DISCUSSION PAPER SERIES

IZA DP No. 7531

If You're Happy and You Know It, Clap Your Hands: How Do Mothers and Fathers Really Feel about Child Caregiving?

Rachel Connelly
Jean Kimmel

July 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

If You're Happy and You Know It, Clap Your Hands: How Do Mothers and Fathers Really Feel about Child Caregiving?

Rachel Connelly
Bowdoin College
and IZA

Jean Kimmel
Western Michigan University
and IZA

Discussion Paper No. 7531
July 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7531
July 2013

ABSTRACT

If You're Happy and You Know It, Clap Your Hands: How Do Mothers and Fathers Really Feel about Child Caregiving?

This paper considers the question posed by popular media, do women like doing child care more than men? Using experienced emotions data paired with 24 hour time diaries from the 2010 American Time Use Survey, the paper explores gender differences in how men and women who have done some child caregiving on the previous day feel when engaged in a set of common daily activities. We find that both men and women enjoy their time in child caregiving, men as much, or even more so, than women as evidenced by their average values for happiness, tiredness, and stress, their predicted values for the same three emotions and via an aggregated statistic, the unpleasantness index. Counter-factual unpleasantness indices provide evidence that difference between men and women come almost completely from differences in their experience emotions rather than from differences in how they use their time.

JEL Classification: D13, J13, J16

Keywords: time use, subjective well-being, child care, gender wage gap, experienced emotions, happiness

Corresponding author:

Rachel Connelly
Bowdoin College
Department of Economics
9700 College Station
Brunswick ME 04011
USA
E-mail: connelly@bowdoin.edu

If You're Happy and You Know It, Clap Your Hands: How Do Mothers and Fathers Really Feel about Child Caregiving?

In March of 2012, *The New York Times* website ran the headline, "Do Women Like Child Care More than Men?"[1] The article answered this question in the affirmative, basing its conclusions on an academic paper by Rhoads and Rhoads (2012), published in the *Journal of Social, Evolutionary, and Cultural Psychology*. The NYT's broad promotion of this research finding, despite the authors' own concerns about reliability and relevance,[2] serves to reinforce public perceptions about gendered preferences. The notion that women enjoy child caregiving more than men could serve as a partial explanation for the slowdown in the gender revolution (Coontz, 2013). Substantial gender differences in time spent with children persist despite the fact that young women now attend college at a higher rate than young men and that wages of recent male and female college graduates are now very similar. (Ryan and Siebens, 2012; McDonald and Thornton, 2007) It would be nice to think that women are devoting more time to caring for their children due to personal choice rather than as a result of pressure to adhere to social norms and workplace discrimination; however, we are skeptical. Fortunately, data have become available to address this question more precisely. In 2010, the American Time Use Survey (ATUS) included a special module in which time diary respondents were asked about the strength of five emotions experienced during three randomly choosen activities during their time diary day. Using these emotions data, we look more broadly at the question of whether there is a gender difference in self-reported emotions while engaged in child caregiving? If women "like

[1] This article was published three days later in the New York Times magazine with the title, "Diaper Changing Index." March 25, 2012, p. 16.

[2] Specifically, the research was based on a survey of 184 academic couples with at least one child under the age 2, all on the tenure-track at universities. Additionally, the authors discuss concerns with potential respondent bias inherent in the survey instrument design.

child care more" as the *NYT* reported, then women should report stronger positive emotions and weaker negative emotions while engaged in child caregiving activities. We also use these broad-based data to compare gender differences in child caregiving emotions with emotions during other activities of the day. Finally, we examine specific activities included within the broad "umbrella" category of child caregiving time to explore if all child caregiving activities generate the same emotional responses.

Such an examination is important because both policymakers and the public need broader based and more objective information than provided by Rhoads and Rhoads (2012). Mistaken information about gendered preferences serves to ameliorate concerns about persistent gender wage and achievement gaps. After all, if mothers earn lower wages because they enjoy their time with children more than fathers, then why worry about the resulting wage gap? In addition, to the extent that men also enjoy time spent with their young children (and we find that they do, as much or even more so than women), men would also benefit from institutional and policy changes that allow both parents to take active roles in parenting, while maintaining their strong continuous labor force commitment

We use data from the 2010 ATUS to examine gender differences in emotions during daily activities. Emotions data are not new to the world of rigorous statistical analysis, but the type of the emotions data collected by the ATUS differ substantially from many of the other sources of emotions data available. To date, surveyors have relied on three main types of measures of emotions. The most common are global measure of life satisfaction. These measures result from survey questions like those asked by the General Social Survey, "Taken all together, how would you say things are these days? Would you say that you are very happy,

pretty happy, or not too happy?" This type of measure leads to a broad-based multidimensional assessment of one's total time use.

The second type of emotion measure is known as a general activity judgment measure. Juster and Stafford's time use collection effort (1985) included questions asking about satisfaction with various activities in one's life. The analysis of these data led to the surprising result that individuals report greater enjoyment associated with paid work and time spent with children than most other activities. The Rhoads and Rhoads (2012) survey questions fall under the categorization of general activity judgment measures, however, they only asked about child caregiving activities. One of the problems associated with this type of question is that survey respondents have a sense of how one "should" feel about the activity that might result in conflating emotions with senses of meaning and family responsibility. Spending time with children is meaningful and we "should" enjoy it. After all, most people choose to have children and make tremendous sacrifices for their children. But they don't usually like being awoken at 3 am by a crying baby.

