Master Thesis

October 18, 2011

The effects of happiness on productivity: Panel evidence from the United Kingdom

Author:
Kirk GEALE

Supervisor:
Therese NILSSON

Abstract

This paper examines the causal effects of happiness on productivity, using cross sectional time series (panel) data from the UK, spanning the years 1996 to 2008. Innovations in this paper include the use of total compensation instead of wage rates as a proxy for productivity, the use of overall happiness instead of operationalized measures of happiness, and the use of GIS-derived bioclimate variables to instrument for happiness. The main result of the paper finds that happiness has a significant negative effect on productivity, in contrast to a large body of existing literature.

Keywords: Productivity · Happiness · Compensation · Psychological well-being · Bioclimate · LIML · British Household Panel Survey

Contents

1 Introduction

The question of whether happy workers are productive workers is one of interest in many academic areas and professions, particularly to businesses managers and behavioural psychologists. Productivity, defined as the level of output for a given level of input, is central to a firm's production process because higher productivity results in increased revenue, all else equal. In order to maintain the highest levels of productivity, around which classical profit maximizing behaviour is based, employers endeavour to keep their employees happy. It is common for them to do this in a variety of ways, from organizing Christmas parties to casual Fridays to flexible working hours. A clear question that should arise from these common observances is whether or not the ultimate goal of increased productivity is actually achieved. If time and money are being spent on increasing happiness with the ultimate goal of increased productivity, it should be shown that productivity gains are actually caused by increases in happiness. Otherwise, businesses need to reconsider diverting resources to these programs, or redefine the purpose behind these types of happiness-promoting corporate behaviours. This question is of interest because as Wright et al. (2004) notes, despite no clearly defined theoretical reasons predicting this result, the notion that happiness causes productivity has become a common 'truth' in the business world.

A second point of interest relates to the hiring of employees, and the factors that influence managers' choice of applicant. Personality evaluations to determine a fit in a particular corporate culture are common enough, and ask questions pertaining to various attitudes. If these questions lead to some assessment of how happy a particular applicant is, the employer should have an accurate perception of whether happiness benefits or detracts from productivity. This paper will attempt to provide some insight in this area.

In order to evaluate the relationship between happiness and productivity, the problem of how to measure a person's happiness must be resolved. In recent years, there has been a re-evaluation of the appropriateness of

classical measures of utility used in economics literature. Specifically, pecuniary measures such as GDP per capita are recognized as having less value than previously thought, spurring a migration towards more direct measures of well being, such as self-reported happiness. A person's happiness is often operationalized as workplace satisfaction, but this paper uses a broader measure, to capture happiness in a wide sense. Of the five existing five socio-economic panels that measure happiness and income, I use the British Household Panel Survey (BHPS). There are a variety of reasons for using micro panel data: panels are able to control for heterogeneity that cross sections are not, and they are not subject to aggregation bias. In light of the lack of causal research in the area of happiness on productivity, the BHPS panel is a relevant information set to utilize. This paper utilizes an updated definition of productivity from the previous literature, measured specifically by total compensation, building on Feldstein's 2008 paper. In addition, a new set of instrumental variables are used to develop a model that does not suffer from endogeneity. The results in this paper are consistent with the aggregate results of previous literature, when using similar techniques. However, an important departure occurs when accounting for the exigent issue of endogeneity, whereby it is found that happiness has a significant negative effect on productivity.

2 Literature Review

2.1 Productivity (Wages) and Happiness

In-depth and high quality analyses of whether happiness causes greater productivity are not as common as one might expect. A study in 2004 by Graham et. al. uses a Russian panel to analyze the effects of happiness on income during the period from 1995 to 2000. The general method used in Graham et al. is similar to the one applied in this paper, namely studying the same individuals over time in a cross sectional time series. Their research

finds that happiness has a positive effect on future income, and that the effect is of a greater magnitude for individuals who begin with lower income levels. Due to a variety of circumstances, there are problems with Graham et al. that I hope to overcome. First, only two time periods are present in their panel, the first in 1995 and the other in 2000. This is quite limiting in comparison to the BHPS panel that consists of 19 waves in total, of which 12 are utilized. This broader time dimension allows for higher confidence in the inferences of this paper, partly due to more reliable statistical tests that result from more degrees of freedom and variation when compared to a purely cross sectional study. A second concern is that the period of study in Graham et. al. was volatile in terms of Russia's economic policies and structural changes at that time. Structural breaks in the data can lead to a variety of misleading conclusions, a problem that this paper overcomes, since the UK has more political and economic stability than Russia. The authors also assent that there is concern some of the respondents included in the panel may have viewed the questions with suspicion, leading to dishonest responses. Despite these shortcomings, Graham et al. is an important paper with to compare due to its similarities with my research, especially in methodology and data.

In Oswald et. al.'s (2009) experimental setting, a randomized trial was used to treat a group with an increase in happiness in contrast to a control group that was not given the treatment, and found significant effects of happiness on performance. In a second experiment, they measured the effects of happiness on productivity by comparing a group that was "naturally" made happier through life circumstances, against those that were not, and found similar results to the first experiment. This paper is somewhat unique in the sense that it studies the subjects that were paid directly, without using some proxy or indirect method to increase wealth. It also has the advantage of randomized trials, but does not adequately control for the relevant variables. Another problem is the vagueness of questions such as "rate your happiness,"

which is the dependent variable in their study. These problems are discussed later in this paper, where it is explained why different operationalizations of happiness measure significantly different things.

Another perspective on the happiness literature explores whether income causes happiness, and the conclusion is generally that increased income does not lead to higher utility - this is known as the Easterlin Paradox. Again, income is related to this study with regard to the dependent variable compensation, which is similar to income. Tian and Yang (2009) outline possible theoretical constructs underpinning the Easterlin Paradox, showing that as long as there are negative consumption externalities (costs associated with consuming some good), an increase in income may not lead to an increase in happiness. To compliment this view, Eaton and Eswaran (2009) explain this phenomenon as the allocation of income to the purchase of Veblen goods[1], which are a consumption externality. As a result, income increases but utility does not. Causality moving from income to happiness is of interest with respect to the robustness of literature exploration, but businesses are more concerned with whether happiness causes productivity. This is because it is commonly assumed that happiness directly affects productivity which then affects business profits while the effects of income on happiness is only of interest if there is a looping effect, but this is indirect.

Experimental evidence for income causing increased happiness includes the natural experiment using lottery winnings and inheritances, exploited by Gardner and Oswald in 2001. They use the same data set as in this paper - the BHPS - and find that the classic supposition that money makes people happier is actually supported by empirical evidence. A major limitation of the Gardener and Oswald paper is the lack of long enough data to determine whether happiness reverts to a lower level. Second, they were not able to instrument the lottery winnings variable to eliminate self selection bias that

---

[1]A Veblen good is one that increases in demand for some people as a positive function of the good's price

could be present in gamblers.

The literature discussed so far provides support for the proposition that happiness causes an increase in income (or productivity), as well as for income causing happiness. As mentioned above, only one causal channel is of interest in commerce; specifically, whether happiness causes productivity.

There is also a large body of literature, especially in psychology, which finds relationships between happiness and productivity, but without a direction of causality. Hagerty (2003) finds that the relationship between happiness and income is a positive one, but non-linear: happiness is increasing but at a decreasing rate as absolute income increases. This is found in both cross sectional regressions, as well as in a short (3 wave) World Values Survey panel.

Another substantive body of research focuses on the concept of relative and absolute income. As an example, this research indicates that in a neighbourhood of rich people, the least rich person could feel unsuccessful and therefore unhappy, even though they are relatively rich compared to other citizens in the general population. In other words, it is the relative wealth of a person that determines their happiness, not absolute income. The concept of relative utility is explored in 2002 by Stutzer, in a paper that found higher income aspirations reduce people's happiness. This is related to Easterlin's (2001) research, which finds material aspirations begin similarly in various income groups, but these aspirations grow with income, and so the positive effect of income on happiness is reduced by the increase in aspiration. In other words, even though a person achieves some material aspiration decided upon at an earlier time, aspirations have grown since the goal-setting period and so (s)he is not satisfied with achieving the original goal.

