Collegio Carlo Alberto

The Relative Utility Hypothesis With and Without Self-reported Reference Wages

Adrián de la Garza
Giovanni Mastrobuoni
Atsushi Sannabe
Katsunori Yamada

No. 159
November 2010

Carlo Alberto Notebooks

www.carloalberto.org/working_papers

The Relative Utility Hypothesis With and Without Self-reported Reference Wages*

Adrián de la Garza † Giovanni Mastrobuoni ‡ Atsushi Sannabe §
Katsunori Yamada ¶

Abstract

This article uses survey data of 90,000 union employees working in 62 publicly-traded companies in Japan between 1990 and 2004 to study the effect of both own and *self-reported* reference wages on workers' subjective well-being levels. The availability of *self-reported* reference wages generates very robust findings that do not depend on questionable identifying assumptions. These findings confirm that higher estimates by workers of their peers' earnings are associated with lower levels of life and job satisfaction. These comparison effects are statistically and economically strong but are smaller in absolute value than the impact of workers' own wages on their own utility. We compare our results with standard tests of the relative utility hypothesis in the literature that recur to alternative proxies for comparison wages, including: (i) Mincer-predicted wages; (ii) cell averages defined over different groups within our dataset; (iii) cell wage averages estimated from an external data source; and (iv) colleagues' average wages. In spite of their potential flaws—that we discuss—these alternative empirical constructs employed in the literature do not introduce a simple classical measurement error problem and the bias attributed to this measurement error issue can go in both directions. We propose a simple IV strategy when the self-reported reference wage is not available that does not eliminate the bias but delivers a lower bound of the "true" effect. We also address the issue of endogeneity of self-reported reference wages in our subjective well-being regressions by accounting for workers' pessimistic attitudes at the workplace.
Keywords: Subjective well-being; relative utility; reference wages.
JEL classification: D00; J28.

---

*We give many thanks to Gani Aldashev, Prashant Bharadwaj, Michael Boozer, David Card, Andrew Clark, Angus Deaton, Hank Farber, Erzo F. P. Luttmer, Guy Mayraz, Takashi Unayama and seminar participants at Yale University and Hitotsubashi University for their useful comments. Of course, any remaining errors and omissions are our own. Yamada is grateful for the research grant provided by the GCOE program entitled "Human Behavior and Socioeconomic Dynamics" of Osaka University.

†Banco de México, Research Department, 5 de Mayo 18 Col. Centro, Mexico DF 06500, Mexico.

‡*Corresponding author (giovanni.mastrobuoni@carloalberto.org):* Collegio Carlo Alberto, Via Real Collegio 30, Moncalieri (TO) 10024, Italy.

§Oxford University, Japan Society for the Promotion of Science and St. Antony's College, 62 Woodstock Road, Oxford, UK.

¶Institute of Social and Economic Research, Osaka University, Mihogaoka 6-1, Ibaraki 567-0047, Japan.

1 Introduction

The relative utility hypothesis (Duesenberry, 1949) posits that individuals derive utility not only from their own levels of consumption but also from how much they consume relative to a well-defined benchmark group. In the last decade, numerous studies have found empirical support for this theory, documenting the existence of an inverse relationship between an individual's reported level of happiness and the income or wages earned by his peers.[1] However, there is still considerable disagreement regarding the magnitude of these relative effects. In particular, some studies have argued that the impact of comparison income on individuals' well-being is both statistically and economically stronger than that of absolute income—a result that goes against standard neoclassical theories of consumption.

One difficulty in evaluating and comparing the findings in the literature is the different definitions of relative income. One approach is to estimate a Mincer equation that allows the econometrician to predict the wage of an individual with given characteristics. This predicted wage then becomes the relative measure to which individuals compare. An alternative method is to calculate average wages by cells or groups defined by age, gender, education level, and other particular individual characteristics. In general, all of these approaches implicitly assume that individuals compare themselves to a hypothetical "average" individual with given characteristics and that they infer their peers' wages the way econometricians do (Manski, 1993; Sloane and Williams, 2000). More specifically, each method has potential drawbacks. For example, studies that recur to the Mincer solution must identify the reference wage effect on happiness by including additional variables in the wage regression that are excluded from the happiness equation. These exclusion restrictions are frequently unjustified and easily refutable.

This paper empirically tests the relative utility hypothesis and assesses the validity of the different methodologies that the literature has employed to construct comparison income measures. We study the relationship between subjective well-being and both absolute and relative wages by using data on about 90,000 Japanese workers surveyed in repeated annual cross-sections between 1990 and 2004. Our analysis overcomes some of the criticisms that previous studies have faced, as we employ data on workers' actual perceptions of their peers' wages instead of theoretical constructs that may not correspond to the "true" benchmark that individuals use to compare themselves to their reference group. Our results show that individuals report higher levels of both job and life satisfaction when their individual absolute income levels are higher. We also demonstrate the existence of significant relative income effects, as workers experience greater satisfaction when they perceive that their own wages are higher relative to their peers'. The impact of relative income is consistently smaller in absolute value than the effect of absolute income on reported well-being. However, we demonstrate that the absolute and, in particular, the relative effects are much stronger for workers who are better able to accurately predict their peers' wages—a result that we associate with feelings of jealousy and with workers' access to information about job offers and the wage structure in their profession. Our analysis also addresses the possibility of endogenous self-reported reference wages due

[1] For an excellent exposition of recent developments in the economics literature and the validity and applicability of happiness measures in economics, see Kahneman and Krueger (2006).

to underlying pessimistic attitudes that may be correlated with individuals' subjective well-being measures.

Having confirmed the robustness of our results using self-reported reference wages, we then re-estimate the model using instead the theoretical comparison income measures that the literature has employed. Our findings show that the estimated effects of reference wages on well-being are not consistent when using constructed reference wages as proxies for the comparison wage that is actually perceived by the worker. We estimate the bias in these estimates by means of a generalized version of the classical measurement error model and show that the direction of the bias is ambiguous. We propose a simple instrumental-variables strategy that can be implemented when self-reported comparison wages are not available. This approach does not eliminate the measurement error bias, although it delivers a lower bound—in absolute value—of the "true" impact of reference wages on well-being. Our analysis also suggests that the use of Mincer-predicted reference wages—one of the traditional comparison wage measures employed in previous literature—must be avoided whenever possible, as this method tends to yield highly unstable results mainly due to its reliance on questionable identification assumptions.

Additionally, our paper also documents the relationship between subjective well-being and individual characteristics of workers in Japan. Our analysis compares these results with those in previous studies, which have focused primarily on the U.S. and the western European labor forces. Our findings confirm several standard results in the happiness literature. For instance, women and married workers tend to report higher levels of satisfaction than men and single individuals, respectively. However, we also observe a U-shaped relationship between satisfaction and educational attainment, which contrasts starkly with the monotonically increasing association in the U.S. and Europe found in past studies. Given that, to the best of our knowledge, our paper is one of the first happiness studies in Japan, our findings are useful in the construction of a set of stylized facts that may be generalizable to cultures outside of the western hemisphere.

The paper is organized as follows. Section 2 provides a brief synthesis of the literature and places this study in its historical context. In Section 3, we present our empirical framework to test the relative utility hypothesis, introduce our data, and explain how they differ from previous data sets on job satisfaction and happiness. Section 4 presents our main results focusing on the use of workers' expectations on their peers' wages as our primary comparison income measure. In Section 5, we re-estimate the model employing alternative metrics used in the literature and compare the results. We also demonstrate from a theoretical standpoint why the bias present in estimations that employ these theoretical constructs as reference income proxies may go in either direction and propose our instrumental variables approach to bound the reference wage estimate. Finally, Section 6 presents some concluding remarks.

2 Literature Review

In a seminal paper, Easterlin (1974) analyzed the results of thirty individual-level surveys implemented in nineteen different countries between 1946 and 1970, and observed that there is a noticeable positive association between income and happiness *within* countries. That is, in each country survey, individuals in the highest income brackets tend to report, on average, significantly greater levels of happiness relative to those in lower income categories. However, this association is not visible *among* countries: at any given point in time, the differences in happiness between richer and poorer nations are negligible.[2] Similarly, although income per capita rose steadily in the United States during this period, average reported happiness showed no increasing trend and even declined between 1960 and 1970. Easterlin (1974) explained this paradox by alluding to Duesenberry's (1949) relative utility theory, which suggests that people often compare themselves to a reference group, and so they care not only about their own absolute consumption levels, but about how much they consume relative to that benchmark. Thus, Easterlin (1974) concluded that a higher average income level in a richer country or within a country over time increases overall aspirations of all individuals, negating the expected positive impact on welfare.

These results prompted the economics literature to investigate further the impact of absolute and relative income on subjective well-being measures.[3] Van de Stadt et al. (1985) investigate this from a theoretical perspective by means of a dynamic model of habit formation and utility interdependence. They estimate the model using two waves of an annual panel in the Netherlands and conclude that their findings are compatible with the hypothesis that utility is completely relative. However, their estimates cannot exclude the possibility that utility is partly relative and partly absolute.[4] From an empirical perspective, this discussion was later revived following an article by Clark and Oswald (1996). Using job satisfaction data from the British Household Panel Survey, the authors find a strong correlation between job satisfaction and different measures of comparison income. More than that, they also observe that workers' reported levels of well-being are at best weakly correlated with absolute income, which supports the findings by Easterlin (1974) twenty years earlier.

In the subsequent empirical literature, the evidence on these issues has been mixed. In agreement with Clark and Oswald (1996), the majority of the studies find a negative relationship between satisfaction and various comparison income variables, while observing an own-income effect that is either relatively smaller in absolute value or statistically insignificant (Sloane and Williams 2000; McBride 2001; Ferrer-i Carbonell 2005; Brown et al. 2008). Using a panel of Russian individuals between 1994 and 2000, Senik (2004) also shows that the effect of individual income is smaller in magnitude than that of reference income, although she observes that reference income exerts a positive influence on satisfaction. She argued that these findings are consistent with Hirschman and Rothschild's (1973) "tunnel effect" conjecture, whereby rising inequality may increase welfare if it is interpreted as a positive signal with respect to likely future outcomes. Blanchflower and Oswald (2004) used samples of British and U.S. workers to

---

[2] For further relative utility tests using national average data, see Easterlin (1995) and Easterlin (2001).

[3] Discussions on these topics have been widespread in the psychology literature. For instance, see Veenhoven (1991), Diener et al. (1993), and Boyce et al. (2010).

[4] Their model is based on a previous paper by Van Praag (1968) that demonstrates theoretically that every individual can evaluate his welfare position with respect to his income level on a bounded scale, and that a description of this evaluation can be given by a log-normal individual income welfare function. More recently, Rayo and Becker (2007b) combined a happiness function with habit formation and peer comparisons to argue that utility levels depend not only on our current consumption but also on our personal histories and social environment.

show a strong positive correlation between personal income and satisfaction, concluding that "money buys happiness." They also found that relative income had some explanatory power in a happiness equation, although this was not enough to explain away their findings in support of the Easterlin hypothesis. Perhaps implicitly shedding some light on the reasons behind these disparate findings, Luttmer (2005) emphasized the relevance of the definition of the reference group. He finds evidence of an inverse relationship between individual well-being and other people's earnings, but this relationship is much stronger for people who socialize more with neighbors than for people who interact more with friends outside the neighborhood.