The third type of emotions measure is solicited via questions designed to gauge subjective well-being or experienced emotions. These questions are much more specific than the other two types of measures, asking the respondent about how one felt while doing a specific activity at a specific time. These measures are intended to approximate "process utility," the direct utility resulting from time spent engaged in an activity. Conceptually, process utility is distinct from total utility because the latter includes both process utility and outcome utility (i.e., the utility resulting from consuming the outcome of the activity).[3]

[3] Juster, Courant and Dow (1985) define process benefits as the "direct subjective consequences from engaging in some activities to the exclusion of others. . . . For instance, how much an individual likes or dislikes

Csikszentmihalyi (1990) and Stone and Shiffman (1994) experimented with collecting real time subjective well-being data with the Experience Sampling Method, in which participants carried electronic devices which prompted them several times during the day with questions about what they were doing and how they felt. Unfortunately, this data collection methodology is costly which excludes the possibility of large data sets in which important population subgroups can be analyzed separately (like mothers and fathers of young children). Consequently, Kahneman and associates have been experimenting with the Day Reconstruction Method (DRM) which uses time diaries with one day recall to also collect subjective well-being measures (Kahneman *et al* 2004). They find similar patterns of emotions by time of day and type of activity from the DRM as compared with the Experience Sampling Method. However, when the results of the DRM are compared with the general activity judgment questions of Juster and Stafford (1985), large differences are found (Kahneman and Krueger 2006). Respondents answering DRM questions rank employment lower than the averages from the general activity judgment questions. However, child caregiving activities are still ranked quite high.

In a further effort to reduce respondent burden (and collection costs), Kahnemen and Krueger *et al* conducted the 2006 Princeton Affect and Time Use Survey (PATS) in which emotions data are collected on only three activities in which survey respondents had engaged the previous day. Activities, with the exception of sleep and personal care, were randomly selected in proportion to duration and without replacement (Krueger *et al* 2008). The ATUS subjective well-being module has a very similar design to the PATS survey instrument with the exception that questions are asked about five instead of six emotions: happy, sad, tired, stressed, and pain.

---

the activity 'painting one's house,' in conjunction with the amount of time one spends in painting the house, is as important determinant of well-being independent of how satisfied one feels about having a freshly painted house." (pp.120-121)

Both PATS and ATUS respondents were asked to assign values of zero to six to each emotion for the three selected activities with zero being no emotion and six being a very strong emotion. In the ATUS, the survey design was rotated to modify the order in which the five emotions were considered by the respondent because there was some evidence that subsequent emotion values differed if happy was asked first.

The subjective well-being questions have been shown to yield reliable information about the emotions we experience while engaged in specific activities. Thus, these data are useful for answering the question do women like child care more than men? However, it is worth noting that maximizing process utility is not the sole motivation underlying our time use choices. Very few people enjoy cleaning, but many enjoy the resulting clean home. Similarly, even if we do not like changing diapers, many enjoy parenting young children, which comes inextricably bundled with the need to change diapers.

Section 2: Time Use and Our Experienced Emotions by Gender

The ATUS is an annual nationally-representative time use survey that has been collected by the Bureau of Labor Statistics since 2003. For the purpose of our study, we focus on individuals aged 15 to 85 who report positive minutes of child caregiving on their diary day in 2010. These selection criteria result in an analysis sample of 3,536 individuals.[4] When we report time use information only, we include all individuals in the sample with a positive number of minutes of child caregiving time. When we report on emotions, we include only those individuals who both engaged in that activity and who had that activity chosen as one of their

[4] Because of a lower response rate by men than women and the sample selection criteria that one must have engaged in a child care activity on diary day the sample includes 33 percent men and 67 percent women. However, all results reported in the paper are weighted using BLS supplied weights in order to return the sample to the population proportions.

three "emotion" activities. This latter requirement reduces the overall sample size making it impossible to analyze the emotions for some activities that occur too infrequently to produce reliable sample sizes. Six specific child caregiving activities are included in our analyses: the physical care of household children, playing with household children, talking with household children, picking up or dropping off household children, other caring for and helping household children, and activities related to household children's education. We include 13 other frequently occurring activities, in addition to these six child caregiving activities.[5]

It is not surprising that men and women's (all of whom engaged in positive child caregiving on the diary day) time allocation choices differ substantively in 2010, despite the gender revolution of the last fifty years, the convergence of educational attainment and the substantial increase in women's paid employment. Table 1 shows time use in minutes and the proportion of included time (the sum of the time spent in the 19 included activities) and the proportion of total emotion-eligible time (excludes sleep and personal care) spent in each activity category for men and women who have any child caregiving time on diary day. Women spend more time than men in four of the six child caregiving categories with equal time in the "playing with children" category and in the "residual care for and helping household children" category.

[5] The 2010 ATUS included 476 different possible time use activities which are categorized into 17 "2 digit" categories and within each "2 digit" category there are sub categories called "4 digit" categories and then further sub categories called "6 digit" categories. In all there are 21 "6 digit" categories related to child caregiving of children living in the household (excluding all transportation categories). However, many of these categories seldom occur. We only include an activity if there are at least 30 subjective well-being reports in each of the men's and women's sample. When a "6 digit" activity did not pass this threshold, we combined it with other similar activities (within the same "4 digit" classification) and created a "4 digit residual" activity. Again, we subjected that "4 digit residual" to the 30 observation criteria for inclusion. Six child caregiving activities pass the "test" and are included in our analysis. Of these, four are "6 digit" activities, the physical care of household children, playing with household children, talking with household children, and picking up or dropping off household children. The other two are "4 digit residual" categories, the rest of "caring for and helping household children" and all of "activities related to household children's education." To compare child caregiving activities to other time use activities we used the same inclusion rule. In all we have 19 activities (excluding transportation categories) which are either "6 digit" activities or "4 digit residual" activities.

