,implies that an individual subjectively evaluates such an income level by the ranking in her income distribution $H(y)$, with parameters μ and σ:

$$W_1(\kappa * \exp(m)) = \Phi\left(\frac{\ln(\kappa * \exp(m)) - \mu}{\sigma}\right) = \Phi\left(\frac{\ln(\kappa) + m - \mu}{\sigma}\right) \tag{6}$$

Assumptions (1) and (2) imply that within an individual's reference group, incomes rise at the same rate α_τ. This simplifies some expressions.[6]

[6] We use a bar above an expression or part of an expression to denote that this part is constant over time periods.

$$
\begin{aligned}
y_{i\tau} &= \alpha_\tau * y_{i\tau-1}, \forall\tau \\
m_\tau &= \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \ln(y_{i\tau}) \\
&= \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \ln(\alpha_\tau) + \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \ln(y_{i\tau-1}) \\
&= \ln(\alpha_\tau) + m_{\tau-1}
\end{aligned}
$$

$$
\begin{aligned}
\ln(y_{i\tau}) - m_\tau &= \ln(\alpha_\tau) + \ln(y_{i\tau-1}) - \ln(\alpha_\tau) - m_{\tau-1} \\
&= \ln(y_{i\tau-1}) - m_{\tau-1} \\
&= \overline{\ln(y_{i\tau-1}) - m_{\tau-1}} \quad \text{(constant)}
\end{aligned}
$$

$$
\begin{aligned}
\mu_\tau &= \ln(\alpha_\tau) + \mu_{\tau-1} \\
&= q_0 \ln(y_{0\tau}) + (1-q_0) m_\tau \\
&= \overline{q_0[\ln(y_{0\tau}) - m_\tau]} + m_\tau \\
&= \overline{q_0[\ln(y_{0\tau}) - m_\tau]} + \ln(\alpha_\tau) + m_{\tau-1}
\end{aligned}
$$

$$
\begin{aligned}
\mu &= \sum_{\tau=-\infty}^{0} a(\tau)\mu_\tau \\
&= \sum_{\tau=-\infty}^{0} a(\tau) m_\tau + \overline{q_0[\ln(y_{0\tau}) - m_\tau]} \qquad (7)
\end{aligned}
$$

$$
\begin{aligned}
s_\tau^2 &= \sum_{i=1}^{n} \frac{q_i}{(1-q_0)} \big(\ln(y_{i\tau}) - m_\tau\big)^2 \\
&= s_{\tau-1}^2 \\
&\equiv \overline{s_\tau^2} \quad \text{(constant)}
\end{aligned}
$$

$$
\begin{aligned}
\sigma_\tau^2 &= (1-q_0)\{q_0(\ln(y_{0\tau}) - m_\tau)^2 + s_\tau^2\} \\
&= (1-q_0)\{q_0(\ln(y_{0\tau-1}) - m_{\tau-1})^2 + \bar{s_\tau^2}\} \\
&= \sigma_{\tau-1}^2 \\
&= \overline{\sigma_{\tau-1}^2} \equiv \Psi \quad \text{(constant)}
\end{aligned}
$$

$$\sigma^2 = \sum_{\tau=-\infty}^{0} a(\tau)\{\sigma_\tau^2 + (\mu_\tau - \mu)^2\}$$
$$= \Psi + \sum_{\tau=-\infty}^{0} a(\tau)\Big(\sum_{s=-\infty}^{0} a(s)m_s - m_\tau\Big)^2$$

The availability of data on GDP growth rates varies by country. For most countries we have estimates of GDP growth, going back in time by ten years. We make the additional assumption that growth has been zero beyond the nth period (where n=10). If a is small enough, this assumption may have little effect on the estimation results.

$$\sum_{\tau=-\infty}^{0} a(\tau)m_\tau = (1-a)\sum_{\tau=-\infty}^{0} a^{-\tau}m_\tau$$
$$= (1-a)\Big(\sum_{\tau=-n}^{0} a^{-\tau}m_\tau\Big) + (1-a)\Big(\sum_{\tau=-\infty}^{-(n+1)} a^{-\tau}m_{-n}\Big)$$
$$= (1-a)\Big(\sum_{\tau=-n}^{0} a^{-\tau}m_\tau\Big) + a^{n+1}m_{-n}$$
$$\equiv M$$

$$\Rightarrow \mu = M + \overline{q_0[\ln(y_{0\tau}) - m_\tau]}$$

$$\Rightarrow \sigma^2 = \Psi + \sum_{\tau=-\infty}^{0} a(\tau)(M - m_\tau)^2$$
$$= \Psi + (1-a)\sum_{\tau=-n}^{0} a^{-\tau}(M - m_\tau)^2 + a^{n+1}(M - m_n)^2 \tag{8}$$

4.3 Specification

We have thus far focused our attention on W_1, the welfare function of income, however for our main equation (1) we need to specify $K-1$ more functions W_k. We will take an agnostic view on the W_k for $\in \{2, 3, .., K\}$, and represent these domains by a collection of dummies to be specified below. Collecting all assumptions, the general equation that we will be estimating is the following:

$$W(x_{ivc}) = \beta_0 + \beta_1 W_{1ic}\big(\kappa_v * \exp(m_c)\big) + \beta_2 X_2 + \beta_3 X_3 + ... + \beta_k X_k + \epsilon_i \tag{9}$$