One reason that could potentially explain away such discrepancies in the various findings in the literature is the difficulty in obtaining an appropriate reference wage measure. First, it is not obvious who belongs in the comparison group.[5] Do workers compare themselves to other workers in the same company? In the same industry? To their relatives and friends? To their neighbors? To people of their same age and education level? To workers of the opposite sex? The possibilities are vast. A second and perhaps more subtle issue is that, ideally, the econometrician would like to have information on workers' perceptions about their peers' wages but such self-reported beliefs are typically unobserved.[6]

As Clark et al. (2008) have pointed out, the bulk of the literature has defined relative income as income earned by people with similar individual characteristics—such as age, education, or civil status—and confined to a common social sphere—e.g., same company, same neighborhood, or doing the same kind of job. These comparison income measures are generally constructed in a number of ways, which include: (i) predicting individual wages from Mincer equations (Clark and Oswald 1996; Sloane and Williams 2000; Lévy-Garboua and Montmarquette 2004; Senik 2004); (ii) calculating cell averages according to specific individual characteristics within-sample (Ferrer-i Carbonell 2005) or (iii) out-of-sample from external data (Cappelli and Sherer 1988; McBride 2001; Luttmer 2005; Clark et al. 2009a; Card et al. 2010); and (iv) computing average wages of workers' colleagues (Rizzo and Zeckhauser 2003; Brown et al. 2008; Clark et al. 2009b).

These approaches present several potential problems. For instance, studies that estimate comparison income by imputing peers' wages from Mincer equations implicitly assume that individuals will infer peers' wages the way econometricians do (Manski 1993; Sloane and Williams 2000). Moreover, this methodology relies heavily on questionable identification assumptions. In particular, Mincer regressions from which the predicted reference income measures are derived must include variables that are excluded from the satisfaction equation, an exclusion restriction that is often not warranted.[7] More generally, these methods presuppose that individuals compare themselves to a hypothetical average

[5]For example, Falk and Knell (2004) argue that reference groups used as benchmarks may be endogenous. Also, using data on workers' perceptions about their competition, Clark and Senik (2010) show that workers tend to compare more often to their own colleagues, while comparisons to their friends are less common. From a theoretical standpoint, Rayo and Becker (2007a) model happiness as a function with a time-varying reference point, and show how both habits and peer comparisons arise as special cases of this process.

[6]Recent studies by Knight et al. (2009) and Senik (2009) use data on individuals' perceptions about their rivals' income. However, these relative income measures are ordinal, which are generally not appropriate in relative utility tests.

[7]Table A-1 shows a comparison of the results obtained by a selection of papers in the literature employing various reference income measures.

worker within well-defined categories defined by the econometrician, an assumption that is difficult to validate.[8]

To better understand the drawbacks in the literature, the next section discusses in more detail the empirical framework that we will follow to test the relative utility hypothesis.

3 Empirical Framework and Data Description

In order to examine the impact of absolute and relative wages on subjective well-being, researchers typically estimate the following regression:

$$\text{SWB}_i = \alpha_1 y_i + \beta_1 \bar{y}_i + \mathbf{x_i}'\gamma + u_i, \quad (1)$$

where the dependent variable SWB_i represents a measure of worker i's subjective well-being, such as happiness or job satisfaction. This proxy for utility is assumed to depend on the worker's (log) wage, y_i;[9] a (log) reference wage that workers use as a benchmark to determine how well they do relative to their peers, $\bar{y}_i$; and a vector of individual characteristics, $\mathbf{x_i}$, that includes age, tenure, gender, educational attainment, marital status, among other socio-demographic factors. As usual, the term u_i corresponds to an idiosyncratic error term.[10]

One advantage that we have over previous studies is that our dataset contains information on workers' actual perceptions about their peers' wages. Assuming that these self-reported reference wages correspond to the true benchmark that workers use to compare themselves against their peers, we can estimate equation 1 by standard methods[11]

[8] Relative utility tests that are based on "hypothetical choice experiments" may overcome these issues since rivals and reference income levels are well-defined for survey participants in such experiments. See, for example, Solnick and Hemenway (1998); Johansson-Stenman et al. (2002); Alpizar et al. (2005); Carlsson et al. (2009). The general conclusion in these studies show that relative income effects are about as important as absolute income effects. See also Yamada and Sato (2010) for hypothetical choice experiments evidence from Japan suggesting that, on average, utility is only partly relative.

[9] Unless otherwise noted, all wages are in logs.

[10] This specification can be thought of as a reduced-form version of a standard utility function of the form:

$$U = U(c, \bar{c}, h),$$

where c is individual consumption, $\bar{c}$ is the level of consumption of a comparison group, and h is hours worked. The theoretical literature has investigated how much keeping-up-with-the-Joneses ultimately matters to consumers and what role social status plays in determining individual utility levels. These studies find that, in addition to their own levels of consumption, individuals care about their peers' consumption levels and their wealth rank relative to their comparison group, which validates the use of reduced-form models such as equation 1. For instance, see Cole et al. (1992); Corneo and Jeanne (1997); and Yamada (2008). Boskin and Sheshinski (1978) and Oswald (1983) theoretically analyze tax policy implications of relative utility, while Abel (1990); Bakshi and Chen (1996); Gali (1994); and Campbell and Cochrane (1999) examined such relative effects on asset pricing. Frank (1985) and Frank (2005) show that relative concerns are more important for positional goods consumption than other goods such as leisure, and that the structure of utility functions with relative concerns over different types of goods is of key importance in the valuations of social welfare.

[11] Given the discrete, hierarchical nature of the dependent variable, the vast majority of happiness studies in economics have estimated equations similar to 1 using ordered logit or probit estimators. However, it is not clear that these methods are superior to ordinary least-squares since additional assumptions required for these estimates to be valid—such as that of parallel slopes—are often not met.

and obtain an unbiased measurement of the effect of relative wages on subjective well-being.[12] We present these findings in Section 4. Our data also allow us to compare our results to the ones we would have obtained had we followed the alternative methodologies that Manski (1993), Sloane and Williams (2000), and others have deemed potentially flawed. These findings are reported in Section 5.

3.1 Data Description

Our dataset comes from the Comprehensive Survey of Labor Union Members, which was designed and administered by a group of psychologists at the International Economy and Work Research Institute. It comprises repeated cross-sections on about 90,000 union members working in Japanese firms listed on the Tokyo Stock Exchange (TSE) from 1990 to 2004. The survey requests that respondents provide self-assessments on their individual well-being at work and in life in general. In addition to this, other questions attempt to obtain information on workers' perceptions of their work environment.[13] The dataset also allows us to control for individual demographic and socioeconomic characteristics, which include age, gender, educational attainment, marital status, annual wage level, overtime hours worked, and workers' expectations of their peers' wages.

After cleaning the data and removing some inconsistencies, we are left with 78,136 observations.[14] Table 1 shows some statistics that describe our dataset. Workers in our sample are young with an average age of 35 years. Their average tenure is 14 years, which suggests relatively low mobility in the Japanese labor force. Moreover, union workers in Japan seem to be well-educated in general, as almost all of them have graduated from high school, close to 50 percent have some college experience, and 36 percent have completed at least a university-level degree. We also observe that 58 percent of workers are married and work an average of 23 hours of overtime per month. All of these individuals are regular full-time employees and union members. About one-third of them holds blue-collar positions and close to one-fifth of them performs some managerial role in the company. These numbers do not intend to provide an accurate depiction of the representative Japanese worker since, as described above, the survey was administered exclusively to union members employed by major publicly-traded companies. Thus, the results below may not be generalizable to the entire Japanese labor force—for instance, our dataset does not cover employees in high administrative positions as they are not allowed to take part

Further, Luttmer (2005) showed that results from happiness equations obtained by OLS are virtually indistinguishable from those obtained by ordered models. Our OLS estimates in Section 4 facilitate the interpretation of the coefficients, although ordered probit estimates are also reported in the robustness section for completeness.

[12] This of course requires that the typical exogeneity assumptions hold. Our robustness tests in Section 4 address the possibility that wages are endogenous.

[13] The full list of question categories is available in an earlier version of this paper; see Appendix 1 in de la Garza et al. (2008)

[14] The reduction in the number of observations is mainly due to missing information for some of the variables used in the empirical analysis. Observations that showed inconsistencies in the data, such as the worker's tenure being greater than his age, were also dropped from the sample. To ensure that these problems were not due to sample selection on observable characteristics, we compared the full-sample and the working databases along different dimensions including age, education, wages, and reported levels of happiness and job satisfaction. We are *happy* to report that we were unable to find any statistically significant differences between the two datasets.

in unions. Nonetheless, due to the large size of our sample and the breadth of coverage of 62 firms across a variety of industries from food to electronics to finance, we believe that this dataset does capture significant features of the Japanese labor market.

One additional consideration is that our database underrepresents women in the Japanese labor force. According to the World Bank's World Development Indicators, women represent about 40 percent of the country's working population; in contrast, the share of female participation in our working sample is only 22 percent. This is an important issue because male and female workers differ significantly along several dimensions. Table 1 shows some of these gender differences in terms of observable characteristics, such as education and marital status. For instance, while 40 percent of the male subsample obtained a college degree, only 20 percent of the female group achieved this goal. Interestingly, the fraction of married men is twice the proportion of married women (66 percent vs. 33 percent). Since the literature has found significant correlations between various subjective well-being measures and individual characteristics such as gender, educational attainment, and marital status, pooling the male and female subsamples in our analysis may lead to results that would differ had we considered these two groups separately. We thus keep this distinction in mind and discuss these differences accordingly.

3.1.1 Subjective-Well Being

Each respondent provides information on his own subjective well-being by choosing one of five possible categories, from "least satisfied" (category 1) to "most satisfied" (category 5). We employ primarily two distinct subjective well-being measures, life happiness and job satisfaction, which present some significant differences.[15] First, the correlation between the two is only 27 percent. As Figure 1 suggests, the distribution of life happiness is more spread out and more skewed to the left than that of job satisfaction. Since the literature has utilized a wide variety of subjective well-being measures, the availability of these two variables allows us to test the robustness of our results. For brevity, the empirical analysis below highlights our findings utilizing life happiness as the dependent variable. However, our results are robust to the use of job satisfaction as an alternative proxy for subjective well-being.

3.1.2 Workers' Own Wages

The survey also requests that workers mark down their own wage level from a list of 9 categories, where category 1 denotes annual wages of under 2 million yen and category 9 corresponds to an annual income level of over 10 million yen. We measure individual wages as the mid-point in each of the 7 intermediate categories, and use *ad hoc* values for the two extreme categories. Thus, respondents who reported categories 1, 2, ..., 9 as their wage level, were assigned annual wages of 1.5; 2.5; ...; 12 million yen, respectively. Alternative choices do not alter the main results. Additionally, we deflate this nominal measure using the Consumer Price Index to obtain real wages with 1990 as the base year.

[15]At the beginning of the questionnaire workers are first told "from now on, we would like to ask about your general happiness," and are asked to report whether they agree with the statement "I'm very happy!" in general. Later in the questionnaire, workers are asked to report about their satisfaction with respect to all aspects of their job.