Finally, there is research that indicates that there are causal channels running both ways: increased happiness both causes and is caused by increased productivity. Zelenski et. al. (2008) finds that happy people are more productive, and that people are more productive when they are happy

(between and within dimensions, respectively). This study uses self-reported productivity of 75 directors employed in the private and federal sectors in Canada. To create a panel, the respondents were surveyed 16 times over 8 weeks. Although the conclusions of the paper imply that "happiness may indeed foster productivity," there is no definitive support for the direction of causality. The authors find, as in a variety of related papers, a strong correlation of productivity with particular measures of happiness. Papers such as Zelenski et. al. reflect a large amount of literature showing a correlation between productivity and happiness (see Wright et. al., 2004; Hosie et. al., 2007). However, there is room for improvement over these results, since meaningful statistical and policy inferences are difficult to discern with non-causal results.

Another relationship worth describing is the one between health and productivity. This is important for two reasons: first, health is a variable of interest in the regressions that is instrumented. Second, happiness can be thought of as a facet of mental health, in contrast to the physical health component explicitly measured by the 'health' variable. In 2006, Weil found that individual health outcomes have significant, albeit small, effects on GDP per capita. The relatively small effect is in contrast to a large body of literature, including the position of the World Health Organization. In a macroeconomic context, Bloom et. al (2004) combines theory and empirical evidence by using a production function model backed by evidence from a panel of countries. The model predicts that health has a positive and significant effect on economic growth. In contrast to the micro-level paper Weil (2006), this paper shows relatively strong effects of health: an increase of 1 year in life expectancy corresponds to a 4 percent increase in output.

Bloom and Canning (2005) attempt to reconcile the macro- and microeconomic literature on health and productivity. Based on their findings, the papers above provide a good picture of the literature as it stands as of 2005. Bloom and Canning conclude that "health plays a larger role in

explaining cross-country differences in the level of income per worker" than on productivity. In the majority of the literature it is found that health has a positive and significant effect on productivity; it is the size of that effect that is usually the topic of interest. It should also be noted that there are a variety of different health indicators used in these studies, a possible cause for different findings in otherwise similar studies.

2.2   The concept of happiness

After some reflection, it becomes clear that the term happiness is not a particularly useful word to use in the formulation of a scientific question. In order to use this term as a valid form of measurement with any meaningful interpretation, it must be clearly defined: in this paper, what is meant by happiness is overall life satisfaction. It is a purposefully broad definition; more appropriate than operationalized definitions interpreting happiness as job satisfaction, which is often used in economic happiness literature. The job satisfaction interpretation is narrow and may not capture a person's actual level of happiness with their life as a whole, nor is it robust to recency[2] effects. The happiness variable of interest in this paper - overall life satisfaction - is the most appropriate choice due to the global nature of this self-reported measurement, which does not suffer from the affective state (mood) of the respondent at the time at which (s)he reports it (Lucas et. al., 1996). This variable is as broad as is feasible in a survey method, and by using such a measurement, it is possible to look at happiness as a whole without looking at satisfaction that is in some way restricted to a particular time period or area of a person's life. In other words, there is doubt that job satisfaction is a very good proxy for life satisfaction (Zelensky et. al., 2008).

In fact, in the BHPS data used in my research, there is a correlation of only 47 percent between job satisfaction and overall life satisfaction. This

---

[2]Current feelings about one's circumstances take precedence over one's global circumstance.

implies that a person's intangible, overall happiness may not be closely related to particular facets of their happiness (i.e., workplace) as has been assumed in previous research. The 47 percent correlation means that less than half of the variation in overall happiness is reflected by job satisfaction. The discord in the literature's findings regarding the happy-productive worker thesis suggests that small differences in the measurement of variables such as happiness may have a significant effect on the results. It is often complicated to objectively compare results from different authors since the definitions of happiness can differ significantly. Instead of happiness, the psychological community uses the term psychological (affective) well being (PWB), which is also a very broad term - Warr (1990) defines this measure to be a context free variable which is not tied to a particular situation. The definition used in this paper is consistent with PWB, allowing this paper to offer interdisciplinary comparability between the psychological and economic happiness literature.

2.3 Measuring Productivity

The question of how to measure an individual's productivity is certainly a valid one, and since it is not directly observable, it has been the subject of substantial research in its own right. This challenge has been characterized in the literature by the inability to uncover what extent one particular person's labour, often a minuscule unit within the entire production process, contributes to the final product. In neoclassical microeconomic theory, wages paid in competitive markets are equal to a worker's marginal product of labour. The marginal product of labour is defined as the derivative of the production function with respect to labour, holding the other inputs constant. This can be intuitively understood by reasoning that a firm is only willing to hire a worker if the wage paid is equal to or less than the revenue the worker adds to the firm. The worker, then, only accepts payment that is equal to or greater her productive worth, implying that the solution to the

system of equations is to set the wage exactly equal to the marginal product of labour.

A more sophisticated understanding can be obtained from the theory of human capital and wage differentials, where workers can become more productive through various channels, including education, job training, and experience. That being said, workers with equal productivity may accept different wages, which initially seems to be in contradiction with the wage theory described above. These differences in wages can often be attributed to compensating wage differentials; workers gain more utility from jobs in particular fields of work, with closer commutes, that offer more prestige, etc (Nicholson, 2004). Now, instead of the neoclassical monetary based interpretation of utility, it can be seen that favourable work conditions are a form of payment towards utility, that is also paid into by wages. In the extended theoretical construct, there is a disparity between marginal productivity that is not accounted for in the model, in addition to immediate wage adjustment that are not present in reality (refer to literature on sticky wages). It is assumed that while compensating wage differentials may exist for a variety of individuals, it is a zero-sum game. This implies that some workers will accept lower pay for favourable work conditions, and that some employers will pay more for unfavourable working conditions. As discussed next, empirical evidence from the United States and the UK indicates that a modification of the traditional notion that intra-period wages reflect productivity allows for meaningful estimations in this paper.

While neoclassical theory of competition is simplistic, yet logical within its own limited context, research has been conducted to determine whether wages actually reflect productivity, and in particular if it only applies to some demographics. There is controversy in this research; Hellerstein et. al. (1999) finds that in general, productivity is reflected in an individual's wages, in contrast to Crepon et. al. (2002) who find that workers' wage profiles rise more quickly than their productivity profiles. Hence, productiv-

ity is not reflected in their wage. Skirbekk (2003) agrees with this finding on one level, but points out that despite this empirical result, older workers may possess particular qualities that are important to the success of the firm, but hard to quantify. If these unquantifiable qualities exist, it would be very difficult to accurately determine whether productivity is reflected in the wage rate by using an empirical method such as econometrics. These examples suggest that studies attempting to determine whether productivity is reflected in wages have drawn inconsistent conclusions, due in part to potentially unmeasurable variables. Feldstein (2008) reveals a reasonable solution to this issue based on recent evidence from the US, that total compensation - not wages - fully reflects productivity, in a one-to-one relationship that is not statistically different from zero. The intuition behind this idea is that as a percentage of compensation, wages have recently fallen while benefits like health have increased proportionally. In relation to the discussion of the wage rate equaling the marginal product of labour, the idea is that the firm pays the worker at the rate (s)he is worth to the firm. Payment then, should include benefits as they are a cost to the firm directly to the worker. Making use of total compensation also allows for its components to vary, while compensation remains steady in inflation adjusted units.

In the US, wage and salary payments dropped from 89.9 percent of total compensation to 80.9 percent for the economy as a whole, during the period 1970 to 2006. While that period is longer than the 12 year period analyzed in this paper, it can be comfortably assumed that salary as a percentage of compensation dropped a large amount in the UK and is therefore a significant departure from the typical wage proxy. The compensation variable is logged as is standard practice, in order to linearize the equation. The important distinction between total compensation and wage rates is apparent when examining logged compensation and logged wage rates; the correlation between the two is 83 percent in the data used in my research. This discrepancy can clearly play a large role in the magnitude and significance of

regressed variables.

As noted by Feldstein, the amount of intra-year of correlation between productivity and compensation is 0.79 in the United States; the gain in productivity is only fully reflected in compensation after 2 years. Lazear (2006) also finds that wage growth lags productivity growth in the US. Both Lazear and Feldstein report that over the long run, productivity and hourly compensation grow together in a one-for-one relationship. To support this in a UK context, Alexander (1993) finds that while wage increases have an immediate effect on productivity, productivity increases take a little longer to be reflected as a wage increase. As a result, I have created a variable that approximates hourly compensation by taking the number of hours normally worked in a week, multiplying that by 4 to create a monthly value, and dividing total income last month by the normal number of hours worked in a week. This is not an exact value for hourly compensation rate, but one that is probably very close. These new values do not introduce any bias: the process described simply changes the period of measurement of the variable - essentially the variable is being multiplied by a constant, which has no effect on the mathematical expectation of the variable. Also, the respondents are interviewed at similar times each year, which should eliminate any seasonality effects that would potentially bias this variable. In the regressions presented in this paper, productivity is modelled two years ahead of the other variables, in keeping with the results presented in the US and UK evidence above. While this lag length may not be optimal in the UK, it is the best estimate available and is supported indirectly by Alexander (1993).