They also spend more time on interior housekeeping and shopping, less time in paid employment, and less time watching television. Men and women devote equal minutes to financial management and to many leisure activity categories, with women spending somewhat more time reading and relaxing alone and men spending more time in physical activity.

There are several reasons why men and women may use their times differently; they may have different opportunity costs of their time due to different marginal wage rates; they may have differing productivities due to learning by doing, education or inherent differences between men and women; or they may have different preferences. In addition, productivities in one task may be a function of other tasks performed. For example, if a parent is home caring for children then he or she is more productive in interior housework tasks which must be done at home and can be done in small chunks in between child caregiving tasks. Since most employment hours take place outside the house, those with more time in employment will be less productive on average in household tasks done at home. This bundling of tasks based on location and/or inherent interruptability seems like an important but not well understood piece of gender differences in time use.[6] Finally, habit may play an important role. Schober (2012) shows that in Germany, women who withdraw from the labor market when their children are young, continue to do more of the household chores even after they return to the labor market. The point of this long list is that certainly not all the differences we observe in Table 1 are the result of preferences. Yet it is still interesting to directly consider differences in how men and women feel while engaged in an activity as process utility certainly affects total utility for both men and women.

[6] Hamermesh and Lee (2007) note that women report a greater sense of feeling rushed and this gender difference may be the result of women changing activities more often throughout the diary day.

The emotions reported in the ATUS data, as recalled one day later, likely reflect the average emotion experienced during an activity. Lab experiments have shown that average reported emotion may be different from average experienced emotion since respondents tend to remember the end of the episode more vividly than the middle and also overweight the emotional peaks and troughs (Kahneman, Fredrickson, Schreiber and Redelmeier, 1993). Thus, we may be concerned about the effect diminishing marginal utility plays on reported emotions. Since women perform more child caregiving than men, their marginal happiness at the end of a period of caregiving may be lower and their reported negative emotions may be higher. (Oster, 2013) Using OLS regression analysis, we can control for the total amount of time in a single activity episode and also the total amount of time spent in the same activity on diary day.[7] On the other hand, since to some extent, time use data reveal choices made by the respondents, we would expect that people who do more of something may simply enjoy it more. This would lead to the prediction that women would like child caregiving more, simply because we observe them doing more of it. As a result, a finding based on more rigorous analysis that women like child caregiving more than men will be inconclusive, but alternatively, a finding that men like it more or that men and women like it equally provides evidence against the hypothesis that women like child care more than men.

Table 2 presents the average happiness scores for respondents who had positive minutes of child caregiving on their diary day. Both men and women report high levels of happiness when engaged in child caregiving, higher than most other activities. Employment provides a

[7] The regression equation also includes sex, the person's age, age squared, the day of the week, the month of the year, and the starting time of the activity in intervals.

much lower level of happiness than child caregiving. These results are different from Juster's (1985) previous findings, probably because of the different type of questions being asked. The average happiness level during employment is statistically equivalent for men and women. On the other hand, men have statistically significantly higher happiness scores than women for all categories of child caregiving except physical care.[8] The regression results reported in the last column of Table 2 show that, even controlling for other potential influences on the happiness score including the time of day, the day of the week, month of the year and duration of the activity, women report lower happiness scores for the categories of playing with children, talking to children and the residual caring category. Their average score for the aggregated category of child caregiving activities is also lower than men's. These results dispel some of the alternative explanations for gendered child caregiving preferences, such as "men interact with children at better times during the day" or "men interact with children for fewer minutes a day." Instead, the results provide evidence that while men and women who engage in active child caregiving some time during the day report being equally happy when all activities are aggregated and in most of the more specific included activities in Table 2, men report being happier while engaged in child caregiving.

As discussed above, the subjective well-being data in the ATUS allow us to look beyond just happiness. Table 3 presents results for tired and stressed. Women report being more tired than men in most activities, including the aggregate child caregiving activity, as well as the specific categories of physical care of children and picking up and dropping off children (based on simple t-tests of the weighted averages). Previous research by Hamermesh (2007) and the Pew Research Center (2006) provided early evidence of this gender difference, noting that

[8] The difference is statistically significant using a simple t-test for the category "playing with children" and for the aggregated category of all child caregiving activities.

women are considerably more likely than men to report "feeling rushed" in their daily activities. Regression results show women more tired in many categories of activities including playing with children, talking with children and picking up and dropping off children. The regression results provide evidence that women's extra tiredness is not simply a function of duration of activity nor time of day when the activity is performed.

There are fewer differences between men and women in reported stress than tiredness, but in the cases where differences are observed, it is always the case that women report higher stress than men. The activities with the highest levels of stress are financial management for women, working one's main job for both men and women and grocery shopping. Child caregiving activities have lower levels of stress, except for education-related caregiving for women. T-test results show that women experience statistically significantly more stress than men while engaged in education-related caregiving and that finding is robust to controlling for many of the characteristics of the activity including the duration of the activity and the time of day. In the aggregate child caregiving category, again women are statistically significantly more stressed than men, as is shown in both the t-test and from the regression result.