One salient feature of wages in our sample of union workers in Japan is that individual characteristics explain an atypically high fraction of the variation in wages. Table 2 shows standard Mincer regressions for men and women, separately. Wage regressions for the male subsample that control for age, tenure, education, hours worked, and marital status, obtain R^2 measures of about 70 percent.[16] In contrast, similar specifications for a sample of U.S. workers would generally yield an R^2 of about 30-35 percent. This high coefficient of determination in the Japanese data may be partly due to the very strict seniority system that prevails in the Japanese labor market. For instance, the correlation between wages and age for the full sample of Japanese workers is 69 percent. Using a comparable sample of unionized workers from the Current Population Survey during the same 1990-2004 period, the corresponding correlation measure for the U.S. is only 21 percent. This high coefficient of determination in the Japanese sample may play to our advantage as it reduces possible concerns that wages may depend themselves on happiness, thus minimizing potential endogeneity issues. We nevertheless address this possibility in Section 4.

3.1.3 Self-Reported vs. Constructed Reference Wages

Survey respondents provide information on self-reported comparison wages by answering the question, "What do you think is the average wage of corporate employees who are the same age as you and doing the same job?" Just as in the case of workers' own wages, answers to this question are originally chosen from a list of 9 categories, then matched with individual wage values according to the category mid-points described above. In contrast with the reference wage measures employed in previous studies, the availability of self-reported comparison wages allows us to gauge how workers perceive their peers' wages. We refer to this self-reported measure as the worker's "true" reference wage to distinguish it from the alternative empirical constructs that the literature has traditionally utilized.

To better understand the differences between self-reported and constructed wage measures, we compare our relative wage measure as reported by the worker to some of the most common reference wage estimates used in other studies. A standard method consists in calculating wage averages by cells or groups defined by a set of given observable characteristics of workers. We define cells in three different ways: by gender and age; by gender, age, and education level; and by gender, age, education level, and managerial experience. Following the literature, we estimate these cell averages from two different underlying data sources: our own dataset, described above, and an external dataset containing wage information of workers with similar observable characteristics. This second approach is most commonly employed in the literature and is useful to validate our results. The external data correspond to wages for individuals working in large companies with over 1,000 employees, as provided by the Basic Survey on Wage Structure (BSWS) released by the Japanese Ministry of Health, Labour and Welfare.[17] A third method that is frequently used in relative utility studies is to estimate Mincer regressions such as the ones showed in

[16] This feature depends only marginally on wages being reported in categories. When we smooth wages adding a disturbance term that is uniformly distributed between each cutoff point the R^2 statistic decreases by, at most, 5 percentage points.

[17] The data can be obtained from the following website: http://www.mhlw.go.jp/english/database/db-l/.

Table 2. The theoretical comparison wage measure corresponds to the fitted wage value of the average worker. In what follows we use the prediction from a wage regression that controls for gender, age, education level, and managerial experience—similar to the third cell defined above. When using Mincer regressions to estimate reference wages researchers need to find appropriate exclusion restrictions, meaning at least a variable that influences reference wages but not happiness. Notice that computing cell averages is equivalent to running a Mincer wage regression where the regressors are fully interacted with each other. In this case the identification hinges on these interactions, which, like the exclusion restrictions, are assumed to influence reference wages but not happiness.

The bottom panel of Table 1 displays summary statistics for workers' own wages and self-reported reference wages, as well as comparison wages defined according to the various methods employed in the literature. The table shows that average wages estimated from the BSWS external database are less than 2 percent higher than self-reported reference wages but about 10 percent higher than workers' own wages. We attribute this significant difference in the second case to the inclusion of workers with a *general* assistant manager position (*kacho* in Japanese) in the BSWS data. In contrast, our dataset only includes wage information for employees in non-supervisory, assistant manager positions. This technical subtlety evinces how difficult it can be to compare wage averages across different datasets, which may pose challenges to the use of external databases in the construction of reference wage measures. In addition, self-reported reference wages seem to be significantly more disperse than any of the other measures utilized in the literature. Averaging wages across workers with similar given characteristics may not reflect accurately, for example, how undervalued a worker feels in the labor market. This is further confirmed in Figure 2, where we plot the kernel density of various wage measures. Comparing self-reported reference wages and reference wages constructed from an external data source, for example, illustrates a potential underlying pessimism among Japanese workers regarding their beliefs of what their peers earn. If feelings like this are common among workers, studies utilizing such theoretical constructs would underestimate the effects of comparison wages on workers' self-reported levels of well-being.

The data also hint at a potential underlying pessimism in workers' perceptions of their peers' wages, as suggested by the 10-percentage point difference between self-reported reference wages and workers' own wages. This is an important point since a worker's belief of his peers' wages may be endogenous, for example, if a relatively happier individual has a more optimistic view of his own salary relative to his colleagues'. Conversely, a pessimistic worker may think that he is underpaid with respect to his peers. With panel data, assuming there was enough variation over time in happiness, wages and relative wages, one could difference out time-invariant worker characteristics, such as pessimism, and obtain unbiased estimates of the effect of absolute and relative wages on happiness. As our dataset does not allow us to track individuals over time, we address this problem by using answers to two questions in our survey that are likely to capture workers' pessimistic attitudes along two different dimensions. The first question asks whether a worker believes that his colleagues would help him in times of need; the second question asks the worker if he is satisfied about the possibility of promotion within his company. Controlling for these two effects should capture the level of inherent pessimism of a given worker, which should in turn quell some of the above-mentioned endogeneity concerns.

Having introduced the data, the next section presents the main results. It is important to remark that, in spite of their virtues, self-reported reference wage measures are not a panacea and continue to leave some issues unresolved. For instance, by asking workers to report what they believe that other employees with their same characteristics *and in the same job* earn, the survey restricts workers to confine their reference group to individuals within their same job. Still, on a more regular basis, workers may potentially compare themselves to relatives, friends, or colleagues doing other jobs. Moreover, the relevant reference group may not be stable over time. For instance, a recent graduate may compare himself to other recent graduates; but a person who has been out of school for 10 years may compare himself to former classmates and colleagues, to supervisees and supervisors, and even to former selves.[18] Nonetheless, we believe that self-reported comparison wage data are a superior alternative to reference wage measures constructed from Mincer equations or cell averages, since self-reported refrence wages do not limit the worker's reference group to some cell average, do not assume that individuals compute their peers' wages the way econometricians do, and do not hinge on disputable identifying restrictions (see Section 5).

4 Results

To empirically test the relative utility hypothesis, we estimate a version of equation 1 that substitutes our self-reported reference wage measure, $\bar{y}_i^*$, for the standard proxy used in the literature, $\bar{y}_i$.[19] As discussed in Section 3.1, there are important differences between male and female workers in almost every dimension from wages to hours worked to educational attainment, which justifies our estimation of equation 1 for men and women separately. Pooling the male and female subsamples does not have any significant qualitative impact on our findings. The errors u_i are allowed to be correlated within firms although different clusterings do not affect our conclusions.

The results in Table 3 find strong empirical support in favor of the relative utility hypothesis: holding wages constant, individuals tend to report lower levels of satisfaction when they perceive that their peers' wages are higher. In particular, if a worker believes that his peers' wages have risen by one standard deviation, his happiness level would decrease by 0.10 standard deviations. On the other hand, the absolute wage coefficient is consistently positive and significant at conventional levels, which implies that a worker's reported happiness increases as his wage goes up. The results suggest that a one-standard-deviation increase in a worker's own wage would lead to an increase in happiness of 0.13 standard deviations. This finding is reminiscent of the conclusion by Alesina et al. (2004) that "money buys happiness." Given that the impact of absolute wages on subjective well-being is stronger than that of relative wages, an across-the-board wage hike of, say, 10 percent would be associated with an overall increase in reported happiness. This finding is confirmed by an F-test that rejects the null hypothesis that the sum of the coefficients on absolute and relative wages is equal to zero at conventional levels.

[18] On this and other related issues, see Senik (2009).

[19] Section 5 discusses the consequences of estimating equation 1 using a mismeasured version of the "true" relative wage benchmark used by workers.

The absolute and the relative wage effects are substantially larger for males. A comparison between columns 5 and 6 shows that both the absolute and the comparison wage coefficients are about 35 to 50 percent stronger for men. This implies that women do not derive as much utility as men do either from their own labor earnings or their perception of their peers' wages. Nonetheless, in agreement with the literature, women generally report higher levels of life satisfaction *ceteris paribus*, as the results for the pooled sample in column 7 suggest.

Individual worker characteristics explain a significant amount of the variation in reported levels of life satisfaction. For instance, in the case of males, the inclusion of age, educational attainment, marital status, and dummies for whether the worker has a blue-collar position and whether it performs any managerial tasks in the company, increases the coefficient of determination of the regression from 2 to 7 percent.[20] The inclusion of these individual worker characteristics reduces the absolute wage coefficient by about 22 percent, although the impact of wages on happiness remains positive and strongly significant. Such reduction in the magnitude of the own wage coefficient is unsurprising given that standard Mincerian analysis has proven these individual characteristics to be important determinants of wages.

Pessimism seems to be an important individual trait that affects a worker's sense of well-being.[21] Pessimistic attitudes about the helpfulness of co-workers and about the possibilities of job promotions have large and significant negative effects on life satisfaction. These effects are of similar magnitude for both men and women. As discussed in Section 3.1, the inclusion of these two pessimism variables attempts to minimize concerns about wage endogeneity. Controlling for pessimistic attitudes leads to small decreases (in absolute value) in the magnitudes of both the own and the comparison wage coefficients. However, in both cases, the direction and magnitude of each of these effects are preserved. Additionally, these two pessimism variables by themselves explain an extra 4 percent of the variation in life satisfaction, which corroborates the relevance of a worker's individual characteristics in the determination of subjective well-being.

To verify that the relationship between happiness and both absolute and relative wages is not spurious, we investigate other channels through which these links may arise. Things like a company's wage structure, pay raises, and worker mobility within the establishment may affect how satisfied workers are with the salary they and their co-workers perceive. For instance, several authors have argued that high average wages within the company may provide a signal to the worker about his ability to rise within the firm's wage ladder Hirschman and Rothschild (1973); Manski (2000); Senik (2004); Clark et al. (2009b). If the worker believes that he has greater possibilities of increasing his compensation in the future, the negative impact of higher peers' wages would decrease in absolute terms. In other words, omitting average wages in a happiness regression would overestimate the comparison wage effect. Moreover, greater wage dispersion within the firm may also have a significant impact both on satisfaction and on the effect of absolute and relative wages on happiness. For instance, if workers know that there is greater variance in compensation

[20] Many previous studies control for both age and tenure in subjective well-being regressions. In our case, we opt to exclude tenure as a regressor because of the high correlation of this variable with age due to the characteristic seniority tenure system in the Japanese labor force described in Section 3.1.2.

[21] On a related issue, Stutzer (2004) found that higher income aspirations reduce individuals' life satisfaction. See also Frey and Stutzer (2010).

packages offered by their company, that may influence how they feel about their relative rank in the wage spectrum Brown et al. (2008).