Where i is the individual identifier, $v \in \{A1, A2, A3, A4, A5, A6, B1, B2, B3, B4, B5, B6\}$ is the vignette identifier and c is the country identifier. $X's$ represent vignette, individual and country-level characteristics. The W_{1ic} function varies across individuals according to equation (6) and the assumptions made on the idividual parameters μ and σ. For a given value of κ, $W_1(x_1)$ is fully specified by the set of equations {(6), (7), (8)}. We use the information about the vignettes to define a set of dummy variables that capture the rest of the vignettes' life dimensions, in addition to the information about their income level. These dummies will serve as other x's. To complete the x vector, we also include a number of macro country-level variables collected from the World Bank (Life expectancy, Education expenditure, GDP per capita).[7], as well as a number of individual respondent characteristics. We estimate the parameters of the model using nonlinear least squares clustering at the country level.

5 Results

5.1 Summary statistics

Table 2 summarizes average responses for the twelve vignettes as well as the life satisfaction question.[8] (Respondents use an 11-point discrete scale [0-10]). For ease of exposition and interpretation, we provide the gender and income of each vignette.

Table 2: Average responses for Cantril's ladder - life satisfaction questions.

		Vignette identifier					
A-set	Self	$A1^{(F,m)}$	$A2^{(M,Tm)}$	$A3^{(M,Hm)}$	$A4^{(F,m)}$	$A5^{(F,Hm)}$	$A6^{(M,Tm)}$
Mean	5.36	4.46	6.39	4.03	5.27	3.95	6.6
SD	2.30	2.04	2.39	2.05	2.16	2.18	2.48
SD-Cntry	1.08	0.71	1.15	0.69	1.01	0.94	1.26
SD-Age_grp	-	0.12	0.30	0.05	0.23	0.24	0.35
B-set	Self	$B1^{(M,Tm)}$	$B2^{(F,Hm)}$	$B3^{(M,Hm)}$	$B4^{(F,Tm)}$	$B5^{(F,m)}$	$B6^{(M,m)}$
Mean	5.34	5.14	3.95	4.27	5.05	3.97	4.11
SD	2.30	2.18	2.17	2.17	2.16	1.95	2.00
SD-Cntry	1.10	0.75	0.88	0.88	0.64	0.54	0.59
SD-Age_grp	-	0.14	0.17	0.17	0.15	0.06	0.07

The "Self" column depicts average responses to the subjective evaluation of the own life satisfaction question. The other columns depict average ratings of each vignette. Each vignette identifier is superscripted with the corresponding (Gender, Income) characteristics. For example "$A1^{(F,m)}$" represents vignette A1 which describes a female earning median income (for full vignette description see Appendix). "SD" is the standard deviation across the whole sample, "SD-Cntry" is the standard deviation of the averages responses across countries. "SD-Age_grp" is the standard deviation across eight age groups, each with a 10-year range each: [15, 25], (25, 35], (35, 45], (45, 55], (55, 65], (65, 75], (75, 85], (85, 100].

The average responses to the life satisfaction question is around 5.35, almost identical for

[7] To test for robustness we utilize a different set of vignette dummies and individual-level variables. The results are largely unchanged. See Appendix.

[8] This is the question where respondents were asked to rate their own life satisfaction.

Table 3: Average responses for several variables of interest.

Variable	Observations	Mean	Std. Dev.	Min	Max
$\ln(HHincome)$	101,867	$-2.96*10^{-8}$	(0.87)	-8.86	5.19
Age-15	120,521	25.7	(17.4)	0	84
Female	120,823	0.54	(0.5)	0	1
Married	120,318	0.54	(0.5)	0	1
Divorced	120,318	0.036	(0.19)	0	1
Widowed	120,318	0.072	(0.26)	0	1
Domestic partner	120,318	0.05	(0.213)	0	1
Unemployed	117,819	0.069	(0.25)	0	1
Self employed	117,819	0.15	(0.35)	0	1
Urban	118,138	0.41	(0.49)	0	1
HH adults	120,743	3.08	(1.82)	1	96
HH Children	120,272	1.27	(1.8)	0	97
HH size	120,194	4.35	(2.89)	1	104
Education	120,236	1.81	(0.67)	1	3
Religion important	99,615	0.76	(0.43)	0	1
Home internet access	120,293	0.39	(0.49)	0	1
Life satisfaction in 5 years	111,201	6.65	(2.4)	0	10

The variable "$\ln(HHincome)$" represents the natural logarithm of the household income in international dollars. We use the acronym "HH" for household. The variable "Education" takes the value of 1 for elementary education or less, 2 for secondary and up to three years of tertiary education, 3 for four years of tertiary education or beyond. The variable "Life satisfaction in 5 years" is the same Cantril's ladder question as in Table 2 except that respondents were asked to imagine their life in 5 years time. The reason for the missing observations is that respondents had the option of declining to answer a particular question.

participants across the two sets. This is as expected, since participants were randomly assigned to sets A or B and the question asked prior to the vignette questions. The highest rated vignette is vignette A2 while the lowest ranked vignette is vignette B2. Interestingly, the vignettes in the set B receive significantly lower ratings from respondents compared tor the vignettes in set A. Since incomes, gender and ages are similar across the two sets, this difference is likely caused by the other vignette characteristics. Indeed, one can easily see (from Table 1 or the Appendix) that vignettes in set B, on average, appear to exhibit worse health, family, and job characteristics than the ones in set A. Responses to set A vignettes seem to have higher standard deviations across countries and age groups compared to vignettes in set B.