It is common to use wage as a sufficient statistic[3] for productivity (for example, see Dearden et. al., 2006). And, according to Sparber (2010) this is the case when when there is no direct measure of productivity. These

---

[3]A sufficient statistic is one that has the property that no other statistic can be calculated that gives more information about the true parameter from a given sample. Statistically, $P(x|t, \theta) = P(x|t)$, where θ is the true parameter, and t and x are realizations from the statistic $T(X)$ and the data X respectively.

two suppositions are supported by the neoclassical theory that the wage rate is equal to marginal productivity in a competitive context. However, it is important to point out that in reality, especially in the BHPS that surveys an entire country, it is certain that many individuals do not work in a purely competitive industry. It can be argued however, that real wage increases can provide at least a lower bound on the probable increases in productivity (Dearden et. al., 2006). Following in the spirit of Feldstein (2008) as mentioned above, this paper improves on the lower bound result by aggregating each person's wages and benefit income to give a more reasonable estimate of the effects of happiness on productivity.

2.4   Relationship Between Climate and Happiness

Some variables in the productivity-happiness equations can be thought of as endogenous, including productivity, happiness, and health (Borghesi and Vercelli, 2007; Contoyannis and Rice, 2001). Consequently, exogenous instruments must be utilized, and bioclimate variables are able to serve this purpose with respect to happiness.

Variables used to identify one of the reported IV specifications include mean temperature, mean precipitation, and altitude for each local authority district. Rehdanz and Maddison (2003) found that "climate variables can be used to investigate differences in self-reported subjective well being," precisely the variable used to measure happiness in this paper. These researchers find that people living in high latitude countries are significantly benefited by increased temperature and lower precipitation. There is further reason to use altitude as an instrument, due to the criteria discussed in Section 3.4, all three bioclimate variables can be considered exogenous to productivity in the UK. There is good reason to believe that a variety of climatic variables, however, directly affect happiness. Using geographic variables to identify happiness-productivity equations is not standard practice, but it allows for a different perspective on the results by comparing with more standard iden-

tification methods, both of which are presented in Section 4.1.

3   Data and Empirical Methods

3.1   Data

The annual British Household Panel Survey (BHPS) is administered by the ESRC UK Longitudinal Studies Centre at the University of Essex. The data comes in various forms, and in this paper the individual respondent records are used, as opposed to household records. The study spans the years 1991 to 2009, and covers the geographic area of the United Kingdom. The original 1991 wave contained 5500 households, of which 10 000 individuals who were over the age of 16 were interviewed. This sample has expanded, by the addition of Scotland, Wales, and Northern Ireland. In addition to tracking the original sample members, their children are also tracked once they reach the age of 16, as are people living in the same households as either the original member or their children. This study is nationally representative, and explores a wide variety of socio-economic questions, in addition to common survey questions regarding topics such as demographics, education, and income. The BHPS Conditional Access data is also used in conjunction with the GIS aspect of the analysis, as described below. This data set is also managed by the ESRC at the University of Essex, where Local Authority Districts (LAD) data can be obtained, also known as Unitary Authority Districts, for each individual respondent record. The data used in this paper's regressions are included in Table 1, with the usual summary statistics.

Interesting to note in Table 1 is the mean of happiness, which is quite high given a scale of 7. Health is also rated higher than its scale's average. A set of graphs of the EDF of happiness by health status is presented in Appendix A, showing that as health increases, the EDF of happiness becomes increasingly more skewed to the extreme of "completely happy." This suggests that people are more likely to rate themselves as happy when they

Table 1: Summary statistics

Variable	Mean	(Std. Dev.)	Min.	Max.	N
Controls					
Logged Compensation/Hr.	2.3312	(0.5755)	0.0008	6.3769	87755
Happiness	5.2329	(1.2903)	1	7	160518
Health	3.8043	(0.9526)	1	5	156007
Age	45.6285	(18.6922)	15	101	171682
Cohabiting or Married	0.6405	(0.4799)	0	1	171690
Divorced	0.0547	(0.2273)	0	1	171690
University Education	0.3579	(0.4794)	0	1	171690
A/O-Levels	0.2858	(0.4518)	0	1	171690
Social Class	4.8107	(1.3003)	1	9	102955
Self Employed	0.0687	(0.2529)	0	1	171690
Conservative Supporter	0.1492	(0.3563)	0	1	171690
Labour Supporter	0.2412	(0.4278)	0	1	171690
Union Membership	0.4906	(0.4999)	0	1	84560
Instruments					
Altitude	108.9718	(77.6410)	-2	386.273	147271
Mean Precipitation	851.6534	(226.4842)	543.897	1782.91	146949
Mean Temperature	91.4141	(10.1361)	0	112	147455
Smoking Status	0.2265	(0.4186)	0	1	171677

are in good health. Approximately 36% of respondents have post-secondary education (University Education in Table 1), and approximately 64% are divorced. A/O Levels refers to British high school graduate certificates, and represents the number of respondents who's highest educational is an A or O level diploma. Only 7% of respondents are self employed. Also, the mean age is 46, a little above the median age of 44. Overall, the summary statistics display no unusual components, despite perhaps, the minimum of Altitude; however, it is entirely possible to live at an elevation below sea level. Not enough is know about UK geography to determine whether that is in fact a true value or measurement error, but either way there are only 9 respondents with elevations below 0.

Many of the variables used in the regressions were modified from the original data in the BHPS. First, the logged hourly compensation variable

was modified as described in Section 2.3. Other changes are relatively simple, including the age squared variable which is self-explanatory. The variable for health was rescaled so the highest value was the highest health, in order to give intuitive coefficient estimates. This simply involved taking the variables maximum value, adding one, and subtracting each respondent's health rating from the sum of $\max(health) + 1$. The construction of each indicator variable is also straight forward: setting it equal to 1 if the title is true, and 0 if it is false.

ArcGIS 9 was used to process the geographic data in this paper, which includes administrative boundary files and bioclimatic information. The administrative boundaries for the UK are at the level of LAD's, which is the smallest geographic area that is feasible to use in this paper, and is more localized than the regional variables that are typically employed. These GIS-derived values are matchable to the LAD's in the BHPS Conditional Access data. The bioclimatic variables used are yearly rainfall, temperature, and altitude, and all three use a resolution of 2.5 arc minutes. The data was compiled from a variety of different climate stations wherever available, and is approximately representative of the years 1950 to 2000 (Hijmans, 2005). 2.5 arc minute data is used despite the availability of 30 arc second resolutions, because computer capacity limitations required the use of lower resolution. This should not prove to be a major detractor, due to the overall low density of available climate stations. The administrative boundary data is in vector format, and was obtained from the OS OpenData service provided by the UK Ordinance Survey Department. The bioclimatic data, in raster format, was obtained from the WorldClim group. The two bioclimatic variables were averaged by LAD, and then matched to each individual's LAD in the BHPS sample. The three bioclimate variables can be represented visually as in Figure 1, which shows the variation in precipitation (mm), temperature ($10 \times °C$), and altitude levels (m) geographically.

A brief note on the programs that were used to derive, process, and report

(a) Mean annual   (b) Mean annual   (c) Altitude
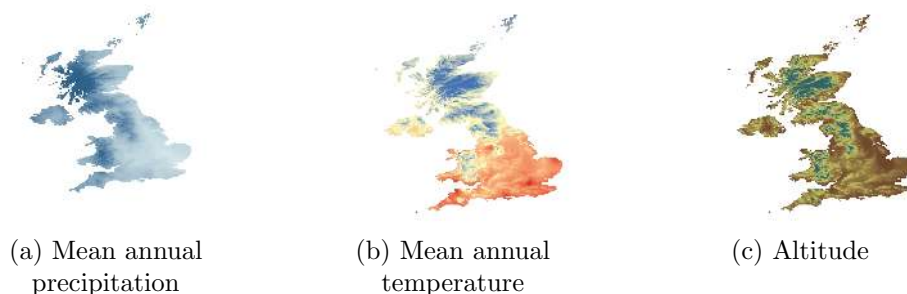   precipitation   temperature

Figure 1: Variation in bioclimate data

these data can be found in Appendix D.