We explore these possibilities by controlling for the logarithm of average wages and the interquartile range of log wages within the firm. As expected, higher average wages have a positive effect on reported levels of life satisfaction—although this effect is statistically significant only in the case of male workers. In contrast, greater inequality has a negative and significant impact on happiness for women, while it has no effect on men. Using the pooled sample, neither mean wages nor wage dispersion have a statistically significant effect on life satisfaction. In all cases, the inclusion of these variables has no discernible impact on the own wage and the reference wage coefficients; that is, the relative utility results continue to hold.

4.1 Heterogeneity in Wage Effects

We have now confirmed that the negative correlation between reference wages and subjective well-being holds even after accounting for a number of individual worker characteristics and controlling for other factors that may potentially explain away this relationship. In this section, we now explore in further detail one additional possible mechanism that may link these two variables.

Perhaps the most logical or straightforward reason that may justify such relationship is jealousy. Workers who like to compare themselves to other workers or who care about what their peers earn are likely to be better informed about the prevalent wage differentials at any given time. If this is true, workers who make smaller prediction errors should suffer greater disutility from any increases in wages earned by their peers.

To test this, we first calculate a worker's wage prediction accuracy and then estimate the impacts of comparison wages on his reported levels of life satisfaction. We define such prediction accuracy as the percentage difference (in levels) between the peer wage reported by the worker (i.e., his belief) and the wage average within each worker group or cell identified by the age of the worker, his gender, and his education level.[22] We sort the whole subsample of male workers by their prediction accuracy, and then group them into four broad categories, ranging from worst (i.e., 1st quartile of workers) to best (i.e., workers in the top 5 percent of predictors). Finally, we re-estimate the SWB equations for each of these groups, including the full set of controls that were accounted for in column 5 of Table 3.

The results appear in Table 4. The estimates corroborate our initial hypothesis that better predictors are more severely affected by changes in their peers' wages. Those who fare worst at predicting the wage of other workers with similar characteristics, do experience negative disutility from an increase in their peers' wages, although this effect is smallest relative to that of other workers who are more accurate in their predictions. The differences in these reference wage effects over well-being are large. For instance, workers in the the top 5 percent according to their prediction accuracy experience a negative impact on life satisfaction that is 5 times as large as that reported by workers in the bottom 25 percent of the sample (-0.21 for worst predictors vs. -1.06 for best predictors).

[22] Although we continue to focus on the male subsample only, in this calculation we control for gender because we expect male workers to compare themselves more often to other male, not female, peers.

Moreover, good predictors not only care more about their peers' wages, but also about *their own*. The differences in these own wage effects are also relatively large: about three times as large for best predictors compared to worst predictors. Nonetheless, these increases in the own wage effect as workers' prediction accuracy improves are not as marked as those observed for reference wages. One possible interpretation of this finding is that better predictors experience stronger feelings of jealousy, and so increases in their peers' wages will outweigh any increase in well-being they may get from a similar rise in their own wage. In turn, this result may imply that workers who care more about their peers' and their own wages will be more likely to be better informed about job offers and the wage structure in their profession.

As an additional remark, the estimates in Table 4 suggest that one cannot ignore cultural differences when comparing reference and own wage effects on well-being across studies that use different samples, especially if the data employed correspond to individuals in different countries. In Section 5.2, we compare our findings with those in previous literature keeping this caveat in mind. Before we do that, we corroborate the robustness of our results in the following section.

4.2 Robustness Checks for Self-reported Reference Wages

For brevity we show our robustness checks for the male sample only. Table 5 shows that our main results hold when we consider alternative specifications. For the sake of comparison, column 1 shows our preferred specification from Table 3, which includes individual characteristics (not shown), underlying pessimism, and firm's average wages and wage dispersion, with the addition of industry fixed effects. As is shown in column 2, accounting for year fixed effects, instead of industry fixed effects, leaves our main results unchanged.

In column 3, we address a possible endogeneity issue: a worker's reference group might depend on the company the worker chooses to work for, and unobserved amenities within the company might be correlated with both wages, reference wages, and happiness. Even though these companies are large[23] and might thus have more than one establishment, company fixed effects will also capture average neighborhood characteristics. There is no evidence that these endogeneities alter our results. Controlling for company fixed effects leaves them practically unchanged.

In column 4 we perform a horse-race between self-reported reference wage and the reference wage from external data (see Section 3.1.3). The coefficient on the external reference wage has the right sign and is significantly different from zero, but, most importantly, the coefficient on the self-reported reference wages is practically unchanged. One possibility for the negative sign on the external reference wage is that it proxies for the worker's permanent wages. The *difference* between the own wage and the external reference wage will thus proxy for unexpected positive shocks to income.[24]

In column 5, we use an ordered probit instead of the OLS. The coefficients are not directly comparable with the OLS ones but the relative size of the coefficient on wages and

[23]We observe approximately 90,000 workers and 62 firms, which means that each company has on average 1,500 workers.

[24]We thank Angus Deaton for pointing out this possibility.

the one on reference wages can be compared and the results are basically unchanged.[25] This result is in agreement with findings by Luttmer (2005) that the use of least-squares estimators in happiness studies does not impact negatively the general conclusions, even when ordered probit or logit models would be methodologically preferred given the hierarchical and non-linear nature of the dependent variable.

Column 6 shows that using job satisfaction as our subjective well-being measure does not alter the main findings: the coefficient on absolute wages, although slightly smaller in magnitude, is positive, strongly significant, and greater in absolute terms than the reference wage coefficient. As expected, the latter estimate is negative and significant at the 1 percent level.

5 Testing the Relative Utility Hypothesis Using Constructed Reference Wages

In the previous section, we used reference wages reported by workers to demonstrate how robust is the negative relationship between these self-reported reference wages and well-being. However, self-reported reference wages are not always available. Instead, most tests of the relative utility hypothesis found in the literature generally rely on reference wage measures constructed as some average wage defined in various ways, as discussed in Sections 2 and 3. In this section, we investigate whether these alternate wage measures deviate from the comparison wage benchmark truly perceived by workers, and if so, whether these differences introduce any significant biases in the estimated impacts of reference wages on well-being.

To preview our results, the findings below show that, in general, the estimated effect of reference wages on well-being is not consistent when using constructed reference wages as proxies for the truly perceived comparison wage. Although the theory suggests that the direction of the bias introduced cannot be determined, our empirical estimates show slightly smaller reference wage effects when we define this comparison wage variable using different types of cell averages. However, when reference wages are constructed from Mincer regressions, the estimated effects are vastly different from those suggested by all of our previous estimates, and highly unstable. We attribute these significant discrepancies to the difficulty in finding valid exclusion restrictions to justify the Mincer approach.

5.1 Revisiting the Subjective Well-Being Regressions

To investigate how the use of constructed reference wage measures may affect the empirical tests of the relative utility hypothesis, we first re-estimate the regressions in Table 3, now substituting the self-reported reference wage variable with the alternate measures used in the literature. The estimates are shown in Table 6. The first regression uses self-reported reference wages and produces estimates similar to the ones shown earlier in Table 3. The coefficients in columns 2-4 are derived from specifications that construct reference wages

[25] Nevertheless, the coefficients turn out to be similar in magnitude to the OLS case partly because the root mean squared error of the baseline regression is 1.04 and thus close to unity, to which the error term in the latent model of the ordered probit is normalized.

using various cell definitions. Cell 1 is defined by the age and education level of the worker and the average is computed using our own wage data. Cells 2 and 3 are both defined by the age of the worker, his educational attainment, and whether he is responsible for any managerial tasks within his company. The difference between these two averages is that cell average 2 is also computed from our own wage data, while cell average 3 uses the external BSWS wage database described in Section 3. Finally, the reference wage in column 5 comes from a standard Mincer regression similar to the ones shown in Table 2.

The results suggest that life satisfaction regressions that use constructed reference wages instead of self-reported ones do produce somewhat different results. The top panel illustrates how the raw effects of own and reference wages change when using alternate comparison wage measures, and so these first five specifications do not control for individual worker characteristics. In the case of specifications that use cell averages, both the own wage and the reference wage effects are always about 20-30 percent smaller in absolute value relative to the benchmark estimates. However, although the differences between each of the own wage and reference wage coefficients in columns 2-4 and the analogous estimates in column 1 are statistically significant, all of these coefficients are important in a statistical sense and the suggested directions of these effects are preserved. By contrast, the specification that uses Mincer-constructed reference wages tells a different story. Although the estimated own wage effect is positive and statistically significant, as before, the coefficient is 85 percent smaller than the one derived from a specification that uses self-reported reference wages. Moreover, the estimated impact of reference wages on life satisfaction is more than twice the size of the corresponding coefficient in column 1, *and positive*, which contrasts with the suggested effects in all of the previous specifications.

These findings are echoed in the bottom panel of the table, where we re-estimate the same five equations, this time controlling for worker individual characteristics. As shown in Table 3, the inclusion of these regressors slightly lowers the magnitude of both the own wage and the reference wage coefficients, although the general results are preserved. Once again, the estimates derived from regressions that use constructed reference wages are somewhat smaller in absolute value than the corresponding benchmark coefficients obtained when self-reported reference wages are employed instead. Nonetheless, the own wage effect on life satisfaction continues to be positive, the negative impact of reference wages on well-being persists, and all of the coefficients are strongly significant. The stark difference in the results is given, once more, by the specification with Mincer-constructed reference wages. This time, the size, direction, and statistical significance of the own wage coefficient are almost identical to those obtained using cell averages. Similar to specification 5 in the top panel that did not control for individual characteristics, the estimated reference wage effect under the Mincer approach continues to be positive, although the coefficient is statistically indistinguishable from zero.

In sum, empirical tests of the relative utility hypothesis that employ reference wage measures constructed by the econometrician imply reference wage effects on subjective well-being that deviate somewhat from those estimated when self-reported reference wages are used. In what follows, we attempt to uncover the reasons behind this bias.

5.2 Estimating the Bias

To better understand the differences described above, we utilize a generalized version of a the standard classical measurement error model. Recall that equation 1 gives the relationship between subjective well-being and the reference wage variable, $\bar{y}_i$. Now, suppose that the reference wage variable employed is a mismeasured version of the worker's correct perception of what his peers earn, $\bar{y}_i^*$. That is,

$$\bar{y}_i = \alpha_0 + \beta_0 \bar{y}_i^* + \epsilon_i. \tag{2}$$

This represents a generalization of the classical measurement error model, where $\alpha_0 = 0$ and $\beta_0 = 1$. We believe this generalization is appropriate in our case because the data suggest that reported reference wages differ significantly depending on workers' wage levels. For instance, people with low wages systematically perceive that their peers earn more than they do; that is, $\bar{y}_i^* > \bar{y}_i$ for low earners, which would imply that $\alpha_0 > 0$ and $\beta_0 < 1$. Similar to the classical measurement error model, we assume that the error term is orthogonal to the true reference wage, and that the relationship between the latter and the mismeasured variable is linear.