In addition to responses to vignette satisfaction questions that constitute our dependent variable, for our model estimations we make use of several other measures including log of household income, age, gender, etc. Table 3 summarizes the independent variables in our sample.

The average age in our sample is 40.7 years (25.7 + 15), female and/or married respondents correspond to about 54% of the sample, while less than 10% are either/and divorced, widowed, have a domestic partner and unemployed. Self employed constitute about 15%. For 76%, religion is important and just 39% have home internet access, while 41% live in urban areas. Interestingly, the average responses for the estimated life satisfaction five years after the interview date is much higher than the responses for current life satisfaction ($\sim$ 5.35 - Table 2).

5.2 Structural estimation

To estimate our model (equation (9)) we use nonlinear least squares clustering at the country-level. As starting values for the estimation we use $\{\beta_0, \beta_1, q_0, a\} = \{0, 1, 0.5, 0.8\}$. These values represent findings in past literature (see Kapteyn & Teppa, 2003, Van de Stadt et al., 1985). To ascertain robustness, the Appendix shows results based on different starting values. The estimates of the effects of the vignette characteristics are broadly plausible: vignettes are rated higher when health is better, and when the job is better. Bad family relations lead to lower evaluations, although our operationalization of a good family life seems to lead to a lower rating than the left out category. The estimated memory parameter a is close to one, which makes the approximation of setting GDP growth to zero for periods more than 10 years before the current data poorer than we would have liked. The estimate of q_0 is small compared to estimates from earlier papers (Van de Stadt et al., 1985). The country-level macro variables life expectancy and education expenditure seem to not significantly impact the vignette evaluation while GDP per capita is only significant at the 10% level and shows a positive impact on the vignette evaluation i.e. individuals from countries with a higher GDP per capita would tend to evaluate a vignette higher.

Table 2: Model estimation results.

	Main specification
β_0	-0.134
	(0.304)
β_1	0.274***
	(0.017)
q_0	0.07***
	(0.014)
a	0.98***
	(0.037)
$\beta_{HealthGood}$	0.033***
	(0.004)
$\beta_{HealthBad}$	-0.048***
	(0.004)
$\beta_{FamilyGood}$	-0.017***
	(0.003)
$\beta_{FamilyBad}$	-0.07***
	(0.008)
$\beta_{JobGood}$	0.017***
	(0.005)
β_{JobBad}	-0.03***
	(0.005)
$\beta_{ln(LifeExpectancy)}$	0.050
	(0.086)
$\beta_{ln(EducationExpenditure)}$	0.003
	(0.011)
$\beta_{ln(GDPperCapita)}$	0.035*
	(0.018)
N	405,100
R^2	0.202

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

In the Appendix, we provide additional model estimation results where we code vignette characteristic in a different way and we add respondent characteristics. The results in table 2 are largely robust to these specifications. Where we find significant differences, the mean squared error is higher, suggesting that our estimates correspond a global minimum of the sum of squares.

5.3 Simulations

To gain some insight into the meaning of the estimated parameters, we provide a number of simulated outcomes, based on the parameter estimates in Table 2. We present results for three countries that vary in growth rates over the previous decade: China (high growth), Italy (stagnation), Greece (negative growth). To make outcomes comparable, all macro variables and vignette dummies are set equal to the overall sample mean.

We then simulate three scenarios per country: (1) own income history and own income inequality;[9] (2) own income history and low income inequality (s^2=.21329 - Belarus in our sample); (3) own income history and high income inequality (s^2=2.212288 - Namibia in our sample). For each scenario, we consider evaluations by three individuals/households; with their income being at either the 20^{th}, 50^{th}, or 80^{th} percentile of their respective sample country income distributions. Thus we simulate 3 countries * 3 scenarios * 3 incomes/households * 3 vignette income levels ($\kappa = \{0.5, 1, 2\} * median$). The simulated numbers (representing the expected evaluation of life satisfaction) are presented in Table 3.

The table shows very clearly the effect of recent economic growth: according to the model, Chinese respondents evaluate vignettes higher than Italians or Greeks, because their frame of references is strongly influenced by their past incomes, which were substantially lower than their current incomes. As one would expect, the evaluations increase with κ. Note that respondents with higher incomes evaluate a given vignette lower than respondents with lower incomes (within columns, compare vignettes with the same value of κ, but different respondent incomes). The differences are not large, due to the small estimate of q_0. Income inequality within a country influences the sensitivity of evaluations to the value of κ. To provide one example, consider an Italian respondent with median income. In the scenario with low income dispersion, the ratings of the highest and lowest income vignettes differ by .340 (.624-.384); in the scenario with the high income dispersion, the difference is only .102 (.555-.453).