Section 3: Aggregating the Subjective Well-Being Information—An Unpleasantness Index

While there is some correlation between the five emotions, with happy negatively correlated with sadness, pain, tired and stressed, and tired and stressed positively correlated, these correlations are not strong. If we consider all included activities with reported emotions, the correlation for happy and tired is -.13 for men and -.16 for women. Looking just at the child care activities, the correlation between happy and tired is -.12 for both men and women.[9] Table

[9] The correlations of average emotion scores by activity weighted by activity durations are much higher. The correlation of the average happy and tired emotion in 19 included activities is -.21 for men and -.58 for women. The correlation is -.86 for men and -.79 for women for happy and stressed.

4 shows the full correlation matrix for all included activities in panel A and for the six child care activities for panel B.

The relative weakness of the correlations implies that we should consider all emotional responses to an activity instead of relying on a single emotion gauge (say, happiness) for assessing process utility. Krueger (2007) suggests two methods for aggregating the multiple emotion information: cluster analysis and the creation of an unpleasantness index (the so-called U-index). In this section, we consider first cluster analysis and then the U-index separately for men and women, still with a focus on child caregiving activities. Our goal is to provide a more definitive answer to the question of whether women like child caregiving more than men.

Section 3.1: Cluster Analysis

Cluster analysis is a descriptive statistical technique for combining activities with multiple attributes into groups that are similar in their pattern of attributes. For example, we might consider 150 different automobile models which differ on margins of safety, reliability, styling, speed, miles per gallon, price, and size and choose to identify some small number (say, five or six) of groups of cars. The cars within each group would be similar to each other in all dimensions and differences across groups would again be in a combination of dimensions. Various statistical methods are available for performing cluster analysis. Krueger (2007) uses the weighted emotion averages by activity and the K-means method to create six clusters. We experimented with that method but rejected it because the resulting groupings were not robust to the inclusion/exclusion of a single activity nor to changes in the number of clusters. Instead, we found that using a hierarchal cluster method with Ward linkage which allows the data (and the researcher as observer) to decide on the best number of clusters produced more stable results.

Five clusters provide a reasonable fit for the data for both men and women.[10] Table 5 presents means for the five emotions for each of the five clusters for both men and women. We have renumbered the clusters so that lower numbers are associated with lower utility levels. This is a somewhat arbitrary ranking because for both men and women, there were two clusters that were difficult to rank; clusters 2 and 3 for men and clusters 3 and 4 for women.[11] Cluster analysis is useful for bringing activities together, but it does not provide any sure fire way of rank ordering total well being among the clusters.

It is also important to note that we cannot compare cluster numbers between men and women. Women's two lowest clusters, if merged together, would have the same happiness average as men's cluster 1. By looking at the composition of the clusters and our rank ordering, we can conclude that both men and women rank working one's main job quite low, similar to interior cleaning and grocery shopping. Similarly, men and women agree that playing with children and sports and exercise are among the most enjoyable categories.

Large differences exist between men and women in their assessment of financial management tasks and relaxing alone. Women rank both of these activities much lower overall (i.e., they appear in a lower ranked cluster) than men. Shopping (not grocery shopping) and socializing with others are two activities that men rank lower overall than women.

---

[10] Like Kruger (2007) our cluster analysis was performed on a sample of activity emotion averages weighted by relative frequency of time duration on diary day.

[11] For men, clusters 2 and 3 have very similar level of average happiness but the activities in cluster 2 are more stressful than the activities in cluster 3. For men, Food Preparation, Shopping, Socializing with Others are more stressful activities than Kitchen Cleanup, Physical Care of Children, Talking with Children, Relaxing/Doing Nothing Alone, Watching TV, Rest of Relaxing and Leisure, especially if one also took the tiredness average into account. For women, clusters 3 and 4 had a similar ambiguity as cluster 3 contains activities that are less stressful on average than those in cluster 4. But in this case the happiness averages are more different, with the activities in cluster 4 producing a meaningfully higher level of happiness than cluster 3, causing us to rank the activities of cluster 4 as higher utility than cluster three.

It is interesting to note that for both men and women, the six child caregiving activities are scattered across three clusters. Table 6 presents the same information arranged differently to make this type of comparison easier. Physical care of children is in the same category as watching TV (not our favorite leisure time activity even though we spend a substantial amount of time doing it each day). Playing with children is always ranked in the highest category. The other activities differ between men and women, for reasons we have already observed in Tables 2 and 3, that is, they produce differing levels of happiness, stress and tiredness.

Section 3.2: The Unpleasantness Index

Kahnemen and Krueger (2006) suggest an unpleasantness index (the U-Index) as an alternative to cluster analysis to aggregate the information from the multiple measures of emotions for each activity. Given the nonrobustness of cluster analysis, the many choices left up to the discretion of the researcher in terms of the "right" number of clusters, the method used in clustering and the agnosticism of the method in terms of the rank ordering of the clusters, the U-index is a welcome alternative.

The U-index, calculated for each individual, measures the percent of time one is engaged in unpleasant activities. There are a number of ways that one could categorize an activity as pleasant or unpleasant. One must choose which emotions to include and whether to use averages or individual scores.[12] Kahneman and Krueger's U-index defines an activity as unpleasant if any of the individual's negative emotion scores in that activity are higher (stronger emotion) than any of the individual's positive emotion scores. In the PATS data, there are two positive emotions and four negative ones. We use the same approach with the ATUS data, but there is only one

[12] Bertrand (2013) using the same ATUS data includes meaningfulness as a positive emotion and excludes tiredness. We exclude meaningfulness as it has very low variance and we consider tiredness an important negative emotion on which men and women differ.

positive emotion (happiness). We include all four negative emotions: sad, tired, stressed and pain. Once an activity is defined as unpleasant for the individual, the indicator value is multiplied by the duration spent in the activity then averaged over all individuals with time spent in that activity to calculate the average activity-level U-index. This measure reflects the average percent of time in the activity that people characterized as unpleasant. These average activity-level U-indices are then used along with an individual's daily time use to calculate a person-level U-index which is interpretable as the percent of one's time (in the included activities) that is unpleasant.