Substituting equation 2 into equation 1 obtains:

$$\text{SWB}_i = \alpha_1 y_i + \beta_1(\beta_0 \bar{y}_i^* + \epsilon_i) + \mathbf{x_i'}\gamma + \beta_1 \alpha_0 + u_i. \tag{3}$$

In such a model, a bias arises as a function of both the variance of the ϵ residuals and the distance between β_0 and 1. To see this, assume for simplicity that subjective well-being does not depend on worker individual characteristics.[26] In this case, the probability limit of the OLS estimate of β_1 is

$$\text{plim } \widehat{\beta}_1^{ols} = \frac{\text{cov}(\beta_0 \bar{y}^* + \epsilon, \text{SWB})}{\text{var}(\beta_0 \bar{y}^* + \epsilon)} = \frac{\beta_0 \text{cov}(\bar{y}^*, \text{SWB})}{\beta_0^2 \text{var}(\bar{y}^*) + \text{var}(\epsilon)} = \beta_1 \frac{\beta_0}{\beta_0^2 + \text{var}(\epsilon)/\text{var}(\bar{y}^*)}$$

The equation above suggests that, in general, $\widehat{\beta}_1^{ols}$ is not a consistent estimator of the true effect of comparison wages on subjective well-being when constructed reference wages are used. This statement is true as long as β_0 is different from $\beta_0^2 + \text{var}(\epsilon)/\text{var}(\bar{y}^*)$. Moreover, unless we can accurately determine whether β_0 is above or below 1, the direction of the bias is ambiguous.

Can we use instrumental variables (IV) to consistently estimate $\widehat{\beta}_1^{ols}$? In the classical measurement error model, an IV estimator would normally get rid of the bias introduced by the mismeasurement of the independent variable. By contrast, in the generalized version of the model, an IV approach does not lead to a consistent estimate of $\widehat{\beta}_1^{ols}$, although it does allow us to bound the estimated effect. To see this, suppose that there are two independent proxies of $\bar{y}^*$, proxy a and proxy b. We can then use proxy a as an instrument for proxy b:

$$\text{plim } \widehat{\beta}_1^{iv} = \frac{\text{cov}(\beta_0^a \bar{y}^* + \epsilon^a, \text{SWB})}{\text{cov}(\beta_0^a \bar{y}^* + \epsilon^a, \beta_0^b \bar{y}^* + \epsilon^b)} = \frac{\text{cov}(\bar{y}^*, \text{SWB})}{\beta_0^b \text{var}(\bar{y}^*)} = \frac{\beta_1}{\beta_0^b}. \tag{4}$$

[26]The results below do not hinge on this assumption. Using the Frisch-Waugh-Lovell theorem we could alternatively work with the residuals of a projection onto the space orthogonal to the individual characteristics. Alternatively we can derive the coefficient for $\bar{y}^*$ independently from $\mathbf{x}$ if self-reported reference wages are not dependent on workers' individual characteristics. The data suggest that this is a reasonable assumption. For instance, compare columns 1 and 2 in Table 3, and note that the estimated self-reported reference wage does not change when worker characteristics are introduced.

The expression above implies that an IV estimate of β_1 would still be biased, although now this bias depends only on the distance between β_0^b and 1, not on the variance of the ϵ residuals. Taking the ratio between $\widehat{\beta}_1^{iv}$ and $\widehat{\beta}_1^{ols}$ one can see that whenever OLS and IV give similar results, the signal-to-noise ratio $\text{var}(\bar{y}^*)/\text{var}(\epsilon)$ will be large. Unfortunately, β_0 might still be different from one and bias both estimates. Since workers with low wages tend to report a reference wage that is higher than their own wage, and *vice versa*, the regression of predicted reference wages on self-reported ones shows a positive constant term and a slope that is smaller than one. The IV estimate will thus tend to be larger, in absolute value, than the true coefficient.

Table 7 compares the OLS and the IV estimates. The OLS regressions are identical to the ones that appear in Panel A of Table 6 and are shown to facilitate the comparison with the IV estimates. The four OLS regressions are exactly the same, except each uses a different reference wage measure. The analogous results for Panel B are not reproduced for brevity. We focus on estimations that use self-reported and constructed reference wages since, as discussed above, these coefficients are stable across specifications. Results that use reference wages derived from Mincer regressions are explored further in Table 8 below.

The findings confirm the corresponding effects that own wages and reference wages have on life satisfaction. In all cases, each of the IV estimates is statistically indistinguishable from the OLS estimate obtained when self-reported reference wages are employed. Looking at equation 4, the direct implication of this is that $\widehat{\beta}_0$ is very close to 1, and that the IV coefficients do provide an unbiased estimate of the effect of reference wages on subjective well-being. Of course, it is possible that this result may not be generalizable to other similar databases measuring well-being and wages for different populations in other regions or countries. A generalized result, though, is that the IV procedure gets rid of one of the sources of bias. In our case, this is shown by the significant differences between each of the IV coefficients and the OLS regression that employs the corresponding mismeasured (constructed) reference wage measure. Moreover, with respect to the own wage effects, we observe that the IV coefficient is always somewhat larger than the corresponding OLS estimate. We attribute this result to the construction of the reference wage measure used in the OLS regressions as some average of the own wage variable. In this case, the inclusion of another variable that accounts for the mean wage of some subsample of the population that has similar characteristics to the worker's would most likely absorb some of the effect that would otherwise be (correctly) attributed to the own wage variable.

5.3 Understanding Mincer-Predicted Reference Wages

Thus far, our findings have been quite robust to different specifications and subsamples. They have also been relatively stable when using alternate definitions of comparison wages, except in the case of Mincer-predicted reference wages. In this section, we further investigate this discrepancy and argue that one of the main reasons behind the instability of the Mincer reference wage effects on subjective well-being is the indiscriminate use of exclusion restrictions that may not necessarily help accurately identify the reference wage effect. A proper instrument should be able to predict a reference wage, or define a reference group without influencing life satisfaction.[27]

Table 8 displays results for various tests of the relative utility hypothesis according to other papers in the literature that construct reference wages following a Mincer approach. Each column shows the own wage and the reference wage coefficients we obtain by running a specification as close as possible to the one employed in each of the papers, but using our own data. We also report the direction and significance of the reference wage effect on well-being that the respective authors obtain in their paper using their own data.

The differences in the estimated reference wage effects, both across specifications and across samples, are striking. First, the coefficients vary significantly depending on the specification, going from -0.41 to 0.64. Thus, the discrepancies arise not only in the magnitude of the coefficients, but also in the direction of the estimated effect. Such large variation in the implied impact of reference wages on well-being is surprising given that the underlying sample is exactly the same and the only thing that changes is the exclusion restriction assumed to identify the reference wage effect. Second, note how the same (or nearly the same) specification produces very different results from those reported in the original papers for their respective samples. For instance, while Sloane and Williams (2000) find a positive and insignificant effect of reference wages on well-being when they control for variables such as hours worked, tenure, marital status, and union management experience in the Mincer equation, we obtain a negative and very significant coefficient using the same specification for our sample.

We believe that what gives rise to such disparate results is the indiscriminate use of different exclusion restrictions and the unjustified categorization of variables as belonging in either the Mincer or the well-being equation. As noted earlier, the fact that our findings are quite robust to different specifications, including those that use various constructed reference wages, suggests that something not inherently related to the estimation of the subjective well-being equation may be driving the Mincer results. Under the Mincer framework, it is up to the econometrician to decide whether a given variable belongs in one or the other equation. For example, some of the authors that follow the Mincer approach have used different combinations of variables including age, education, and marital status in their exclusion equations, even when these variables have shown to be both economically and statistically relevant in the subjective well-being equation.

Of course, it is also possible that the discrepancies outlined above are due to other reasons. For instance, different papers use different samples. It is possible that the estimated reference wage effects on well-being are stronger or weaker, depending on the working population. It is also entirely possible that the "tunnel effects" described in Section 2 are more prevalent in other populations. Moreover, note that, for the sake of comparison, we are restricting our estimations to the same sample of male workers with which we have worked throughout the paper. Other papers in the literature have pooled males and females and used individuals that differ in their characteristics from our sample of unionized workers employed by large, publicly-listed firms in Japan. Nonetheless, estimations using different working subsamples (not reported for brevity) produce results that are, in some cases, significantly different from the ones reported for each of the specifications in the table, which further underlines the instability of the Mincer estimates.

[27]A recent paper by Card et al. (2010) does this in a clever way by using randomized manipulation of access to information on peers' wages.

Summing up, the results above suggest that the use of constructed reference wages in empirical tests of the utility hypothesis produces inconsistent estimates of the effects of comparison wages on subjective well-being. Our theoretical framework demonstrates that, in the presence of several mismeasured proxies of the reference wage that is truly perceived by the worker, an instrumental variables approach may help bound the comparison wage estimate, although it does not eliminate the bias entirely. An evaluation of the various empirical constructs that the literature has previously used shows that various cell wage averages, computed both within our own dataset and from an external wage data source, perform best in the estimation of the reference wage effects compared to the self-reported reference wage measure. However, the use of Mincer-predicted reference wages obtains reference wage effects that are unstable due to multicollinearity that ad hoc exclusion restrictions typically do not solve. It is thus advised that these Mincer-predicted reference wages are employed with care.

6 Conclusions

The results in this article find strong support for the relative utility hypothesis. Using data on self-reported reference wages, we observe that individuals report lower levels of both life and job satisfaction when they perceive that their co-workers earn higher salaries. However, unlike Easterlin (1974), we find that the association between absolute wage effects and subjective well-being is economically stronger than that between the latter and comparison wages.

Our analysis also shows that standard methods employed in the literature to estimate reference income measures yield inconsistent estimators of the reference income effect. Nonetheless, we demonstrate that these theoretical approaches do not suffer from the classical measurement error problem, and that the bias in the estimated relative effects cannot be signed. More particularly, we show that linear predictions of benchmark wage measures obtained by Mincer estimations perform poorly. This is mainly due to weak identifying assumptions of the comparison wage effect and a strong multicollinearity problem between predicted wages and the rest of the individual worker characteristics. In contrast, cell averages based on age, gender, and education levels and estimated using external datasets generate more reliable results.

From a practical point of view, our findings strengthen the case to increase resources that improve the quality of surveys and data collection on life happiness. Relative wage effects have obvious consequences in redistributive policies, both within a firm in the form of salary increases, as within a country in the shape of tax considerations. Only through a better understanding of these potential externalities on workers can firms and governments ameliorate the design of welfare-enhancing pay schedules and fiscal programs.

References

Abel, A. B. (1990): "Asset Prices under Habit Formation and Catching Up with the Joneses," *American Economic Review*, 80, 38–42.

Alesina, A., R. Di Tella, and R. MacCulloch (2004): "Inequality and Happiness: are Europeans and Americans Different?" *Journal of Public Economics*, 88, 2009–2042.

Alpizar, F., F. Carlsson, and O. Johansson-Stenman (2005): "How much do we care about absolute versus relative income and consumption?" *Journal of Economic Behavior & Organization*, 56, 405–421.

Bakshi, G. S. and Z. Chen (1996): "The Spirit of Capitalism and Stock-Market Prices," *American Economic Review*, 86, 133–57.

Blanchflower, D. G. and A. J. Oswald (2004): "Well-being Over Time in Britain and the USA," *Journal of Public Economics*, 88, 1359–1386.

Boskin, M. J. and E. Sheshinski (1978): "Optimal Redistributive Taxation when Individual Welfare Depends upon Relative Income," *The Quarterly Journal of Economics*, 92, 589–601.

Boyce, C. J., G. D. A. Brown, and S. C. Moore (2010): "Money and happiness: Rank of income, not income, affects life satisfaction," *Psychological Science*, forthcoming.