3.2   Estimated Models

There are 4 models presented in this section, of which specification (3) is the
main empirical result. While the specifications change slightly, the control
variables are consistent throughout. A summary of these controls can be
seen in Table 1 in addition to the dependent variable and two endogenous
variables, with the exception of the regional and time indicator variables.

The data is initially analyzed with the use of the fixed effects estimator in
Equation (1), which allows individual heterogeneity to be accounted for. The
error term is made up of two components: $\varepsilon_{i,t} = \alpha_i + v_{i,t}$. By subtracting the
each time averaged variable from the original variable, the individual fixed
effects, α_i, are removed, since α_i is invariant over time. Equation (1) is the
model used in the estimation of Specification (1) in Table 3. X is the set
of regressors, α is the individual error component that does not vary over t,

and v is the stochastic error component that varies over t and i.

$$Qy = QX\beta + Q\alpha + Qv \tag{1}$$

$$(I_{NT} - P)y_1 = (I_{NT} - P)X_1\beta_1 + (I_{NT} - P)v_1 \tag{2}$$

$$(I_{NT} - I_N \otimes \bar{J}_T)y_1 = (I_{NT} - I_N \otimes \bar{J}_T)X_1\beta_1 + (I_{NT} - I_N \otimes \bar{J}_T)v_1 \tag{3}$$

The expansion from (2) to (3) shows that P is equal to the identity matrix multiplied by $\bar{J}_T = \frac{J_T}{T}$ (operated on by the Kronecker product) where $\underset{T \times P}{J_T}$ is a matrix of ones, thereby averaging the observations across time. P is the reduced form of the projection matrix on Z. $Q = I_{NT} - P$ is the orthogonal projection matrix on Z, so Q is a matrix that results in deviations from the individual's time-dimensional mean. See (Baltagi, 2005) for a more detailed explanation. Using this logic, or simply applying an expansion of Q to α, reveals that the individual fixed effect is eliminated. Second, in Table 3 Specifications (2)-(4), instrumental variables are used and therefore a new estimator is necessary. Using a method very similar to the one-way fixed effects estimator in Equation (1), instrumental variables are used, and the regression parameters are estimated using two stage least squares (2SLS). The estimator as described can be seen mathematically in Equation (4), and is known as the Within 2SLS estimator (W2SLS). The conditions required for an IV are presented in Equation (5), and the corresponding fulfilment of those conditions are discussed in Section 3.4. In Equation (4), Y_1 is the set of endogenous variables with coefficient γ, X_1 is the set of exogenous variables with coefficient β, α_1 is the one way error component, and v_1 is the two-dimensional error.

$$Qy_1 = QY_1\gamma + QX_1\beta + Q\alpha_1 + Qv_1 \tag{4}$$

3.3   Model Refinement

The models are checked for the appropriateness of random or fixed effects using a Hausman test. The test rejects that the coefficients are the same for the two models, indicating that there is information in the fixed effects model that is not present in the random effects model. The random effects model is more efficient than the fixed effects model, but is biased if any of the independent variables and individual effects are correlated. This is very likely the case, and the Hausman test confirms it by strongly rejecting the models' equality, at the 1% level.

The model is further examined for the statistical significance of the regional fixed effects and the time fixed effects. This is done by testing if the regional effects are jointly equal to zero, and whether the time effects are equal to zero, using a Wald test in both cases. The null hypothesis of both tests are strongly rejected, indicating that both sets of fixed effects are important in explaining the model. Not too much emphasis should be placed on this test[4], but it is included here for standardization with other literature and comparability to other papers and models.

The data is tested for group-wise heteroskedasticity, and unsurprisingly, the null hypothesis of homoskedasticity is strongly rejected. This form of heteroskedasticity probably exists, since there is likely to be heteroskedasticity between individuals as is common in regressions with dependent variables similar to wage. A second typical concern is autocorrelation. However, it is neither necessary nor feasable to test this model for general autocorrelation. Despite requiring too much computational power, this panel is a micro panel (having many more panel observations than time observations) and when N is

---

[4]Tests such as this, where a simple or sharp hypothesis is tested in the context of a continuous variable, is standard in the economics literature despite being somewhat nonsensical. This is because asymptotically, any test with a simple or sharp hypothesis is always rejected, since a single point under any continuous distribution has zero width – and thereby has zero probability of occurring. So, the test is irrelevant because for a small sample size the test may not be rejected, but as the sample size approches infinity, it will surely be rejected, and no real information is gained from this test.

large relative to T, contemporaneous correlation (cross sectional dependence) is not an issue (Baltagi, 2008; Angrist and Kruger (2001); Torres).

Finally, in Specification (4) using bioclimate IV's, interaction terms are included, chosen by the minimization of the Bayesian information criteria (BIC). In the other 3 specifications, the BIC indicates that no interaction terms are required as instruments.

3.4   W2SLS Identification

There is an inherent difficulty in measuring the effects of happiness on productivity that is often ignored in the literature. Some studies suffice to say that there is a link between happiness and productivity, without showing any causality in these variables. The size and significance of this relationship is reflected in a simple OLS regression, shown in Table 2. This table reports only the coefficients of correlation, SE's, and significance levels, as more information is not of interest, and similar relationships can be observed in many existing studies. The same control variables are used in this regression as are mentioned in Section 1.5, and are in fact used in all regressions. As Sutzer (2002) points out, even in panel data there is still a problem es-

Table 2: Bootstrapped OLS Results

	β	SE	p-value
Happiness	0.0120	0.0023	0.000
Health	0.02557	0.0032	0.000
No. Obs	71365		
R^2	0.3762		

tablishing causality. The fixed effects estimator in Equation (1) eliminates the individual heterogeneity, it does not remove the endogeneity between

happiness, health, and productivity; meaning that these three variables are jointly determined. The intuition behind this endogeneity is fairly clear, and it is confirmed by a Davidson-MacKinnon (1993) test of exogeneity, designed specifically for fixed effects regressions.

$$\mathbf{0} = E[\mathbf{Z}_{is}|\varepsilon_{it}] \tag{5a}$$

$$\mathbf{0} \neq E[\mathbf{Z}_{is}|\mathbf{X}_{it}] \tag{5b}$$

In more detail, the IV method involves using a set of instrument(s) that satisfy the moment conditions in Equation 5. 5.a shows that if the instrument set $\mathbf{Z}$ is orthogonal to the error term, it thereby solves the problem of endogeneity, and it is therefore considered a valid instrument. If $\mathbf{Z}$ is chosen such that it is not orthogonal to the error vector, then $\hat{\beta}$ is biased. While the potential to remove endogeneity using IV's is appealing, it is unfortunately not possible to objectively test the validity of the instrument statistically[5], and so a good argument must be made for each instrument's orthogonality condition. The second condition 5.b is technically not required for instruments, but without this moment condition the instrumental regression would have absolutely no predictive power, which is of no economic value. Intuitively, the stronger the correlation between the instrument set and the instrumented variable, the better the result due to increased accuracy of the model's predictions resulting from smaller standard errors. Unlike 5.a, condition 5.b is clearly testable, simply by measuring the correlation between the instrument and endogenous regressor. Another benefit of the IV is to remove bias created by measurement error, which is likely in the self-reported health variable due to inaccurate reporting of particular types of illnesses (Baker et. al., 2001). The problems of heterogeneity, endogeneity, and measurement error are in general mitigated by the use of the fixed effect estimator in conjunction with

---

[5]This is because testing an instrument's validity requires one of the instruments to be valid, and so it is not possible to be sure that any instruments are valid beyond a theoretical explanation

the IV method.

This paper uses three different instrumental variable regression specifications. The first uses lagged values of the two endogenous right-hand-side variables, happiness and health. The lags of each of these two variables are highly correlated with their counterpart in the current time period. Following the assumption of no autocorrelation as discussed earlier, due to the data's micro panel appearance, the error term in this period (ε_t) should have no correlation with ε_{t-1}. See Appendix A for a statistical explanation. Due to my inability to test this assumption statistically despite the underlying logic, I have used a second lag instead of only one in order to be conservative, with the expectation that if there was some small amount of autocorrelation present with the period t-1, it would have diminished by the period t-2. In addition, changes in levels of happiness or health are not something that fluctuate significantly over short periods of time. Hence, the lagged values make for very good instruments since they are very closely related to the present values both theoretically and empirically. This close relationship does not undermine condition 5.b, since each panel has its own equation with a current value that is independent of the value in periods past.