The construction of the unpleasantness index addresses concerns with individual differences in scoring. It is also possible that groups of people (men versus women) could have systematic differences in the relative "generosity" of their scores. For example, it could be the case that women are simply more emotive than men, so that they are both happier and sadder, more tired and more stressed. The unpleasantness index is designed to alleviate this problem by making within-person assessments instead of across person assessments.

Table 7 presents the U-indices calculated separately by gender. Row 1 reports average person-level U-indices that includes all 19 included activities. Row 2 reports average person-level U-indices for the six child caregiving activities. They measure the average percent of child caregiving time that men and women report as unpleasant. Both row 1 and 2 reveal statistically significant gender differences, with women reporting a greater proportion of their time as unpleasant. Men report time spent in child caregiving is unpleasant about 10 percent of the time while women report that their child caregiving is unpleasant 19 percent of the time. This is further evidence that women don't "like" child care activities more than men. Overall, survey

respondents consider child care activities to be quite pleasant, as the average person-level U-index for all 19 included activities is 21 percent for men and 24.5 percent for women.

Rows 3-21 of Table 7 present the average activity-level U-indices by gender. Comparing across activities, we find that men report greater unpleasantness in some child caregiving activities, but the largest gender differences are in activities that women find unpleasant more often than men. We have already shown that playing with children is an activity that makes most people happy. Men report slightly higher level of unpleasantness in that activity than women, but both activity-level U-indices are among the lowest for all included activities, 5 percent for men and 4.2 percent for women. Talking with children is also an activity which men report more often unpleasant than women. However, women report noticeably greater unpleasantness for physical care of children, picking up or dropping off and the residual caregiving category. There is also a small difference in education-related caregiving, but in the direction of women more likely to report as unpleasant.

Overall, the activity-level U-indices show that men and women are quite different in their assessment of activities. These results bring into question Krueger's (2007) strategy of using a gender neutral U-index to examine changing levels of well-being for men and women over time. It seems that analyzing the evolution of well-being over time may depend on whether we are using men or women's emotions' measures.

The U-index can also be used to produce a "gender counter-factual" which allows us to disentangle the source of overall U-indices' gender difference between differences in time use

choices versus differences in self-reported subjective well-being.[13] Results for these counter-factuals are shown in Table 8. The results vary only slightly when the reference group (i.e., which gender's actual time use is used as the weights) is varied. Table 8 shows both possibilities, with the first two rows using women's time use proportions as weights (women are the reference group) and rows 3 and 4 using men's time use as the weights (men are the reference group). The results of this exercise provide strong evidence that gender differences in the average person-level U-indices are the result of women's stronger negative emotions in most activities rather than because women devote more time to inherently unpleasant activities. If men used their time the way women use theirs, they would be slightly better off as their average person-level U-index would be 20.9 instead of 22.1 (representing a smaller percentage of time experienced as unpleasant) and their child caregiving only person-level U-index would be essentially unchanged (10.0 instead of 9.9.) Thus, more than 100 percent of the gender gap in the actual person-level U-indices is attributable to gender differences in the activity-level U-indices and close to 100 percent for the child caregiving U-index gender gap. These counter-factuals allow us to conclude that, even controlling for differences in how women and men use their time, women find child caregiving more unpleasant than men.

Section 4: Concluding Remarks

The newly available 2010 American Time Use Survey data, which provides time diary data along with self-reported measures of subjective well-being for a large, nationally representative sample, offer new insights into variations in process utility across activities and

[13] This is analogous to a Oaxaca decomposition of the wage gap which decomposes differences attributable to composition effects from behavioral effects. The composition effects in this case as differences in how time is used, the behavioral effect is how unpleasant activities are.

individuals. Researchers have suggested a variety of uses for these data, including a new type of national time accounting (Krueger *et al*, 2008). Our use of the data is less grand, but we believe, equally important, as it provides empirical evidence to test the hypothesis that women "like" child caregiving time more than men, with the unstated corollary that differences in preferences help explain why women, even women in couples in which both spouses hold fulltime paid jobs, still do the majority of child caregiving. Policy-makers have an interest in understanding this persistent gender difference because of the established link between unpaid work in the home and average labor market earnings (Hersch 2009). In other words, women perform more unpaid work in the home and this behavior has been linked causally with the gender wage gap. Additionally, recent research has shown that both fathers and mothers report concerns with balancing work and family, suggesting that this topic is not purely a "women's issue." (Parker and Wang, 2013)

Our research shows that both men and women "like" child caregiving in the sense that they report high levels of happiness while engaged in child caregiving than in other daily activities. However, while engaged in caring for one's own children, men report even higher happy scores than women and their happiness scores remain higher even after we control (via regression) for time in the activity, the timing of the activity in day, week and year, and age. Looking at the six sub-categories of child caring for which we have large enough sample sizes, we find that there are substantial differences among the child care activities in terms of happiness scores. Playing with children makes us very happy; picking up and dropping off children, the physical care of children and education-related child caregiving have lower happiness scores.