Figures A1 in the Appendix contain graphs of how the three households would evaluate different incomes. The vertical lines in the graphs show the three different income levels mentioned in the graphs (half median, median, twice median). Of particular interest are the intersections of the evaluation functions with the vertical lines, as these represent the simulated evaluations of the vignette incomes. Table 3 summarizes all simulated evaluations. In the graphs blue lines represent the households at the 20^{th} percentile of the respective country's income distribution. Red lines represent the households at the 50^{th} percentile of the respective country's income distribution. Green lines represent the households at the

[9] For own income history we use the country's history of GDP growth. For own income inequality we use the sample variance of reported incomes in that country.

80^{th} percentile of the respective country's income distribution. As expected, blue lines are higher than red, higher than green representing that wealthier households would rate the same-income vignette lower since their inclusive perceived income distribution is more to the right than poorer households. Further, we can see that the higher the inequality in a country the more dispersed the evaluations by different income-level households.

Table 3: Simulation results.

		China			Italy			Greece		
Income %	κ	Own	Low	High	Own	Low	High	Own	Low	High
20	0.5	0.538	0.555	0.535	0.412	0.384	0.460	0.401	0.380	0.454
20	1	0.602	0.635	0.583	0.504	0.498	0.514	0.491	0.480	0.508
20	2	0.633	0.644	0.616	0.600	0.625	0.567	0.596	0.618	0.562
50	0.5	0.532	0.553	0.524	0.410	0.384	0.453	0.399	0.380	0.449
50	1	0.597	0.635	0.572	0.500	0.495	0.504	0.487	0.478	0.498
50	2	0.631	0.644	0.608	0.596	0.624	0.555	0.592	0.616	0.550
80	0.5	0.525	0.550	0.514	0.408	0.383	0.448	0.398	0.380	0.443
80	1	0.591	0.634	0.561	0.495	0.492	0.495	0.483	0.475	0.489
80	2	0.628	0.644	0.599	0.592	0.544	0.588	0.588	0.615	0.539

The results represent the simulated evaluation of life satisfaction by a particular household in a particular country evaluating a particular vignette in three scenarios. For example a Chinese household at the 20th percentile of the Chinese income distribution simulated at the scenario with the Chinese GDP growth and Chinese income inequality would rate the life satisfaction of a vignette with income $0.5 * median$ *as 0.538. The column "Income %" shows the representative household at the % of the income distribution in the country; "*κ*" shows the vignette median income multiplier; "Own" shows the respective country with own income history and income inequality; "Low" shows the recpective country with own income history and low income inequality (*$s^2 = 0.21329$ *- corresponding to Belarus in our sample); "High" shows the respective country with own income history and high income inequality (*$s^2 = 2.212288$ *- corresponding to Namibia in our sample); One can get a sense of the effects of economic growth alone (holding inequality fixed) by comparing the "Low" and "High" columns across the three countries.*

6 Discussion

We construct and estimate a structural model of life satisfaction where an individuals' life satisfaction is a weighted sum of their satisfaction with K life domains. We assume that an individual's satisfaction with income is equal to her ranking in a perceived income distribution. We call this the individual's perceived inclusive income distribution and is characterized by two main features: 1. It is a weighted distribution of all incomes in one's reference group with weights representing the weight that the individual places on the incomes of others and herself. 2. It is a weighted average of all such perceived distributions over the individual's lifetime with a memory function representing the weight that the individual places on present versus past periods.

The unique feature of our approach is that we model how respondents rate the life satisfaction of others. These others are described in a number of anchoring vignettes that describe individuals who vary in a number of life domains as well as their income. In utilizing our data to estimate our model we were forced to make a number of simplifying assumptions,such as equal income growth for everyone in a country and reference groups comprising the whole population of a country. These are obviously oversimplifications. On the other hand, by using data on a large number of countries, we can plausibly assume that the reference groups of individuals are for the vast majority contained within their country of residence.

References

Barrington-Leigh, C. P., & Helliwell, J. F. (2008). Empathy and emulation: Life satisfaction and the urban geography of comparison groups (No. w14593). *National Bureau of Economic Research.*

Blanchflower, D. G., & Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of public economics*, 88(7-8), 1359-1386.

Brown, G. D., Gardner, J., Oswald, A. J., & Qian, J. (2008). Does wage rank affect employees wellbeing?. *Industrial Relations: A Journal of Economy and Society*, 47(3), 355-389.

Cantril, H. (1965). Pattern of human concerns.

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic literature*, 46(1), 95-144.

Clark, A. E., & Senik, C. (2014). Income comparisons in Chinese villages. *Happiness and Economic Growth: Lessons from Developing Countries*, 217-239.

Conti, R. M., Berndt, E. R., & Frank, R. G. (2006). Early retirement and public disability insurance applications: Exploring the impact of depression (No. w12237). *National Bureau of Economic Research.*

Dahlin, M., Kapteyn, A., & Tassot, C. (2014). Who are the Joneses?. *CESR-Schaeffer working paper*, (2014-004).