While the lagged health and lagged happiness variables are quite highly correlated with the instrumented variables, the bioclimate variables from specification (4) and smoking from specification (3) and (4) are not. From the IV discussion above, condition 5.b is fulfilled but with a low level of correlation, and the instruments are called weak. There is no exact definition of a weak instrument, but a common guideline is an F-value of less than 10 in the first stage regression. Improving on the F-value criteria is the comparison of the Stock-Yogo weak instrument test statistic to critical values based on relative and maximum bias size (Stock and Yogo, 2005). Both specifications (3) and (4) are weakly identified by the F-test criteria as well as the Stock-Yogo statistic. Thanks to recent research dealing with weak but valid instruments, it is possible to estimate these parameters with much

more reasonable confidence intervals. Andrews and Stock (2005) recommend using Fuller's 1977 modified limited information maximum likelihood (LIML) estimator[6] as it allows for the use of weak instruments, and exhibits good overall properties. The alpha value chosen is 4, which is also recommended by Andrews and Stock. Further reasoning for the use of the Fuller$_{\alpha=4}$ estimator is that it does well with a small number of instruments in terms of mean squared error (MSE), in comparison to the classic IV technique. In addition, estimators with well defined sample moments should be used, and typically, LIML estimators do not have sample moments. However, the Fuller LIML does, and performs best in Hahn et. al.'s Monte Carlo simulations (Hahn et al., 2004).

While a person's level of happiness may always be a subjective experience, them same is not quite so true for health. This variable is each person's self-assessed measure of health, which is subject to reporting error for a variety of reasons. These reasons could include expectations of their health, what health means exactly, and how often they use healthcare (Bago d'Uva et al., 2006). By matching self-reported health measures to actual health records, Baker et al. (2004) find that there is indeed measurement error by strongly rejecting the hypothesis that the self-reported measure is equal to the medical record. Baker et al. (2004) also find that there is a significant amount of false positive and negative reporting of specific ailments. While it may be argued that there is error in the self-reported smoking variable as well, it is reasonable to infer that it is a *more* objective measure than overall health. First, a person who smokes likely does not see it as a social stigma as frequently as a non-smoker, and will be likely to report this activity. Second, since smoking is not an illness in itself, Baker et al.'s conclusion that specific ailments are inaccurately reported does not apply to smoking. For these reasons, specification (3) and (4) use an indicator variable for whether or not the person smokes as an instrument for health, with the goal of using an

---

[6]The estimator is $\frac{x'Py - (\phi - \frac{\alpha}{n-K}) \cdot x'My}{x'Px - (\phi - \frac{\alpha}{n-K}) \cdot x'Mx}$, where α is a parameter to be chosen.

alternative self-reported variable that represents health more objectivity.

More important than providing a decrease in measurement error for health,is for smoking status to be exogenous to productivity in order to be a valid instrument. The explanation is relatively simple, in that any effect of smoking on productivity would intuitively come from health. There is no reason to assume that the act of smoking has any effect whatsoever on productivity. This could be imagined if employees were not allowed working breaks to smoke, and took time out of their working hours. However, modern workplace legislation guarantees regular breaks, a time at which employees may smoke. In addition, smoking does not take a large amount of time, even for heavy smokers that may in fact take extra time from working hours since break may be longer apart than nicotine cravings . A final support is the increasing prevalence of products such as nicotine gum, which allow workers to satiate nicotine cravings without stopping work. This still allows them to be classified as smokers, but they may choose to refrain from smoking at work.

The Fuller (1977) estimator is used in specification (4) in conjunction with three bioclimate variables: mean annual temperature, mean annual precipitation, and altitude of each respondent each year, over each LAD; these were discussed in Section 3.1. Temperature and precipitation were selected based on the 2003 paper by Rehdanz and Maddison, which shows that bioclimate variables, in particular temperature and precipitation, have a significant effect on happiness. A third bioclimate variable measuring altitude improves identification and can also be considered valid, and as such it has also been included as an instrument. As discussed earlier, these variables are averaged over a longer timespan than the panel itself. Obviously, these instruments only vary over time for respondents who relocated, but there are sufficient numbers of people moving that the FE estimator works for these bioclimate variables.

Bioclimate variables should be considered exogenous because they do not

vary appreciably between LAD's within the UK. Hence, any increase in these variables should have no direct effect on a person's productivity, as there is no effect on their working conditions, type or hours of work, or other related concepts. One might argue that if the amount of winter precipitation, snow, increased dramatically, then a person could have spent a lot of energy getting to and from work and therefore have less productivity on the job. This is valid, except that the UK does not exhibit a lot of winter precipitation, and as mentioned previously, increases in this type of precipitation would be moderate at best.

Also, even if the temperature were to dramatically increase (which is not reflected in the data), a large number of people work in offices which are governed by allowable ranges of temperature fluctuation. In other words, a large increase in temperature would be moderated by higher usage of an air conditioner in the workplace, leaving productivity unchanged.

In terms of altitude's exogeneity, a critic could question the atmosphere's oxygen content, but in the UK it would be completely unnoticeable, as the altitude variation is not so extreme – less than 400 meters – that workers would be short of breathe. Even if the variation was high enough for that possibility, the human body adjusts relatively quickly to that sort of environmental change. Elevation is likely linked to happiness since Great Britain is an island with a large number of port cities. I argue that people living in port cities, which by definition are low altitude, are living in larger cities and are therefore happier since there are more opportunities for enjoyable entertainment, cultural diversity, and a variety of other factors that make life in general easier and more enjoyable. The significance of altitude in the first stage regression in (4) is not significant, and so this conjecture cannot be directly evaluated.

In summary, these three bioclimate variables affect productivity, but only indirectly through happiness. Bioclimate variables in a low-variation environment such as the UK do not suffer from instrument invalidity, for the

reasons described above. As Rehdanz and Maddison (2003) outline, there is evidence for significant effects of temperature and precipitation on happiness. Elevation is also used as an instrument since there is support for believing both its exogeneity and effect on happiness. The instruments in specification (4) using the bioclimate variables pass the test of validity using the Hansen J Statistic under the null hypothesis of validity, with a p-value of 71% – but as discussed earlier, I am of the opinion that the reader should focus primarily on the theoretical arguments for validity, using the overidentification test as a confirmation.

3.5   VCV Estimates

In all of the fixed effects estimations, the variance-covariance (VCV) matrix is estimated using bootstrapping instead of using asymptotic values or robust VCV matricies. Bootstrap estimates can provide a number of benefits. First, the distribution of a test statistic does not need to be assumed. This allows the data to provide a test statistic based on the EDF, allowing for more reliable tests. Related to this is the tendency for bootstrapped tests to outperform asymptotic tests, improving the likelihood of garnering valuable inferences.

The number of bootstrap repetitions in each regression is 999, in order to create intuitive p-values, as well as to ensure that there is very little loss of power in the computation of the variables' significance levels[7] (Davidson and MacKinnon, 1999). Bootstrapping is more complicated, however, for IV panels. In this case, the bootstraps are clustered, and each resampled drawing is assigned a new id number, so that the bootstrap takes into account the panel nature of the data (Sanchez, 2011). By doing so the bootstrap resamples from each cluster, in this case the cross-wave person identifier, instead

---

[7]Because of repetitions are not infinite as in an ideal bootstrap, there is random variation in the finite number of bootstrap samples that reduces the power of the computation. This power loss is reduced by increasing the number of bootstrap samples computed(Davidson and Mackinnon, 1999).

of selecting randomly from any of the panel observations. Hounkannounon (2008) proves that this individual bootstrap method is consistent for a one way error component model in the dimension of i, which is the FE estimator in Equation (1) used in the reported regressions.

4 Results

4.1 Regression Results

The regression results found in Table 3 show 4 specifications with productivity measured by logged total compensation two periods in the future as the dependent variable: (1) is the standard panel fixed effect regression with no instruments, (2) is a fixed effects regression with happiness and health instrumented by their second lags, (3) shows an IV fixed effects regression using a smoking indicator and happiness lagged twice as instruments, and (4) is the fixed effects IV regression using the bioclimate variables and smoker indicator as instruments. Specifications (3) and (4) are weakly identified as determined by Stock-Yogo (2005) critical values, and are estimated accordingly using Fuller's (1977) LIML with a Fuller parameter of 4. Results for the regional and wave indicator variables are not reported, as they are purely control variables and are not of interest here. The first number is $\hat{\beta}$, and the number in parentheses below is the standard error.