Brown, G. D. A., J. Gardner, A. J. Oswald, and J. Qian (2008): "Does Wage Rank Affect Employees' Well-being?" *Industrial Relations*, 47, 355–389.

Campbell, J. Y. and J. Cochrane (1999): "Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Cappelli, P. and P. D. Sherer (1988): "Satisfaction, Market Wages, & Labor Relations: An Airline Study," *Industrial Relations*, 27, 56–73.

Card, D., A. Mas, E. Moretti, and E. Saez (2010): "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," Nber working papers, National Bureau of Economic Research, Inc.

Carlsson, F., G. Gupta, and O. Johansson-Stenman (2009): "Keeping up with the Vaishyas? Caste and relative standing in India," *Oxford Economic Papers*, 61, 52–73.

Clark, A. E., P. Frijters, and M. A. Shields (2008): "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles," *Journal of Economic Literature*, 46, 95–144.

Clark, A. E., N. Kristensen, and N. Westergård-Nielsen (2009a): "Economic Satisfaction and Income Rank in Small Neighbourhoods," *Journal of European Economic Association, forthcoming.*

——— (2009b): "Job Satisfaction and Co-worker Wages: Status or Signal?" *Economic Journal*, 119, 430–447.

Clark, A. E. and A. J. Oswald (1996): "Satisfaction and Comparison Income," *Journal of Public Economics*, 61, 359–381.

Clark, A. E. and C. Senik (2010): "Who Compares to Whom? The Anatomy of Income Comparisons in Europe," *Economic Journal*, 120, 573–594.

Cole, H. L., G. J. Mailath, and A. Postlewaite (1992): "Social Norms, Savings Behavior, and Growth," *Journal of Political Economy*, 100, 1092–1125.

Corneo, G. and O. Jeanne (1997): "Conspicuous Consumption, Snobbism and Conformism," *Journal of Public Economics*, 66, 55–71.

de la Garza, A., A. Sannabe, and K. Yamada (2008): "Job Satisfaction and Happiness: New Evidence from Japanese Union Workers," Tech. Rep. 08-10.

Diener, E., E. Sandvik, L. Seidlitz, and M. Diener (1993): "The Relationship between Income and Subjective Well-Being: Relative or Absolute," *Social Indicators Research*, 28, 195–223.

Duesenberry, J. S. (1949): *Income, Saving and the Theory of Consumer Behavior*, Harvard University Press.

Easterlin, R. A. (1974): "Does Economic Growth Improve the Human Lot? Some Empirical Evidence," *In: David, P.A., Reder, M.W. (Eds.), Nations and Households in Economic Growth: Essays in Honour of Moses Abramowitz.*

——— (1995): "Will raising the incomes of all increase the happiness of all?" *Journal of Economic Behavior & Organization*, 27, 35–47.

——— (2001): "Income and Happiness: Towards an Unified Theory," *Economic Journal*, 111, 465–84.

Falk, A. and M. Knell (2004): "Choosing the Joneses: Endogenous Goals and Reference Standards," *Scandinavian Journal of Economics*, 106, 417–435.

Ferrer-i Carbonell, A. (2005): "Income and Well-being: an Empirical Analysis of the Comparison Income Effect," *Journal of Public Economics*, 89, 997–1019.

Frank, R. H. (1985): *Choosing the Right Pond*, New York, NY: Oxford University Press.

——— (2005): "Positional Externalities Cause Large and Preventable Welfare Losses," *American Economic Review*, 95, 137–141.

Frey, B. S. and A. Stutzer (2010): "Recent Advances in the Economics of Individual Subjective Well-Being," Tech. rep.

Gali, J. (1994): "Keeping Up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices," *Journal of Money, Credit and Banking*, 26, 1–8.

Hirschman, A. O. and M. Rothschild (1973): "The Changing Tolerance for Income Inequality in the Course of Economic Development; with a Mathematical Appendix," *The Quarterly Journal of Economics*, 87, 544–66.

Johansson-Stenman, O., F. Carlsson, and D. Daruvala (2002): "Measuring Future Grandparents' Preferences for Equality and Relative Standing," *Economic Journal*, 112, 362–383.

Kahneman, D. and A. B. Krueger (2006): "Developments in the Measurement of Subjective Well-Being," *Journal of Economic Perspectives*, 20, 3–24.

Knight, J., L. Song, and R. Gunatilaka (2009): "Subjective well-being and its determinants in rural China," *China Economic Review*, 20, 635–649.

Lévy-Garboua, L. and C. Montmarquette (2004): "Reported Job Satisfaction: What Does it Mean?" *The Journal of Socio-Economics*, 33, 135–151.

Luttmer, E. F. P. (2005): "Neighbors as Negatives: Relative Earnings and Well-Being," *The Quarterly Journal of Economics*, 120, 963–1002.

Manski, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531–42.

——— (2000): "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, 14, 115–136.

McBride, M. (2001): "Relative-income Effects on Subjective Well-being in the Cross-section," *Journal of Economic Behavior & Organization*, 45, 251–278.

Oswald, A. J. (1983): "Altruism, jealousy and the theory of optimal non-linear taxation," *Journal of Public Economics*, 20, 77–87.

Rayo, L. and G. Becker (2007a): "Evolutionary efficiency and happiness," *Journal of Political Economy*, 115, 302–337.

——— (2007b): "Habits, peers, and happiness: an evolutionary perspective," *The American Economic Review*, 97, 487–491.

Rizzo, J. A. and R. J. Zeckhauser (2003): "Reference Incomes, Loss Aversion, and Physician Behavior," *The Review of Economics and Statistics*, 85, 909–922.

Senik, C. (2004): "When Information Dominates Comparison: Learning from Russian Subjective Panel Data," *Journal of Public Economics*, 88, 2099–2123.

——— (2009): "Direct Evidence on Income Comparisons and Their Welfare Effects," *Journal of Economic Behavior & Organization*, 72, 408–424.

Sloane, P. J. and H. Williams (2000): "Job Satisfaction, Comparison Earnings, and Gender," *LABOUR*, 14, 473–502.

Solnick, S. J. and D. Hemenway (1998): "Is more always better?: A survey on positional concerns," *Journal of Economic Behavior & Organization*, 37, 373–383.

Stutzer, A. (2004): "The role of income aspirations in individual happiness," *Journal of Economic Behavior & Organization*, 54, 89–109.

Van de Stadt, H., A. Kapteyn, and S. Van de Geer (1985): "The Relativity of Utility: Evidence from Panel Data," *The Review of Economics and Statistics*, 67, 179–87.

Van Praag, B. M. S. (1968): *Individual Welfare Functions and Consumer Behavior*, North-Holland Publishing Company.

Veenhoven, R. (1991): "Is happiness relative?" *Social Indicators Research*, 24, 1–34.

Yamada, K. (2008): "Macroeconomic Implications of Conspicuous Consumption: A Sombartian Dynamic Model," *Journal of Economic Behavior & Organization*, 67, 322–337.

Yamada, K. and M. Sato (2010): "The Easterlin Paradox and Another Anatomy of Income Comparisons: Evidence from Hypothetical Choice Experiments," ISER Discussion Paper 0795, Institute of Social and Economic Research, Osaka University.

Table 1: Summary statistics[a]

	Whole sample	*Working subsamples* All	Males	Females
Worker characteristics				
Age (years)	34.926	34.676	35.921	30.153
	(9.96)	(9.64)	(9.58)	(8.46)
Tenure (years)	13.870	13.525	14.554	9.786
	(10.3)	(9.94)	(10.11)	(8.3)
Middle school [0,1]	0.069	0.044	0.048	0.029
	(0.25)	(0.2)	(0.21)	(0.17)
High school [0,1]	0.464	0.474	0.480	0.454
	(0.5)	(0.5)	(0.5)	(0.5)
Technical school [0,1]	0.024	0.025	0.024	0.030
	(0.15)	(0.16)	(0.15)	(0.17)
Some college [0,1]	0.095	0.095	0.042	0.286
	(0.29)	(0.29)	(0.2)	(0.45)
College [0,1]	0.270	0.283	0.315	0.167
	(0.44)	(0.45)	(0.46)	(0.37)
Post-graduate [0,1]	0.067	0.067	0.081	0.015
	(0.25)	(0.25)	(0.27)	(0.12)
Married [0,1]	0.581	0.583	0.653	0.329
	(0.493)	(0.493)	(0.476)	(0.469)
Blue collar [0,1]	0.341	0.336	0.369	0.219
	(0.47)	(0.47)	(0.48)	(0.41)
Managerial tasks [0,1]	0.171	0.181	0.225	0.023
	(0.38)	(0.39)	(0.42)	(0.15)
Hours worked (monthly overtime)	22.679	23.235	26.404	11.718
	(19.83)	(20.38)	(21.16)	(11.34)
Life satisfaction (1=lowest. . . 5=highest)	3.274	3.270	3.202	3.516
	(1.1)	(1.1)	(1.11)	(1.04)
Job satisfaction (1=lowest. . . 5=highest)	3.066	3.070	3.085	3.017
	(0.93)	(0.93)	(0.94)	(0.91)
Wage data[b]				
Own wages	15.401	15.405	15.508	15.028
	(0.42)	(0.42)	(0.37)	(0.37)
Self-reported reference wages	15.486	15.491	15.583	15.155
	(0.41)	(0.41)	(0.37)	(0.38)
Cell wage average from our data[c]	15.401	15.405	15.508	15.028
	(0.35)	(0.34)	(0.28)	(0.25)
Cell wage average from external data[c]	15.501	15.508	15.612	15.133
	(0.35)	(0.34)	(0.29)	(0.23)
Mincer-predicted wages[d]	15.406	15.405	15.508	15.028
	(0.26)	(0.26)	(0.16)	(0.21)
Observations	91896	78136	61278	16858

[a] For each worker characteristic, first number corresponds to sample average; second number (in parentheses) corresponds to standard deviation.
[b] Wage figures are in 2005 Japanese yen and are reported in logs.
[c] Cells defined as gender × age × education level.
[d] Predicted wages from a Mincer regression that controls for age, tenure, gender, education level, managerial task, blue collar, job shock, occupation, industry, hours worked, marital status, and year fixed effects.