Deaton, A., & Stone, A. A. (2013). Two happiness puzzles. *American Economic Review*, 103(3), 591-97.

Ferrer-i-Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect. *Journal of public economics*, 89(5-6), 997-1019.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., & Falk, A. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *science*, 318(5854), 1305-1308.

Frank, R. H. (2012). The Easterlin paradox revisited. *Emotion*, 12(6), 1188.

Goerke, L., & Pannenberg, M. (2013). Direct evidence on income comparisons and subjective well-being.

Ifcher, J., Zarghamee, H., & Graham, C. (2016). Income inequality and well-being in the US: evidence of geographic-scale-and measure-dependence. *The Journal of Economic Inequality*, 1-20.

Kapteyn, A. (1977). A Theory of Preference Formation, Unpublished Ph.D. thesis, Leiden University.

Kapteyn, A., Smith, J. P., and Van Soest, A. (2007). Vignettes and Self-Report of Work Disability in the United States and the Netherlands. *American Economic Review*, 97(1):461473.

Kapteyn, A., Smith, J. P., and Van Soest, A. (2010). Life Satisfaction. In Diener, E., Kahneman, D., and Helliwell, J., editors, *International Differences in Well-Being*, chapter 4, pages 70104. Oxford University Press, Oxford, UK.

Kapteyn, A. and Teppa, F., (2003). Hypothetical intertemporal consumption choices. *The Economic Journal*, 113(486), pp.C140-C152.

Luttmer, E. F. (2005). Neighbors as negatives: Relative earnings and well-being. *The Quarterly journal of economics*, 120(3), 963-1002.

McBride, M. (2001). Relative-income effects on subjective well-being in the cross-section. *Journal of Economic Behavior Organization*, 45(3), 251-278.

Van de Stadt, H., Kapteyn, A., & Van de Geer, S. (1985). The relativity of utility: Evidence from panel data. *The review of Economics and Statistics*, 179-187.

Van Praag, B. M. (1971). The welfare function of income in Belgium: An empirical investigation. *European Economic Review*, 2(3), 337-369.

Van Praag, B. M., & Kapteyn, A. (1973). Further evidence on the individual welfare function of income: An empirical investigatiion in The Netherlands. *European Economic Review*, 4(1), 33-62.

Van Soest, A., Delaney, L., Harmon, C., Kapteyn, A., and Smith, J. P. (2011). Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 174(3):575595.

Van Vugt, M., & Tybur, J. M. (2015). The evolutionary foundations of status hierarchy. *Handbook of Evolutionary Psychology* (2), 788-809.

Appendix

Vignette descriptions

Vignette questions were asked after the question of own life satisfaction on the $0-10$ scale. Each individual was asked to answer either the A set or the B set of vignettes (about half in each country). Given the set that each individual answered, the order with which vignettes were presented was randomized.

Vignette set A:

A1 Think of a female who is 40 years old and happily married with a good family life. Her monthly family income is about (median income). She has severe back pain, which keeps her awake at night. On which step of the ladder do you think this person stands?

A2 Think of a male who is 50 years old and divorced. He has a daughter with whom he has a good relationship. He has a secure job that pays about (twice median income) per month. He has no serious health problems. On which step of the ladder do you think this person stands?

A3 Think of a male who is 25 years old and single without many friends. He makes about (half median income) per month. He feels he has little control over his job and worries about losing it. He has no health problems but feels stressed sometimes. On which step of the ladder do you think this person stands?

A4 Think of a female who is 35 years old and married, with no children. Her monthly family income is about (median income). Her work is a bit dull sometimes, but it is a very secure job. On which step of the ladder do you think this person stands?

A5 Think of a female who is a 70-year-old widow. She receives about (half median income) in income each month. She has many friends. Lately, she suffers from back pain, which makes housework painful. On which step of the ladder do you think this person stands?

A6 Think of a male who is 60 years old. His is single but has many friends his age. He no longer works but is comfortable with his decision to stop working. He receives about (twice median income) in income each month. He is very physically active. On which step of the ladder do you think this person stands?

Vignette set B:

B1 Think of a male who is 40 years old and happily married with a good family life. His monthly family income is about (twice median income). He likes to work but suffers from serious back pain, which keeps him awake at night. On which step of the ladder do you think this person stands?

B2 Think of a female who is a 65-year-old widow. She misses her husband a lot but has good relationships with her children and grandchildren. She receives about (half median income) in income each month. She has heart problems, which caused her to stop working. On which step of the ladder do you think this person stands?

B3 Think of a male who is 35 years old and married, with no children. His monthly family income is about (half median income). His work is a bit dull sometimes, but it is a very secure job. On which step of the ladder do you think this person stands?

B4 Think of a female who is 60 years old and divorced. She has children from her marriage but has little contact with them. She has an interesting job. Her monthly income is about (twice median income). She often has trouble sleeping. On which step of the ladder do you think this person stands?