It is easy to see that the sign of the happiness variable changed once the instrumental variables are introduced, suggesting that it is very important to account for the endogeneity of the productivity, health, and happiness variables. This is typically not done in case studies in actual business scenarios, or in the empirical literature. The high significance of health in (1) followed by its insignificance in specification (2)-(4) is probably due to the simultaneity effect of happiness and health in (1), where health takes most of the effect of happiness. This is mitigated in the subsequent regressions through the IV method.

Table 3: Regression Results

Productivity Specification:	(1)	(2)	(3)	(4)
Happiness	0.0036*	-0.0721**	-0.0704**	-0.0918
	(0.002)	(0.028)	(0.028)	(0.168)
Health	0.0072***	-0.0352	0.0378	-0.0327
	(0.002)	(0.034)	(0.119)	(0.182)
Union Membership	0.0057	0.0059	0.0055	0.0077
	(0.006)	(0.009)	(0.009)	(0.009)
Age	0.0332***	0.0151	0.0117	0.0309**
	(0.007)	(0.011)	(0.012)	(0.016)
Age^2	-0.0005	-0.0003***	-0.0003***	-0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)
Social Class	-0.0008	-0.0011	-0.0013	0.0005
	(0.003)	(0.004)	(0.004)	(0.005)
Self Employed	-0.0019	0.0008	0.0053	-0.0080
	(0.037)	(0.052)	(0.053)	(0.045)
Supports Conservative Party	0.0169**	0.0081	0.0073	0.0147
	(0.008)	(0.012)	(0.011)	(0.010)
Supports Labour Party	0.0246***	0.0085	0.0081	0.0251***
	(0.005)	(0.008)	(0.008)	(0.009)
Divorced	0.0036	0.0074	0.0050	0.0271
	(0.018)	(0.023)	(0.025)	(0.035)
Cohabiting or Married	0.0338***	0.0552***	0.0517***	0.0579
	(0.009)	(0.013)	(0.015)	(0.039)
University Education	-0.0035	-0.0087	-0.0050	0.0019
	(0.017)	(0.020)	(0.021)	(0.025)
A/O-Levels	-0.0225	-0.0340	-0.0309	-0.0228
	(0.018)	(0.021)	(0.022)	(0.025)
Centered R^2	0.25	0.16	0.17	0.19
No. obs.	44008	25607	25611	38259

*10% significance, **5% significance, ***1% significance

Specification (2) shows a causal effect of happiness on productivity using an instrumental variable regression in conjunction with the fixed effects estimator. The starred significance is somewhat misleading in the context of specification (2) in the sense that it portrays happiness to be significant at the 5 percent level, when in fact the p-value is 1.10%. Similarly, the p-value of happiness in (3) is 1.13% The point estimate of happiness in (2) is very similar to that of (3), which is robust since the purpose of changing from a lag of self-reported health to a more objective indicator is only to gain a more accurate perspective of the effects of health. The change from (2) to (3) in terms of the sign of health makes the results much more intuitive: it seems surprising that health would negatively affect productivity. The estimate of health is not significant in any of the IV regressions, but the intuitive sign of the estimate along with its consistency across specification (3) and (4) is encouraging.

In specification (3) and (4), the fixed effects estimator is used with a corresponding set of instruments to explain productivity through health and happiness. (3), using two lags of happiness and a smoking indicator as instruments in an exactly identified equation, has an estimate that is not large when considering that the happiness scale is only out of seven.

The significant negative effect of happiness on productivity in (3) may be surprising at first glance, due to a combination of misleading surface level intuition, and results from non-comprehensive research. However, it is not difficult to imagine happier people socializing more at work, or engaging in other non-productive activities, allowing for a negative effect of happiness on productivity that does not require one to ignore theory or common sense. In addition, higher stress levels can lead to lower levels of happiness, but clearly stress often provides motivation to complete tasks, thereby enhancing a labourer's productivity.

The instruments in (4) for health and happiness are mean temperature, mean precipitation, altitude, and a smoking indicator (plus interactions).

Unfortunately, even with the Fuller (1977) LIML estimator, the instruments prove too weak to be able to identify the health and happiness variables well enough to extract informative inference. The validity of the IV's are tested using the Hansen J test, and are accepted under the null with a probability of 71%. While (4) by itself may not be able to provide inference on its own, its comparability to regressions (2) and (3) sheds some positive light: by looking at the standard errors of the happiness estimates in (2) and (3) it can be seen that they are in the range of the corresponding estimate in (4). This suggests that (4) may show a similar result with regard to the effects of happiness on productivity as (2) and (3), but the instruments are too weak to show significance.

The sign of the linear age component suggests that as a person gets older they make more money, and this is to be expected; the negative coefficient of age squared reveals that compensation decelerates with age. It is intuitive, as well as being common in the literature, that wage profiles are increasing in age. Also commonly found, is that being married is positively related to increased compensation, and in this paper that translates to productivity. Union membership exhibits the correct sign according to theory, since the purpose of unions in some sense is to look out for the welfare of its members. As such, one would expect that wages are higher for those with a union membership. This easily extends to the idea that unions should be more likely to accept those workers who are more productive (See Sørensen and Whitta-Jacobsen, 2005).

A variable that has been left out of the regression, that is present in other wage-related literature, is an indicator for whether a respondent's partner is employed. This has been excluded from the regressions presented here because of selection bias. Including a dummy for partner employment restricts the sample to only those people who have a partner, since the variable would have to be set to 'missing' for all respondents for which the questions is not applicable: those without a partner. Including this variable is a mistake in

the other literature.

A possibly surprising result could be the negative sign of social class in (2) and (3). As an explanation, social class' negative coefficients may follow from the concept of diminishing marginal returns to income. Those with relatively higher levels of wealth may be less inclined to work hard, since the marginal benefit of working hard is not high enough to satisfy the required marginal compensation high-wealth workers would like to receive. Those with lower income have more incentive to work harder than the rich. Second, A Levels and O levels are relatively low on the educational spectrum relative to university education, accounting for their negative coefficients. However, the negative coefficient on university education is a surprise: this is most likely not representative of the true value, and it would seem this unrealistic value stems from a large confidence interval. In fact, both education dummies have high p-values, certainly a result of collinearity. This does not affect inference on happiness however, since the collinearity is only between the education indicators. In future research it would be beneficial to construct an objective measure of education based on years of schooling, instead of the categorical variable that is provided in the BHPS. In summary, the education coefficients are not believable given common sense and economic theory, and this can probably be attributed to collinearity, resulting in inaccurate point estimates.

An important result to examine is the R^2 reported for each regression, clearly pointing out the importance of fixed effects. It can be seen that the models' fits are relatively bad, with the best result at only 25%. Since the controls are reasonable and are similar to other papers measuring a similar relationship, it can be concluded that there are sizeable individual effects present. The R^2 does not reflect the fit of these fixed effects, since they are subtracted out in the FE model's equation. If one were to do a regression (although infeasible) explicitly representing the fixed effects using dummy variables, it is expected that the goodness of fit measured by R^2 would be

much better.

A second note is regarding the number of observations listed under N in Table 3. The high number of individual observations reflects a direct benefit of using panel data, in that there is a large amount of variation and individuals present in order to achieve high levels of confidence in the results. The reason regressions (2) and (3) have significantly fewer observations than (1) and (4) are the direct result of using lagged instruments: it causes the first two waves to be inadmissible in the regressions with lower N's.

4.2   Robustness

Robustness checks of these results include the estimation of regressions in Table 3 using robust standard errors instead of bootstrapping, because as Davidson and MacKinnon (1999) note, bootstrapping of simultaneous equations is still the subject of intensive research. The results are in fact very similar, but the significance levels are even higher than in Table 3, with analogous robust SE specifications (2) and (3) reporting happiness at the 5% and 1% level respectively.

The regressions in Table 3 use twice-lagged instruments in specifications (2) and (3) in order to protect against the possibility of short autocorrelation. For robustness the regressions are reproduced but with the endogenous instruments lagged only once, reported in Appendix E. This exercise provides partial support for the regression results in Table 3. First, analogous regression (2) is fully supported, except that the coefficient of happiness is twice as large size in the effect of happiness. (3) shows similar coefficient size, but is not significant. This is a little troublesome, as it is expected that the most recent lagged happiness and health variables have a higher correlation with the current values, and should provide a better instruments under the assumption of the instruments' exogeneity. This leads directly to a possible explanation, that is, the possibility that first order autocorrelation in fact does exist, and the instruments lagged only once suffer from bias

because of it. However, that certainly does not ensure the validity of the twice-lagged instruments, since the regressions in Table 3 were conducted under the assumption of no autocorrelation and therefore instrumental validity - an empirically untestable assumption. It is not even possible to test the validity indirectly, by examining whether the panel errors exhibit auto-correlation. The conjecture that health is positively related to productivity is supported in these regressions, with an insignificant negative result in (2), and a large positive significant result in (3).