Table 2: Mincer regressions

	(1)	(2)	(3)	(4)	(5)	(6)
		Males			Females	
Dependent variable:			*Fixed effects:*			
Annual income		Industry +Year	Company +Year		Industry +Year	Company +Year
Age in years	0.08***	0.08***	0.08***	0.06***	0.04***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age squared	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Tenure in years	0.02***	0.02***	0.02***	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Tenure squared	-0.00***	-0.00***	-0.00***	-0.00**	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Middle school [0,1]	-0.14***	-0.08***	-0.06***	-0.19***	-0.11***	-0.08***
	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.02)
Technical school [0,1]	-0.00	-0.02	-0.01	0.09***	0.07***	0.06***
	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.02)
Some college [0,1]	0.07***	0.06***	0.05***	0.09***	0.07***	0.06***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
College [0,1]	0.15***	0.12***	0.11***	0.25***	0.21***	0.19***
	(0.01)	(0.01)	(0.01)	(0.04)	(0.03)	(0.03)
Post-graduate [0,1]	0.20***	0.15***	0.12***	0.42***	0.35***	0.33***
	(0.02)	(0.02)	(0.01)	(0.04)	(0.04)	(0.04)
Other degree [0,1]	-0.06	-0.08***	-0.05***	-0.08**	-0.12***	-0.11***
	(0.04)	(0.02)	(0.01)	(0.03)	(0.03)	(0.03)
Hours worked (monthly overtime)	0.02*	0.04***	0.04***	0.03**	0.04***	0.04***
	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
Married [0,1]	0.09***	0.09***	0.09***	0.01	-0.00	-0.01*
	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Managerial tasks [0,1]	0.07***	0.08***	0.08***	0.11***	0.13***	0.15***
	(0.02)	(0.01)	(0.01)	(0.03)	(0.02)	(0.02)
Blue collar [0,1]	-0.01	-0.01	-0.01	-0.09***	-0.10***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	61278	61278	61278	16858	16858	16858
Adjusted R-squared	0.664	0.723	0.755	0.538	0.598	0.636

*** p<0.01, ** p<0.05, * p<0.1
Notes: Standard errors clustered by firm in parentheses. All wage measures and hours worked reported in logs. High school education dummy excluded from regression.

Table 3: Testing the relative utility hypothesis using self-reported reference wages

Dependent variable: Life satisfaction	Males (1)	(2)	(3)	(4)	(5)	Females (6)	All (7)
Own wage	0.69***	0.54***	0.46***	0.42***	0.43***	0.26***	0.40***
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.03)
Self-reported reference wage	-0.39***	-0.38***	-0.30***	-0.29***	-0.29***	-0.16***	-0.26***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
Hours worked (monthly overtime)	-0.02	-0.06***	-0.06***	-0.05***	-0.05***	-0.08***	-0.06***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Firm's average wages				0.11*	0.09	-0.03	0.05
				(0.06)	(0.06)	(0.09)	(0.06)
Firm's wage interquartile range					-0.09	-0.19**	-0.12*
					(0.08)	(0.08)	(0.07)
Pessimism about colleagues' helpfulness			-0.37***	-0.37***	-0.37***	-0.41***	-0.38***
			(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Pessimism about future promotions			-0.22***	-0.22***	-0.22***	-0.15***	-0.21***
			(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Female [0,1]							0.43***
							(0.02)
Age (years)		-0.05***	-0.04***	-0.04***	-0.04***	-0.05***	-0.04***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age squared (/100)		0.05***	0.04***	0.04***	0.04***	0.06***	0.04***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Married [0,1]		0.61***	0.60***	0.61***	0.61***	0.43***	0.56***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Middle school [0,1]		0.08***	0.07**	0.08***	0.08***	0.08	0.08**
		(0.03)	(0.03)	(0.03)	(0.03)	(0.11)	(0.03)
Technical school [0,1]		0.09***	0.08***	0.08**	0.08**	0.04	0.07**
		(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)
Some college [0,1]		0.04*	0.03	0.03	0.03	0.17***	0.13***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
College [0,1]		0.17***	0.13***	0.13***	0.13***	0.18***	0.14***
		(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
Post-graduate [0,1]		0.22***	0.17***	0.16***	0.16***	0.22***	0.17***
		(0.04)	(0.04)	(0.04)	(0.04)	(0.07)	(0.04)
Other degree [0,1]		0.06	0.05	0.05	0.04	-0.05	0.01
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Managerial tasks [0,1]		0.01	-0.03*	-0.03	-0.03	-0.06	-0.03*
		(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.02)
Blue collar [0,1]		-0.04**	-0.03*	-0.03*	-0.03*	-0.10***	-0.04***
		(0.02)	(0.02)	(0.01)	(0.01)	(0.03)	(0.02)
Observations	61278	61278	61278	61278	61278	16858	78136
Adjusted R-squared	0.022	0.074	0.111	0.111	0.111	0.087	0.117
Own wage + reference wage = 0? *p* value	0.000	0.000	0.000	0.000	0.000	0.113	0.001

*** p<0.01, ** p<0.05, * p<0.1
Notes: Standard errors clustered by firm in parentheses. All wage measures and hours worked reported in logs. High school education dummy excluded from regression. Except for column (1), all specifications account for industry fixed effects.

Table 4: Effects of own wages and self-reported reference wages on life satisfaction by accuracy of workers' prediction of peers' wages

Dependent variable:	(1)	(2)	(3)	(4)
Life satisfaction	*Prediction accuracy (quartile)*			
	1	2+3	4	Top 5%
Own wage	0.24***	0.66***	0.63***	0.65***
	(0.05)	(0.05)	(0.05)	(0.13)
Self-reported reference wage	-0.21***	-0.58***	-0.91***	-1.06***
	(0.04)	(0.04)	(0.13)	(0.32)
Observations	14216	30174	17588	3702
Adjusted R-squared	0.060	0.077	0.088	0.089
Own wage + reference wage = 0?				
p value	0.678	0.260	0.028	0.171

*** p<0.01, ** p<0.05, * p<0.1

Notes: Standard errors clustered by firm in parentheses. All wage measures and hours worked reported in logs. Prediction accuracy defined as the percentage difference (in levels) between peer reference wage reported by worker and wage average by cell defined as age × gender × education. All specifications control for individual worker characteristics including age, age squared, marital status, education level, managerial and blue collar dummies, as well as firm's average wages, firm's wage interquartile range, and industry fixed effects.

Table 5: Robustness tests

	(1)	(2)	(3)	(4)	(5)	(6)
				External	Ordered	
	Benchmark	Year	Year+Firm	ref. wage	probit	
Dependent variable:	Life satisfaction					Job sat.
Own wage	0.43***	0.43***	0.41***	0.43***	0.42***	0.25***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.04)
Self-reported reference wage	-0.29***	-0.30***	-0.29***	-0.29***	-0.30***	-0.17***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
Hours worked (monthly overtime)	-0.05***	-0.05***	-0.04***	-0.06***	-0.05***	-0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Firm's average wages	0.09	0.13*		0.09	0.09	0.15*
	(0.06)	(0.07)		(0.06)	(0.06)	(0.08)
Firm's wage interquartile range	-0.09	-0.14*		-0.09	-0.10	-0.13
	(0.08)	(0.08)		(0.08)	(0.08)	(0.10)
Pessimism about colleagues' helpfulness [0,1]	-0.37***	-0.37***	-0.37***	-0.37***	-0.38***	-0.38***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Pessimism about future promotions [0,1]	-0.22***	-0.22***	-0.22***	-0.22***	-0.21***	-0.44***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Cell wage average from external data				-0.14*		
				(0.07)		
Fixed effects	Industry	Year	Year+Firm	Industry	Industry	Industry
Observations	61278	61278	61278	61278	61278	61278
Adjusted R-squared	0.111	0.111	0.114	0.111	0.040[a]	0.121
Own wage + reference wage = 0?						
p value	0.000	0.000	0.000	0.000	0.000	0.029

*** p<0.01, ** p<0.05, * p<0.1

Notes: Standard errors clustered by firm in parentheses. All wage measures and hours worked reported in logs. All specifications control for age, age squared, education level, marital status, manager and blue collar positions.

[a] Adjusted R-squared not available; pseudo R-squared reported instead.

Table 6: Testing the relative utility hypothesis using alternate reference wage measures

	Panel A. Without individual characteristics				
	(1)	(2)	(3)	(4)	(5)
Dependent variable:	*Reference wage:*				
Life satisfaction	Self-reported	Cell average 1	Cell average 2	Cell average 3	Mincer
Own wage	0.67***	0.49***	0.48***	0.52***	0.10**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Reference wage	-0.39***	-0.22***	-0.20***	-0.26***	0.85***
	(0.03)	(0.07)	(0.05)	(0.05)	(0.13)
Observations	61278	61278	61278	61278	61278
Adjusted R-squared	0.025	0.021	0.019	0.020	0.024
	Panel B. With individual characteristics				
	(6)	(7)	(8)	(9)	(10)
Own wage	0.54***	0.37***	0.37***	0.37***	0.36***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Reference wage	-0.38***	-0.18***	-0.15***	-0.28***	0.06
	(0.03)	(0.05)	(0.05)	(0.07)	(0.14)
Observations	61278	61278	61278	61278	61278
Adjusted R-squared	0.074	0.069	0.069	0.069	0.069

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by firm in parentheses. All wage measures reported in logs. All specifications control for (log) hours worked and industry fixed effects. Specifications in bottom panel additionally account for individual worker characteristics including age, age squared, marital status, education level, managerial and blue collar dummies. Cell averages 1 and 2 calculated from our own data set and defined as (age × gender × education) and (age × gender × education × managerial tasks), respectively. Cell average 3 calculated using BSWS data and defined as (age × gender × education × managerial tasks). Mincer reference wage predicted from a standard regression of own wages on age and tenure dummies, hours worked, education level, marital status, manager, blue collar, job shock, and occupation dummies, as well as industry and year fixed effects.

Table 7: Use of IV estimates when self-reported reference wages are not available

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Reference wage:						
Dependent variable:	Self-reported	Cell average 1		Cell average 2		Cell average 3	
Life satisfaction	OLS	OLS	IV	OLS	IV	OLS	IV
Own wage	0.67***	0.49***	0.56***	0.48***	0.57***	0.52***	0.53***
	(0.032)	(0.025)	(0.033)	(0.027)	(0.035)	(0.026)	(0.031)
Reference wage	-0.39***	-0.22***	-0.34***	-0.20***	-0.35***	-0.26***	-0.28***
	(0.025)	(0.049)	(0.067)	(0.052)	(0.069)	(0.047)	(0.061)
Instrumental variable:			Cell av. 3		Cell av. 3		Cell av. 1
Observations	61278	61278	61278	61278	61278	61278	61278
Adjusted R-squared	0.025	0.021	0.021	0.019	0.019	0.020	0.020

** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by firm in parentheses. All wage measures reported in logs. Cell averages 1 and 2 calculated from our own data set and defined as (age × gender × education) and (age × gender × education × managerial tasks), respectively. Cell average 3 calculated using BSWS data and defined as (age × gender × education × managerial tasks). All specifications control for (log) hours worked and industry fixed effects.

Table 8: Impact of different exclusion restrictions on own wage and reference wage coefficients

Dependent variable: Life satisfaction	(1) Clark & Oswald (1996)[a]	(2) Sloane & Williams (2000)[b]	(3) Lévy-Garboua & Montmarquette (2004)[c]	(4) Senik (2004)[d]
Own wage	0.39***	0.20***	0.28***	0.22***
	(0.02)	(0.02)	(0.03)	(0.04)
Mincer-predicted reference wage	0.64***	-0.41***	0.45***	0.18**
	(0.10)	(0.06)	(0.06)	(0.08)
Effect of reference income in original paper	Negative and significant	Positive and insignificant	Negative and significant	Positive and significant
Exclusion restriction	Age categories Tenure, Tenure squared Education level Marital status Union management Firm and year fixed effects	Hours worked Tenure, Tenure squared Marital status categories married female married male Union management	Age Hours worked Tenure, Tenure squared Marital status Managerial tasks Occupation fixed effects	Age, Tenure Education level Marital status categories Managerial tasks Industry and occupation fixed effects
Observations	61278	61278	61278	55801
Adjusted R-squared	0.03	0.14	0.12	0.06
Own wage + reference wage = 0?				
p value	0.000	0.001	0.000	0.000

Standard errors clustered by firm in parentheses. All wage measures reported in logs. ∗ ∗ ∗*p*¡0.01,∗ ∗ *p*¡0.05,∗*p*¡0.1

[a] Additional variables in life satisfaction equation include hours worked, age, age squared, managerial tasks, industry and occupation fixed effects.