B5 Think of a female who is 70 years old and married. She and her husband lead their own lives and dont do many things together. They have two children but rarely see them. Her monthly family income is about (median income). She is overweight and gets tired when walking for more than a few minutes. On which step of the ladder do you think this person stands?

B6 Think of a male who is 50 years old. He does not exercise and is obese. He has pain in his knees almost all the time. He is very secure in his job. He has been married for a long time, but he and his wife spend very little time together. His monthly family income is about (median income). On which step of the ladder do you think this person stands?

Coding of vignette dummy variables

Coding of vignette dummies used in Table 2.

Variable	Value	Vignette set
"healthgood"	1	{A2, A3, A6}
"healthbad"	1	{A1, A5, B1, B2, B4, B5, B6}
"familygood"	1	{A1, A3, A5, A6, B1, B2}
"familybad"	1	{B4, B5, B6}
"jobgood"	1	{A2, A6, B1, B4, B6}
"jobbad"	1	{A3, B2}

Coding of vignette dummies used in Table A3.

Variable	Value	Vignette set
"Vig-female"	1	{A1, A4, A5, B2, B4, B5}
"Vig-Backpain"	1	{A1, A5, B2}
"Vig-Securejob"	1	{A2, A4, B3}
"Vig-Interestingjob"	1	{A4}
"Vig-Married"	1	{A1, A4, A6, B3, B5, B6}
"Vig-Children"	1	{A2, B2, B4, B5}
"Vig-Widow"	1	{A5, B2}
"Vig-Working"	1	{A2, A3, A4, B1, B3, B4, B6}
"Vig-Healthprobs"	1	{A1, A5, B1, B2, B5, B6}

Evidence of non-systematic assignment of vignette characteristics

We perform two tests to examine whether vignette characteristics are randomly assigned to vignettes. First, we report pairwise correlations between the vignettes' income multipliers κ and the five other vignette characteristics. Secondly, we report a multiple regression of κ on the other characteristics (Table A1). We find that vignette characteristics are mostly as-good-as randomly assigned. The only characteristics that show some evidence of correlation with κ are age and job which is not surprising since it is expected that as one gets older, her job is better and her income increases.

Table A1: Parwise correlations and multiple regression of κ on other vignette characteristics.

	Female	Age	Health	Family	Job
Pairwise correlations	-0.27	0.73***	0.18	-0.16	0.83***
	(0.4)	(0.007)	(0.57)	(0.62)	(<0.001)
Multiple regression	0.34	0.0002	0.22	0.13	0.86*
	(0.34)	(0.99)	(0.27)	(0.46)	(0.06)

*Numbers in brackets indicate p-values. *p<0.1, **p<0.05, ***p<0.01.*

Robustness checks

As a first test of robustness of our estimation we change the coding of vignette characteristics and add respondent-level characteristics. More specifically we define the following variables capturing various vignette characteristics except income as follows: "$Vig - Backpain$" = 1 if the vignette description mentions the existence of back pain, = 0 otherwise; "$Vig - Securejob$" = 1 if the vignette description mentions that the hypothetical individual has a secure job, = 0 otherwise; "$Vig - Interestingjob$" = 1 if the vignette description mentions that the hypothetical individual has an interesting job, = 0 otherwise; "$Vig - Married$" = 1 if the vignette description mentions that the hypothetical individual is married, = 0 otherwise; "$Vig - Children$" = 1 if the vignette description mentions that the hypothetical individual has children, = 0 otherwise; "$Vig - Widow$" = 1 if the vignette description mentions that the hypothetical individual is a widow, = 0 otherwise; "$Vig - Working$" = 1 if the vignette description mentions that the hypothetical individual is currently working, = 0 otherwise; "$Vig - Healthprobs$" = 1 if the vignette description mentions that the hypothetical individual has any health problems, = 0 otherwise. We also include several variables capturing respondent characteristics, the meanings of which are self explanatory (We exclude some variables from Table A3 for the sake of space. None of the variables we excluded was significant at any level). The results are shown in column "A" which has the same starting values as our main specification in Table 2.

As a second test of robustness of our estimation, we present results of specification "A" with three different starting values of our parameters. These are shown in columns "B", "C" of table A3.

Table A3 shows that the main specification presented in Table 2 is robust to changing the variables that code the vignette characteristics and to adding additional respondent-specific characteristics. Our main specification is mostly robust to starting values. It seems, however that there exist two local optima for the objective function, each of which is reached depending on the starting values. However the local maximum identified in our main specification appears to be the global maximum when looking at the sum of residual deviations.

Table A3: Model estimation results.