The business practice that this paper evaluates is the effectiveness of programs that are intended to make employees happier, such as dress down fridays, as an example. The common operationalization of happiness into job satisfaction is relevant here, since productivity is measured in the context of the workplace. When using workplace satisfaction and the corresponding lagged instrument instead of overall happiness in the productivity specifications, *ceritis paribus*, the results are strikingly similar to the presented regressions in terms of point estimates. The workplace satisfaction variable is significant at the 1 % level for representative regressions (2) and (3), and not significant for (4). The latter should be expected, since the bioclimate IV's should affect general happiness, not happiness at work.

Further results include the stratification of the sample by age and gender, in an attempt to localize the effects of the results. Detailed results can be seen in Appendix C. Consisting of 8 regressions, they suggest that that life satisfaction is significant for both men and women, but very weakly significant for either the oldest or youngest half of respondents. When stratifying on both the age and gender variables, happiness is very significant for the youngest half of female respondents, and very weakly significant for older males. This helps locate the groups that drive the full regression results, although it is clear that additional benefit could derived by further restricting the samples by age for more than 2 quantiles. It is also of course possible to use strata other than gender and age.

4.3 Anent Past and Future Research

In comparison to previous findings on this topic, these results are different in the sense that the effect of happiness is negative. This is important to examine since it is contrary to a large body of previous literature. As discussed in the literature review, a great number papers only measured the significance of correlation between happiness and productivity, which was positive. This paper found the same result, as presented in Table 2. Second, papers that used panel data to eliminate heterogeneity also found positive results. This is also reflected in the data presented in this paper, as can be seen in Table 3, specification (1). The negative effect of happiness is only found when instrumenting the two RHS endogenous variables (health and happiness). This is not found in other research on this topic, and so in fact the results presented in this paper are actually consistent with those found in other literature, but with an added level of sophistication that allows us to see that the underlying effect of happiness on productivity is actually negative.

This paper is clearly not without flaws. Caveats include the assumption that the errors are not correlated across panels, which should be tested empirically with more powerful computers. Second, the use of a proxy variable (compensation) in place of individual productivity, standard though it may be, obviously introduces bias into the regressions. Since it is unrealistic to assume that compensation has a 100% correlation with productivity, the independent variables are being regressed on a dependent variable that does not exactly represent productivity. Third, the lack of knowledge on the speed of compensation's adjustment to an increase in productivity in the UK is only estimated here using evidence from the US.

Future research could look at this speed of adjustment, and further refine the estimations presented in this paper. In fact, related areas for future research are numerous. The IV method is only one way to correct for the endogeneity problem; another is by using a panel-modified version of Zellner's seemingly unrelated regression estimation, which simultaneously estimates

endogenous parameters in a set of endogenous equations. This routine is available for Stata using a random effects model, but not fixed. Another focus for future research is simply examining other geographic regions, in order to provide evidence for or against the external validity of this paper's findings. Similarly, other panels in the UK are necessary to determine whether the unbiased result in this paper is actually near the true parameter, or if its seemingly unintuitive result is unrepresentative of the true population parameter. Also, the finding in this paper may be a product of an over-representation of particular industries where, for example, high stress is common place. It could be argued these types of workers are less happy but more productive.

The results have an impact on the way we view productivity in a broader context that is likely of interest to psychologists, in terms of the determinants of hard work and effort. Important as productivity is in economics, the mechanisms of what makes people productive, whether they are genetic, experiential, or through some other channel(s), should be further examined.

5   Conclusions

The results found in the fixed effects IV specification provide evidence that happiness has a significant negative effect on productivity. Three different IV specifications correct for endogeneity, heterogeneity, and two additionally correct for possible measurement error. Results from (2) and (3) both provide evidence for the conjecture that happiness negatively affects productivity with 5 percent significance, while results from (4) suggest the same conclusions, but suffer from the problem of weak instruments.

Comparisons with previous literature, such as the closely related panel analysis by Graham et al. (2004), reveals several differences that may account for the non-standard and unexpected findings. In particular, this paper uses a panel with many observations and a relatively large time dimension,

increasing inferential capacity. Second, instrumental variables are used to account for endogeneity and measurement error, in part through the use of geographic variables which has previously been unexplored. Third, the analysis in this paper represents a geographic region (the UK) that is stable in a variety of dimensions and relatively transparent, adding to the reliability of individual responses. Despite some drawbacks in the area of IV assumptions, this paper contributes both a new type of analysis using GIS to the existing body of research, as well as deeper econometric techniques providing new results that challenge the findings of other authors with regard to the happy-productive worker thesis.

The primary results of this paper lead to the conclusion that workplace initiatives aiming to improve productivity through the increase of their employees' happiness is misplaced. I do not suggest that it is wrong to keep employees happy, as this is more of a philosophical, humanistic issue, and most would argue that individuals have the right to be happy. However, businesses should explore avenues other than happiness enhancement programs, if their goal is to upgrade productivity levels.

References

ANDREWS, D. AND J. STOCK (2005), "Inference with Weak Instruments", Cowles Foundation Discussion Paper No. 1530.

BAKER, M., C. DERI, AND M. STABILE (2001), "What do Self-Reported, Objective, Measures of Health Measure?", NBER Working Paper No. 8419.

BALTAGI, B. (2008), "Econometric Analysis of Panel Data," 4th ed. New Jersey: Wiley.

BAUM, C. (2010), "xttest3: Stata module to compute Modified Wald statistic for groupwise heteroskedasticity", Boston College Department of Economics. http://ideas.repec.org/c/boc/bocode/s414801.html

Bloom, D. and D. Canning (2005), "Health and Economic Growth: Reconciling the Micro and Macro Evidence", Freeman Spogli Institute for International Economics Working Paper Series.

Bloom, D., D. Canning, and J. Sevilla (2004), "The Effect of Health on Economic Growth: A Production Function Approach", *World Development*, 32(1), 1–13.

Borghesi, S., and A. Vercelli (2008), "Health and happiness: two paradoxes", DEPFID Working Paper No. 1/2008.

Contoyannis, P., and N. Rice (2001), "The impact of health on wages: Evidence from the British Household Panel Survey", *Empirical Economics*, 26, 599–622.

Crépon, B., N. Deniau, and S. Pérez-Duarte (2002), "Wages, Productivity, and Worker Characteristics: A French Perspective", Mimeo, INSEE.

Davidson, R. and J. MacKinnon (1999), "Econometric Theory and Methods", New York: Oxford University Press.

Dearden, L., H. Reed and J. Van Reenen (2006), "The Impact of Training on Productivity and Wages: Evidence from British Panel Data", The Institute for Fiscal Studies Working Paper No. 05/16.

Easterlin, R. (2001), "Income and Happiness: Towards a Unified Theory" *The Economic Journal*, 111, 465–484.

Eaton, B. and M. Eswaran (2005), "Well-Being and Affluence in the Presence of a Veblen Good" *JOURNAL*, 30(4), 591–594.

Feldstein, M. (2008), "Did growth in productivity?", *Journal of policy modeling*, 30(4), 591–594.

Fuller, W. (1977), "Some Properties of a Modification of the Limited Information Estomator", *Econometrica*, 45(4), 939–953.

Gardner, J., and A. Oswald (2001), "Does Money Buy Happiness? A Longitudinal Study Using Data on Windfalls", Mimeo, Warwick University.

Graham, C., A. Eggers, and S. Sukhtankar (2004), "Does Happiness Pay? An Exploration Based on Panel Data from Russia", *Journal of Economic Behavior Organization*, 55, 319–342.

Hagerty, M., and R. Veenhoven (2003). "Wealth and Happiness Revisited: Growing wealth of nations *does* go with greater happiness", *Social Indicators Research*, 64, 1–27.

Hahn, J., J. Hausman, and G. Kuersteiner (2004), "Estimation with Weak Instruments: Accuracy of Higher Order Bias and MSE Approximations", Mimeo, MIT.

Hellerstein, J., D. Neumark, and K. Troske (1999), "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations", *Journal of Labor Economics*, 17(3), 409–446.