Women and men also differ in the tiredness and stress they report while engaged in child caregiving. Women report higher tired scores in almost every activity, including some of the

child care sub-categories and in the aggregated child care activity category. Similarly, women report higher levels of stress in the aggregated child care activity category than men, a result that remains even after we control for the duration in the activity.

One of the challenges that researchers face with the new ATUS emotions data is that there is a lot of information to process: five emotions, many activities. In addition, there is substantial independent variation in the emotions, such that high happy scores do not tell a complete utility story. We rely on two distinct methods used to date in this literature to collapse the multiple measures of well-being and time use activities. The results from both endeavors, cluster analysis and the unpleasantness index continue to cast substantial doubt on the hypothesis that women "like" child caregiving more than men.

One of the main findings of the cluster analysis is that the traditional sense that housework is unpleasant and leisure activities are pleasant is too simplistic; reality is more nuanced. Employment and most unpaid housework activities are grouped in the two or three lowest happiness clusters, but for both men and women there are a group of activities with higher happiness but higher stress levels that are difficult to rank order. Watching TV ranks much lower than many of the other leisure activities (yet both men and women spend a lot of time doing it, men more than women). The child caregiving activities are not all in the same cluster, but rather are scattered among the top three ranked categories for both men and women. We also found that men's and women's clusters are substantially different and it is not correct to compare the third happiest cluster for men with the third happiest cluster for women because the first and second clusters for women have an average weighted happiness score equal to the lowest category for men. These results suggest that using a gender neutral clustering as proposed by Krueger (2007) is problematic.

The aggregation measure preferred by Kahneman and Krueger (2006), the U-index, aggregates all emotions within a given individual/activity. An individual is said to rate an activity as unpleasant when any one of his or her negative emotions scores is higher than his or her highest positive emotion score. Since the ATUS has only one positive emotion, an activity is categorized as unpleasant if any of the other emotion scores are higher than the happy score. Most activities for both men and women are classified using this method as pleasant, but 17 percent of men's activities and 24 percent of women's activities (among the sample of men and women who have positive minutes of child caregiving time on the diary day) are classified as unpleasant. The U-index then weights the activity by the duration of the activity, such that we can say that men who have positive minutes of child caregiving time on the diary day are engaged in unpleasant activities 21 percent of the time compared to 25 percent of the time for women.

Using the U-index we find that child caregiving activities are more often unpleasant for women, with a substantial difference of 9 percentage points. Using activity-specific U-indexes, not all child caregiving activities are more unpleasant for women than men. Playing with children and talking to children are unpleasant a smaller percentage of the time for women, but the differences between men and women are quite small. The other four sub-categories of child caregiving have a larger percentage of unpleasant time for women, especially the categories of physical care, picking up and dropping off children, and the residual caring for children category.

The analysis in this paper provides strong evidence against the popular notion that women perform more unpaid work in the home, particularly activities related to caring for their own children, because they enjoy these activities more than men. Unfortunately, we are not closer to answering the big picture question, why do women perform substantially more of the

child caregiving (and housework) than men. They could still receive more total utility for these endeavors (with the subjective well-being data, we are only "observing" process utility), they may be (real or perceived to be) more productive caregivers, or they may be the logical choice in a heterosexual couple given lower wages and labor market discrimination, particularly at higher managerial levels, that women still face. But what we have been able to do is rule out the hypothesis that women simply "like" child caregiving more. Instead, we find that men like child caregiving as well, in fact, even more than women, but there are differences across the specific caregiving activities. Playing with children is at the top of our lists; picking up and dropping off is near the bottom. Education-related child caregiving is an activity that women and men disagree on the most, with women finding it stressful much of the time.

Finally, this research, simply by delving more deeply into the specifics of the emotional well-being of individuals in the United States, serves to enlighten discussions concerning happiness, a popular subject in our nation's media for generations. Most recently, the cover of *Time Magazine* (July 8, 2013) highlighted the "pursuit of happiness." Based on our findings, researchers looking at happiness or, more broadly, at experienced emotions should consider the role gender plays in their analyses.

Papers Cited

Bertrand, Marianne. 2013. “Career, Family, and the Well-Being of College-Educated Women,” *American Economic Review* 103, 244–250.

Coontz, Stephanie. 2013. “Why Gender Equality Stalled?” *New York Times*, February 16, 2013.

Csikszentmihalyi, Mihaly. 1990. *Flow: The Psychology of Optimal Experience*. HarperCollins, New York.

Hamermesh, D. and J. Lee. 2007. “Stressed Out on Four Continents: Time Crunch or Yuppie Kvetch,” *Review of Economics and Statistics* (May), pp. 374-83.

Hersch, Joni. 2009. “Home Production and Wages: Evidence from the American Time Use Survey,” *Review of Economics of the Household* 7, 159-78.

Juster, F. Thomas. 1985. "Preferences for Work and Leisure." In *Time, Goods, and Well-Being*, edited by Juster F. Thomas and Frank P. Stafford, 335–51. Ann Arbor: University of Michigan Press.

Juster, F. T., P. Courant and G.K. Dow. 1985. “A Conceptual Framework for the Analysis of Time Allocation Data,” in F. Thomas Juster and Frank P. Stafford, (eds.), *Time Goods and Well-Being*. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor.

Juster, F. Thomas and Frank Stafford. 1985, editors, *Time, Goods, and Well-Being*, Ann Arbor: University of Michigan Press.

Kahneman, Daniel, Barbara Fredrickson, Charles A. Schreiber, and Donald Redelmeier. 1993. “When More Pain is Preferred to Less: Adding a Better End.” *Psychological Science* 4(6),401–05.