	A	B	C
β_0	-0.105	-0.105	0.05
	(0.309)	(0.309)	(0.332)
β_1	0.303***	0.303***	0.341***
	(0.032)	(0.032)	(0.03)
q_0	0.053***	0.053***	0.054***
	(0.014)	(0.014)	(0.011)
a	0.971***	0.971***	0.215
	(0.041)	(0.041)	(1.15)
$\beta_{Vig-female}$	-0.059***	-0.059***	-0.038***
	(0.015)	(0.015)	(0.012)
$\beta_{Vig-age}$	0.002***	0.002***	0.0007**
	(0.001)	(0.001)	(0.0003)
$\beta_{Vig-Backpain}$	-0.027	-0.027	-0.0529***
	(0.017)	(0.017)	(0.016)
$\beta_{Vig-SecureJob}$	0.109***	0.109***	0.119***
	(0.013)	(0.013)	(0.013)
$\beta_{Vig-InterestingJob}$	0.096***	0.096***	0.046**
	(0.026)	(0.026)	(0.022)
$\beta_{Vig-Married}$	-0.114***	-0.114***	-0.112***
	(0.009)	(0.009)	(0.008)
$\beta_{Vig-Children}$	-0.074***	-0.074***	-0.097***
	(0.019)	(0.019)	(0.017)
$\beta_{Vig-Widow}$	-0.1***	-0.1***	-0.063***
	(0.016)	(0.016)	(0.015)
$\beta_{Vig-Working}$	-0.143***	-0.143***	-0.139***
	(0.01)	(0.01)	(0.01)
$\beta_{Vig-HealthProbs}$	-0.011*	-0.011*	-0.017***
	(0.007)	(0.007)	(0.006)
N	391,088	391,088	391,088
R^2	0.2101	0.2101	0.2074
Rt(MSE)	0.21068	0.21068	0.211
Res. Dev.	-108336	-108336	-107000
Starting values	{0, 1, 0.5, 0.8}	{0.6, 0.6, 0.6, 0.6}	{0.5, 0.5, 0.5, 0.5}

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table A3: Model estimation results (Cont.)

	A	B	C
β_{age}	-0.0001	-0.0001	-0.0001
	(0.006)	(0.006)	(0.0001)
β_{female}	0.006***	0.006***	0.005***
	(0.002)	(0.002)	(0.002)
$\beta_{married}$	-0.011***	-0.011***	-0.013***
	(0.003)	(0.003)	(0.003)
$\beta_{divorced}$	0.003	0.003	0.0005
	(0.006)	(0.006)	(0.006)
$\beta_{widowed}$	-0.006	-0.006	-0.01**
	(0.005)	(0.005)	(0.005)
$\beta_{domesticpartner}$	0.0006	0.0006	0.0023
	(0.007)	(0.007)	(0.007)
$\beta_{unemployed}$	0.0006	0.0006	0.004
	(0.004)	(0.004)	(0.004)
$\beta_{selfemployed}$	0.003	0.003	0.003
	(0.005)	(0.005)	(0.004)
β_{urban}	-0.008*	-0.008*	-0.008*
	(0.005)	(0.005)	(0.004)
$\beta_{ln(GDPperCapita)}$	0.035*	0.035*	0.017
	(0.019)	(0.019)	(0.018)
$\beta_{ln(EducationExpenditure)}$	0.002	0.002	0.019*
	(0.011)	(0.011)	(0.01)
$\beta_{ln(LifeExpectancy)}$	0.062	0.062	0.041
	(0.87)	(0.87)	(0.09)
N	391,088	391,088	391,088
R^2	0.2101	0.2101	0.2074
Rt(MSE)	0.21068	0.21068	0.211
Res. Dev.	-108336	-108336	-107000
Starting values	{0, 1, 0.5, 0.8}	{0.6, 0.6, 0.6, 0.6}	{0.5, 0.5, 0.5, 0.5}

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Note the following:

- Estimated parameters in columns A and B are essentially the same, apart from some rounding differences.
- All estimation results where starting values are the same and above 0.6 produce the exact same results with column B.
- All estimation results where starting values are the same and below 0.5 produce the exact same results with column C.

List of countries, sample sizes, interview years and non-representative samples

Table A4: Countries, sample size, interview year, non-rep samples

Country	2011	2012	2013	2014	Non-rep sample
Afghanistan			1,000		
Albania				999	
Argentina	1,000				
Armenia		1,000			
Australia			1,002		
Austria				1,000	
Azerbaijan			1,000		✓
Bahrain			1,002		✓
Bangladesh		1,000			
Belarus		1,052			
Benin	1,000				
Bolivia	1,000				
Bosnia & Hertzekovina				1,001	
Botswana				1,000	
Brazil	1,042				
Bulgaria			1,000		
Cambodia		1,000			
Cameroon				1,000	
Canada		1,002			✓
Central African Rep.	1,000				✓
Chad				1,000	✓
Chile	1,009				
China		4,256			✓
Colombia	1,000				
Congo Kinasha		1,000			✓
Costa Rica			1,000		
Croatia				1,000	
Czech Republic			1,001		
Dominican Republic	1,000				
Ecuador	1,003				
Egypt	1,077				
El Salvador		1,000			

Table A4: Countries, sample size, interview year, non-rep samples (Cont.)