[b] Additional variables in life satisfaction equation include hours worked, age, educational attainments, managerial tasks, company fixed effects, six subjective variables regarding working conditions.

[c] Additional variables in life satisfaction equation include age categories, educational attainment, job satisfaction, satisfaction with leisure.

[d] Additional variables in life satisfaction equation include age categories, marital status, family income.

Table A-1: Literature survey

	Mincer wage approach							
Authors	Clark and Osawald (1996) (i)		Sloane and Williams (2000) (ii)		Lévy-Garboua and Montmarquette (2004) (iii)		Senik (2004) (iv)	
Data characteristics	U.K. workers		U.K. workers		Canadian workers		Russia	
Dependent variable	Job satisfaction	Income	Job satisfaction	Income	Job satisfaction	Income	Happiness	Income
Own-income effect	Positive		Positive		Positive		Positive	
Reference income effect	Negative		Positive		Negative (v)		Positive	
$\mid$Ref. income$\mid$ > $\mid$Own income$\mid$? (vi)	Yes **		No		No		No ***	
Hours worked	Yes	No	Yes	Yes	No	Yes	No	No
Age	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Age squared	Yes	No	No	No	No	No	No	No
Tenure	No	Yes	No	Yes	No	Yes	No	Yes
Tenure squared	No	Yes	No	Yes	No	Yes	No	No
Sex	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Marital Status/Children	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Industry	Yes	Yes	No	No	No	No	No	Yes
Occupation	No	Yes	No	No	No	Yes	No	Yes
Job class/ supervisor	Yes	Yes	Yes	Yes	No	Yes	No	Yes
Region	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Table and equation number	4(1)	A	8(1)	7	1,2(1)	A.1	3(3)	A.1
Method	Ordered Probit		Ordered Probit		Ordered Probit		Ordered Probit	

(i) *Other controls in happiness regression*; Renter, Has second job, Temporary contract, Health status. *Other controls in Mincer equation*; Health dummies, Establishment size dummies, Accident dummies, When work dummies, Sex mix at work dummies, Organization type dummies, Trade Union recognized, Pension member,Incentive payment, Union member, Part time, Temporary contract.
(ii) *Other controls in happiness regression*; Temporary contract, Firm size dummies, Part time, Owner-occupier, Promotion prospects, Frequent overtime, Percieved career, 14 Working condition variables, Manual dummy, Union dummy, Training period, Sex mix at work dummies, Race dummies. *Other controls in Mincer equation*; Female married, Male married, Length of training period, Shift work, Merit-related payment, Relevant trade union present, Trade union member, Selectivity, Part time, Temporary contract, Firm size dummies.
(iii) *Other controls in happiness regression*; Job related satisfactions, Country of birth, Mother tongue, Satisfaction with leisure, Religion. *Other controls in Mincer equation*; Canadian borne, Bilingual, Religion, Satisfaction with health, Satisfaction with leisure, Part time.
(iv) *Other controls in happiness regression*; Lagged individual income, Household income, Household size, Mother tongue, Believer, Round, Health.
(v) Current wage gap (own income minus reference income) and reference income controlled. The difference of the estimate of reference income and that of current wage gap is negative, indicating negative effect of reference income.
(vi) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A-1: Literature survey (continued)

	Cell average income from dataset			Cell average income from external source						Subjective reference income	
Authors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Data characteristics	Germany	U.K.	Denmark	U.S.	U.K.	U.S.	U.K. and U.S.	U.S.	Denmark	China	29 countries
Dependent variable (i)	H.	J.S.	J.S.	J.S.	J.S.	H.	H.	H.	H.	H.	H.
Own-income effect	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive
Reference income effect	Negative	Negative (ii)	Positive	Negative (iii)	Negative	Negative	Negative	Negative	Positive	Negative	Negative (iv)
|Ref. income| > |Own income|? (v)	No **	N.A. **	No **	Yes *	Yes **	Yes **	N.A. **	Yes **	No ***	N.A. ***	N.A.***
Hours worked	No	Yes	Yes	No	Yes	No	No	Yes	No	Yes	No
Age	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Age squared	Yes	No	Yes	No	Yes	No	Yes	Yes	No	Yes	No
Tenure	No	No	No	Yes	No	No	No	No	No	No	No
Tenure squared	No	No	No	No	No	No	No	No	No	No	No
Sex	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Education	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes
Marital Status/Children	Yes	Yes	Yes	No	aNo	Yes	Yes	No	Yes	Yes	Yes
Industry	No	Yes	Yes	No	Yes	No	No	No	No	No	Yes
Occupation	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes
Job class/ supervisor	No	No	Yes	No	Yes	No	No	No	No	No	Yes
Region	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Table and equation number	2 (1)	6A (2)	1 (1)	3 (2)	3 (3)	2 (1)	8 (4-8)	1 (1)	1 (1)	6 (4)	1 (1)
Method (vi)	O.P.	O.P.	O.P.	OLS	O.P.	O.P.	O. L.	OLS	OLS	OLS	OLS

(1) Ferrer-i-Carbonell (2005), *Other controls*; Year fixed effects, Familiy income, Number of children at home, Number of adults at home, Living together, Family condition, Social means (familiy income, years of education, number of children at home, and adults at home), Cell average income defined by (education, age, and region).
(2) Brown et al. (2008), *Other controls*; Employer size, Race, Temporary job, Mean pay, Pay range, Union recognition at the work place, Cell defined by work place.
(3) Clark et al. (2009b), *Other controls*; Year fixed effects, Health problem, Plant size, Cell average wages define by plant.
(4) Cappelli and Sherer (1988), *Other controls*; Layoff, Compare mkt, Compare outside, Compare carrier, Manage info, Union info, Union dues, Participations, Part time, Job influence, Change work, Change time.
(5) Clark and Oswald (1996), *Other controls*; Health dummies, Race dummies.
(6) McBride (2001), *Other controls*; Consumption habituation norm dummies, Race, Health dummies.
(7) Blanchflower and Oswald (2004), *Other controls*; Year fixed effects, Unemployed, Regional house price index, Household size, Race, Retired, Student, Keeping home.
(8) Luttmer (2005), *Other controls*; Value of home, Renter, Household size, Population, Non-metropolitan area, Fraction black in PUMA, Religion dummies, Race dummies.
(9) Clark et al. (2009a), *Other controls*; See neighbors often, Socio-economic groups, Number and ages of children, Years in Grid, Year fixed effects, Health problem.
(10) Knight et al. (2009), *Other controls*; Net wealth, Health status, Unemployed, Comparison variables, Community variables, Attitude variables, Race dummies.
(11) Senik (2009), *Other controls*; Family size variables, Owner ship of firms, Member of communist party.
(i) Abbreviations: *J.S.;* Job satisfaction, *H;* Happiness.
(ii) Reference income as "wage rank."
(iii) Reference income as "market wage."
(iv) Subjective comparison perceptions towards neighbors, family members, friends, and work colleagues.
(v) N.A. when magnitudes are not comparable. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
(vi) Abbreviations: *O.P.;* Ordered probit, *O.L.;* Ordered logit.

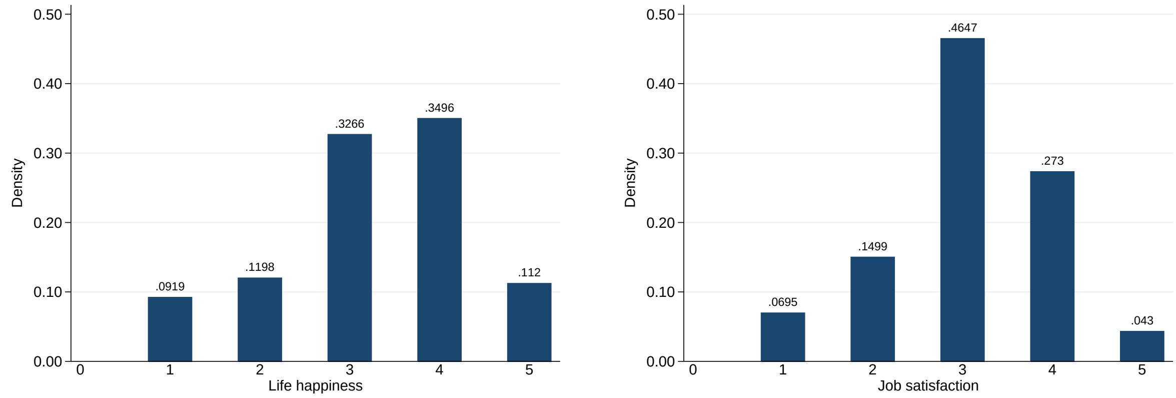

Figure 1: Distributions of life happiness and job satisfaction (1=lowest.. 5=highest)

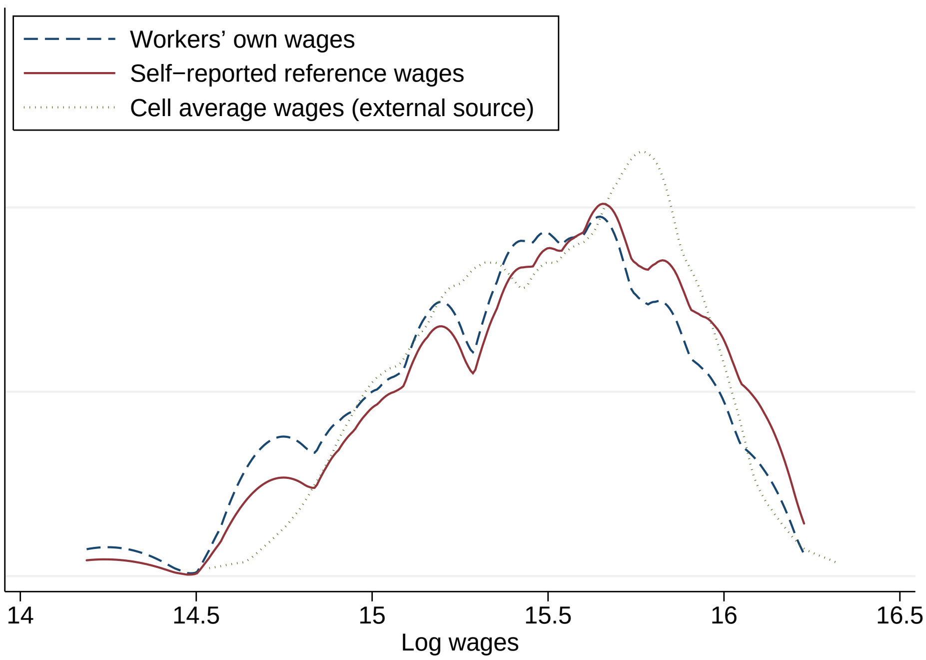

Figure 2: Density of (*log*) wages, self-reported reference wages and reference wages computed from external sources