Kahneman, Daniel, and Alan Krueger. 2006. "Developments in the Measurement of Subjective Well-Being." *Journal of Economic Perspectives* 20(1), 3–24.

Kahneman, Daniel, Alan Krueger, David Schkade, Norbert Schwarz, Arthur Stone. 2004. "A survey method for characterizing daily life experience: the Day Reconstruction Method." *Science* 306: 1776-80.

Krueger, Alan. 2007. "Are We Having More Fun Yet? Categorizing and Evaluating Changes in Time Allocation." *Brookings Papers on Economic Activity* 2, 193–215.

Krueger, Alan, Daniel Kahneman, David Schkade, Norbert Schwarz, Arthur Stone. 2009. "National Time Accounting: The Currency of Life," in Alan Krueger, Ed. *Measuring the Subjective Well-Being of Nations: National Accounts of Time Use and Well-Being* (Univ. of Chicago Press, Chicago), 9-86.

McDonald, Judith, and Robert Thornton. 2007. Do New Male and Female College Graduates Receive Unequal Pay? *Journal of Human Resources* 42(1), 32–48.

Oster, Emily. 2013. "Why You Shouldn't Work Less, Even if You Prefer Spending Time With Your Family." *Slate Magazine*, January 8. http://www.slate.com/articles/double_x/doublex/2013/01/work_life_crunch_why_you_shouldn_t_spend_fewer_hours_at_work.html.

Parker-Pope, Tara, "Do Women Like Child Care More than Men?" http://well.blogs.nytimes.com/2012/03/22/do-women-like-child-care-more-than-men/?emc=eta1 also published as "Diaper Changing Index," New York Times Magazine, March 25, 2012, 16.

Parker, Kim and Wendy Wang. 2013. "Modern Parenthood: Roles of Moms and Dads Converge as They Balance Work and Family." Pew Research Center (March), 68 pgs.

Pew Research Center. 2006. "Who's Feeling Rushed? (Hint: Ask a Working Mom)," *Social Trends Report* (February 28), 8 pgs.

Rhoads, Steven and Christopher Rhoads. 2012. "Gender Roles and Infant/Toddler Care: Male and Female Professors on the Tenure Track," *Journal of Social, Evolutionary and Cultural Psychology* 6(1), 13-31.

Ryan, Camille and Julie Siebens, 2012 .Educational Attainment in the United States: 2009," *Current Population Reports*, P20-566 (February).

Schober, Pia. 2012. "Time to Care (Longer): Maternal Labor Market Interruptions and the Gender Division of Domestic Work." Working paper presented at PAA meetings, San Francisco.

Table 1: Time in 19 Largest Time Use Categories

	Percent of time in included categories		Sig diff?	Percent of time in all swb categories		Sig diff?	Total minutes on diary day		Sig diff?
	Men	Women		Men	Women		Men	Women	
Interior cleaning	1.9%	5.7%	***	1.4%	4.5%	***	13	39	***
Food prep	3.3%	7.2%	***	2.5%	5.5%	***	22	47	***
Kitchen cleanup	0.7%	2.1%	***	0.5%	1.6%	***	5	14	***
Financial management	0.9%	1.1%		0.6%	0.8%		5	7	
Physical care of children	4.1%	7.7%	***	3.0%	5.9%	***	27	51	***
Playing with children	4.5%	3.8%		3.5%	3.1%		31	27	*
Talking with children	0.5%	0.9%	***	0.4%	0.6%	***	3	6	***
Picking up or dropping off children	0.6%	1.0%	***	0.4%	0.7%	***	4	6	***
Rest of caring for children	3.2%	3.4%		2.4%	2.6%		21	22	
Education-related child care	1.1%	1.7%	***	0.8%	1.4%	***	7	12	***
Working main job	34.2%	21.0%	***	29.4%	17.9%	***	276	162	***
Grocery shopping	0.9%	1.4%	***	0.7%	1.0%	***	6	9	***
Shopping not groceries	1.7%	3.0%	***	1.2%	2.2%	***	11	19	***
Eating/drinking	10.1%	9.7%		7.4%	7.2%		66	61	***
Socializing with others	5.1%	5.7%		3.7%	4.2%		34	36	
Relaxing/doing nothing alone	1.2%	1.9%	***	0.9%	1.5%	***	9	14	***
Watching TV	18.7%	16.8%	**	14.5%	13.4%		127	112	***
Rest of relaxing and leisure	4.2%	4.3%		3.1%	3.3%		29	28	
All of participating in sports, exercise and recreation	3.1%	1.4%	***	2.2%	1.1%	***	20	10	***

Notes: Column 1 and 2 use the total time spent in the 19 included activities as the denominator, while columns 3 and 4 use the total time spent in all subjective well-being eligible activities (essentially all activities except sleeping and personal care time) as the denominator. These latter percentages can be interpreted as the percent of awake time. All samples limited to men and women who reported some child caregiving minutes during their diary day. Significance column is based on simple t-tests that control for sample weights. * 10% significance level, ** 5% significance level, *** 1% significance level.