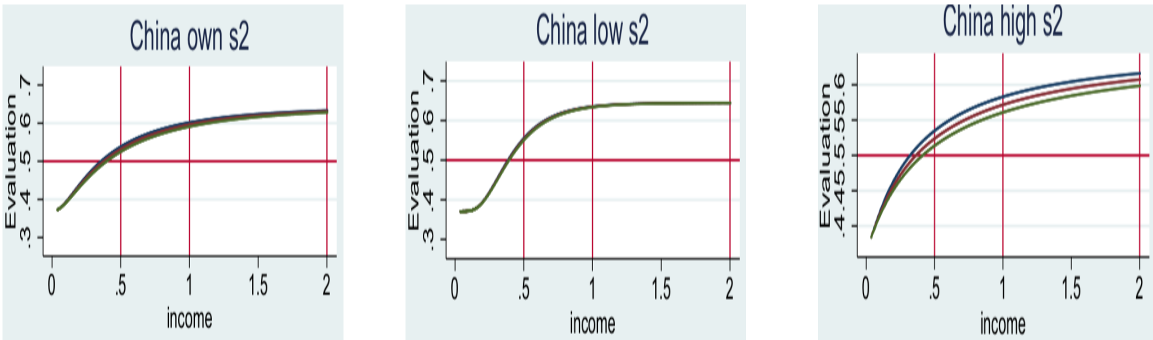
Country	2011	2012	2013	2014	Non-rep sample
Ethiopia		1,000			
France		1,003			
Georgia			1,000		✓
Germany		3,033			
Ghana				1,000	
Greece			1,003		
Guatemala			1,000		
Haiti				504	
Honduras		1,000			
Hungary			1,019		
India		5,000			✓
Indonesia		1,000			
Iran			1,000		
Israel	1,000				
Italy		1,004			
Japan	1,000				✓
Jordan			1,000		
Kazakhstan			1,000		
Kenya			1,000		
Kuwait			1,008		✓
Kyrgyzstan		1,000			
Laos		1,000			✓
Lebanon		1,012			✓
Liberia				1,000	
Macedonia		1,025			
Madagascar				1,008	✓
Malawi	1,000				
Malaysia			1,000		
Mauritania				1,000	
Mexico		1,000			
Moldova				1,000	✓
Mongolia				1,000	
Morocco			1,007		✓
Myanmar			1,020		✓
Namibia				1,000	
Nepal		1,000			
New Zealand		1,008			
Nicaragua			1,000		
Nigeria		1,000			
Pakistan		1,008			✓
Palestine				1,000	

Table A4: Countries, sample size, interview year, non-rep samples (Cont.)

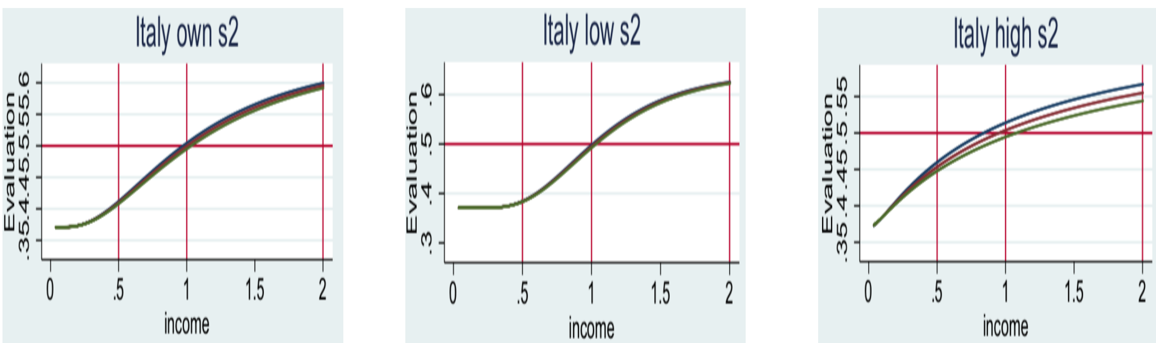
Country	2011	2012	2013	2014	Non-rep sample
Panama		1,001			
Paraguay	1,000				
Peru	1,000				
Philippines		1,000			✓
Poland			1,000		
Portugal				1,007	
Russia		1,500			
Rwanda		1,000			
Saudi Arabia				1,017	✓
Senegal				1,000	
Singapore				1,000	
Slovakia				1,000	
Slovenia				1,017	
South Africa		1,000			
South Korea		1,000			
Spain		1,001			
Sri Lanka			1,030		
Sudan				1,000	✓
Syria		1,025			✓
Taiwan		1,000			
Tajikistan				1,000	
Tanzania			1,008		
Thailand		1,000			
Turkey		1,000			
Uganda			1,000		
United Arab Emirates	1,012				✓
United Kingdom		3,075			
United States		1,019			
Uruguay	1,000				
Uzbekistan		1,000			
Venezuela	1,000				
Vietnam		1,000			
Yemen				1,000	
Zambia				1,000	
Zimbabwe			1,000		
Total (107 countries)	18,143	51,024	26,103	25,553	23 countries

Simulation graphs A1

China.



Italy.



Greece.

Blue lines represent the households at the 20th percentile of the respective country's income distribution. Red lines represent the households at the 50th percentile of the respective country's income distribution. Green lines represent the households at the 80th percentile of the respective country's income distribution. As expected, blue lines are higher than red, higher than green representing that wealthier households would rate the same-income vignette lower since their inclusive perceived income distribution is more to the right than poorer households. Further, we can see that the higher the inequality in a country the more dispersed the evaluations by different income-level households.