Hijmans, R., S. Cameron, J. Parpa, P. Jones, and A. Jarvis (2005), "Very High Resolution Interpolated Climate Surfaces for Global Land Areas", *International Journal of Climatology*, 25, 1965–1978.

Hounkannounon (2008), "Bootstrap for Panel Data", CIREQ, Universite de Montreal.

Hosie, P., N. Forester, and P. Sevastos (2004). "Impact of global pressures on the affective well-being of Australian Managers' Performance", *Research and Practice in Human Resource Management*, 12(1), 134–171.

Lucas, R., E. Diener, and E. Suh (1996), "Discriminant Validity of Well-Being Measures", *Journal of Personality and Social Psychology*, 71(3), 616–628.

Rehdanz, K., and D. Maddison (2003), "Climate and Happiness", Centre for Marine and Climate Research Working Paper No. FNU-20.

Ordinance Survey UK [map]. (2010), "UK Boundary-Line". 1:10 000. Mapping Agency of Great Britain. Available: http://www.ordnancesurvey.co.uk/oswebsite/products/boundary-line/index.html [May 11, 2011].

Oswald, A., E. Proto, and D. Sgroi (2009), "Happiness and Productivity", *JOURNAL*, 30(4), 591–594.

Schaffer, M. (2010), "xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models", Statistical Software Components, Boston College Department of Economics. http://ideas.repec.org/c/boc/bocode/s456501.html.

Shanchez, G. (2011), "Bootstrap with panel data", StataCorp. Retrieved from http://www.stata.com/support/faqs/stat/xt_boot.html on 10/09/2011.

Skirbekk, V. (2003), "Age and Individual Productivity: A Literature Survey", MPIDR Working Paper No. 2003-028.

Sørensen, B. and H. Whitta-Jacobsen (2005), "Introducing Advanced Macroeconomics: Growth and Business Cycles", Berkshire: McGraw-Hill.

Sparber, C. (2010), "Racial Diversity and Macroeconomic Productivity across US States and Cities", *Regional Studies*, 44(1), 71–85.

STOCK, J., AND M. WATSON (2008), "Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regression" *Econometrica*, 76(1), 155–174.

STUTZER, A. (2002), "The role of income aspirations in individual happiness", *Journal of Economic Behaviour Organization*, 54(1), 89–109.

TIAN, G. AND L. YANG (2009), "Theory of negative consumption externalities with applications to the economics of happiness", *Economic Theory*, 39(3), 399–424.

TORRES-RENYA, O. "Panel Data Analysis - Fixed and Random Effects (using Stata 10.x) (ver 4.1), " Data Statistical Services, Princeton University. http://dss.princeton.edu/training/Panel101.pdf.

WARR, P. (1990), "The measurement of well-being and other aspects of mental health", *Journal of Occupational Psychology*, 63, 193-210.

WEIL, D. (2006), "Accounting for the Effect of Health on Economic Growth", Working Paper (Brown University).

WRIGHT, T AND R. CROPANZANO (2004), "The Role of Psychological Well-Being in Job Performance: A Fresh Look at an Age-Old Quest", *Organizational Dynamics*, 33(4), 338–351.

ZELENSKI, J., S. MURPHY, AND D. JENKINS (2008), "The Happy-Productive Worker Thesis Revisited", *Journal of Happiness Studies*, 9(4), 521–537.

Appendices

A EDF of Happiness
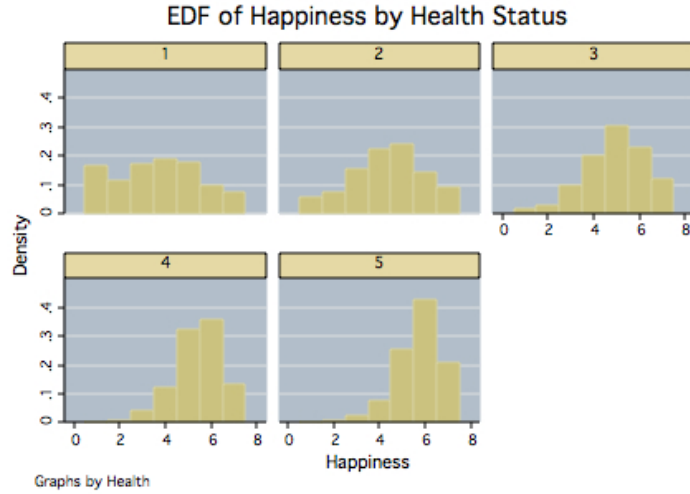
The EDF's show a move towards a skew for being very happy, as health increases. The lowest health status is in Chart 1, and increases to the highest is Chart 5.

B Validity of Lagged Endogenous Instruments

Proof.

A1: $E[\varepsilon_{i,t}|\varepsilon_{i,t-s}] = 0,\ \forall s \neq 0$

The instrument is a lagged endogenous variable, so in general:

$E[\mathbf{Z}_{i,t}|\varepsilon_{i,t}] = E[\mathbf{X}_{i,t-1}^{endog}|\varepsilon_{i,t}]$

Since $y_{i,t-1} = \mathbf{X}_{i,t-1}\beta + \varepsilon_{i,t-1}$, *then:*

$E[\mathbf{X}_{i,t-1}^{endog}|\varepsilon_{i,t}] = E[\frac{1}{\beta}y_{i,t-1} + \frac{1}{\beta}\varepsilon_{i,t-1}|\varepsilon_{i,t}] = \frac{1}{\beta}E[y_{i,t-1}|\varepsilon_{i,t}] + \frac{1}{\beta}E[\varepsilon_{i,t-1}|\varepsilon_{i,t}]$

$= 0 + \frac{1}{\beta}\underbrace{E[\varepsilon_{i,t-1}|\varepsilon_{i,t}]}_{=\ 0\ \text{by A1}}$

It follows that $E[\mathbf{Z}_{i,t}|\varepsilon_{i,t}] = 0$ ∎

C   Stratified Regression Results

Strata	β_{Happy}	β_{Health}	P> $|t_{Happy}|$	SE$_{Happy}$	N
Male	-0.080	0.081	0.064	0.044	11433
Female	-0.081	0.243	0.035	0.038	12340
Young	-0.041	0.159	0.171	0.298	11142
Old	-0.061	0.113	0.218	0.049	12016
Yng males	0.004	-0.187	0.940	0.050	5333
Yng females	-0.124	0.403	0.028	0.056	5786
Old males	-0.156	-0.324	0.114	0.098	5803
Old females	0.001	-0.157	0.991	0.063	6208

D   Utilized Programs

ArcGIS Desktop 9.x: As discussed in the paper, the GIS variables were loaded and processed, including taking the means by LAD's, in ArcGIS. The graphics in Figure 1 were also produced using ArcGIS.

LaTeX: This entire document was typeset using LaTeX, an open source type-setting program. The layout has been modified with the packages: [graphics], [fancyhdr], [booktabs], [setspace], [amsmath], [amsthm], [mathtools], [amssymb], [subfig], and [appendix].

Stata 10: This paper's statistical analysis was conducted using Stata with the additional help of the user-written codes -xtivreg2- and -xttest3-. The user-written commands -eststo- -estout- and -sutex- were used to allow Stata to communicate with Latex.

E   Regressions with Single-Lagged Endogenous IV's

Productivity Specification:	(2)	(3)
Happiness	-0.1507**	-0.1088
	(0.070)	(0.104)
Health	-0.1265	0.1522*
	(0.135)	(0.090)
Union Membership	0.0090	0.0087
	(0.010)	(0.009)
Age	0.0202*	0.0205**
	(0.012)	(0.010)
Age Squared	-0.0004***	-0.0003***
	(0.000)	(0.000)
Social Class	-0.0006	0.0012
	(0.004)	(0.004)
Self Employed	0.0480	0.0070
	(0.047)	(0.040)
Supports Conservative Party	0.0202	0.0072
	(0.013)	(0.010)
Supports Labour Party	0.0220**	0.0267***
	(0.009)	(0.008)
Divorced	0.0152	-0.0033
	(0.033)	(0.028)
Cohabiting or Married	0.0771***	0.0470*
	(0.025)	(0.024)
University Edu	-0.0351	-0.0127
	(0.027)	(0.021)
A/O-Levels	-0.0418	-0.0274
	(0.028)	(0.022)
Centered R^2	-0.01	0.10
F-stat	73.30	96.58
No. obs.	25480	29983

*10% significance, **5% significance, ***1% significance