Table 2: Average Happiness Scores by Activity

	Men	Women	t-test sig diff?	regression result women sig diff?	
Interior cleaning	3.656	4.076			
Food prep	4.514	4.536			
Kitchen cleanup	4.226	3.752		(-)	less happy
Financial management	4.180	2.640	**		
Physical care of children	4.561	4.659			
Playing with children	5.596	5.385	*	(-)	less happy
Talking with children	5.104	4.760		(-)	less happy
Picking up or dropping off children	4.644	4.349			
Rest of caring for children	5.314	4.948		(-)	less happy
Education-related child caregiving	4.591	4.455			
Working main job	3.753	4.004		(+)	more happy
Grocery shopping	3.000	4.068	***	(+)	more happy
Shopping not groceries	4.161	4.794	*	(+)	more happy
Eating/drinking	4.846	4.815			
Socializing with others	4.683	4.778			
Relaxing/doing nothing alone	4.590	3.673	*		
Watching TV	4.489	4.301			
Rest of relaxing and leisure	4.429	4.175			
All of participating in sports, exercise and recreation	5.361	5.356			
All 19 included activities	4.319	4.386			
All 6 included child caregiving activities	5.104	4.869	**	(-)	less happy

Notes: All samples are restricted to those men and women who had some minutes of child caregiving during their diary day. Subjective well-being questions are only asked of respondents who participated in that activity with the probability of being asked about the activity a function of the duration of the activity. The sampling of activities was done without replacement. The last column represents the significance of the gender coefficient (associated with a female 0-1 indicator variable in an OLS regression that also included the time of day of the activity in intervals, the day of the week, month of the year, the duration of the activity questioned, the total duration of that activity on diary day, age and age squared. The regression, t-tests, and averages are weighted by activity duration weights following the methodology provided by the Bureau of Labor Statistics. * 10% significance level, ** 5% significance level, *** 1% significance level. Regression results reported at the 5% significance level.

Table 3: Average Tired and Stress Scores by Activity and Gender

	Tiredness					Stress				
	Men	Women	sig diff?	regression result women sig diff?		Men	Women	sig diff?	regression result women sig diff?	
Interior cleaning	2.325	2.579				2.575	1.831			
Food prep	1.948	2.723	***	(+)	more tired	1.333	1.416			
Kitchen cleanup	2.704	2.908		(+)	more tired	0.721	1.567	**	(+)	more stressed
Financial management	1.778	2.987	**			0.767	3.046	***	(+)	more stressed
Physical care of children	2.503	2.999	**			1.204	1.281		(+)	more stressed
Playing with children	1.851	2.347		(+)	more tired	0.729	0.793			
Talking with children	2.538	2.346		(+)	more tired	1.353	1.428			
Picking up or dropping off children	1.237	2.660	***	(+)	more tired	1.039	1.485			
Rest of caring for children	1.851	1.995				1.284	1.023			
Education-related child caregiving	1.832	2.282				1.532	2.385	**	(+)	more stressed
Working main job	2.095	2.745	***	(+)	more tired	2.420	2.468			
Grocery shopping	1.747	2.416		(+)	more tired	2.354	2.094			
Shopping not groceries	2.065	2.025				1.616	1.787			
Eating/drinking	1.942	2.284	*	(+)	more tired	1.164	1.333			
Socializing with others	1.662	2.402	**	(+)	more tired	1.138	1.421			
Relaxing/doing nothing alone	2.334	3.567	**	(+)	more tired	1.392	2.267			
Watching TV	2.508	2.854				0.753	1.318	**	(+)	more stressed
Rest of relaxing and leisure	2.960	2.389				1.131	1.231		(+)	more stressed
All of participating in sports, exercise and recreation	1.420	2.570	**	(+)	more tired	0.576	1.020		(+)	more stressed
All 19 included activities	2.113	2.618	***	(+)	more tired	1.638	1.696		(+)	more stressed
All 6 included child caregiving activities	2.026	2.520	***	(+)	more tired	1.091	1.203		(+)	more stressed

Note: Same as Table 2

Table 4: Correlation Matricies

All included activities:

Women

	happy	sad	pain	stress	tired
happy	1				
sad	-0.3255	1			
pain	-0.1311	0.3593	1		
stress	-0.329	0.4277	0.298	1	
tired	-0.1653	0.1975	0.3002	0.3283	1

n=4322

Men

	happy	sad	pain	stress	tired
happy	1				
sad	-0.2525	1			
pain	-0.0995	0.317	1		
stress	-0.3327	0.4199	0.2656	1	
tired	-0.1366	0.1737	0.2535	0.2823	1

n=2092

Child caregiving activities only:

Women

	happy	sad	pain	stress	tired
happy	1				
sad	-0.3344	1			
pain	-0.1046	0.2939	1		
stress	-0.3716	0.3941	0.2535	1	
tired	-0.1207	0.1386	0.3164	0.2715	1

n=1306

Men

	happy	sad	pain	stress	tired
happy	1				
sad	-0.2456	1			
pain	-0.0261	0.2377	1		
stress	-0.3566	0.388	0.1593	1	
tired	-0.1208	0.0945	0.1832	0.238	1

n=587

Table 5: Average Emotion Scores by Gender-Differentiated Clusters

Cluster Number	Happy	Sad	Pain	Stress	Tired
Men					
1	3.723	0.597	0.615	2.425	2.095
2	4.545	0.672	1.170	1.276	1.815
3	4.505	0.366	0.799	0.955	2.589
4	4.890	0.270	0.707	1.201	1.876
5	5.537	0.176	0.753	0.690	1.742
Women					
1	3.230	1.670	1.807	2.601	3.318
2	4.009	0.688	0.842	2.276	2.707
3	4.422	0.523	0.734	1.323	2.791
4	4.773	0.393	0.785	1.506	2.282
5	5.183	0.222	0.613	0.924	2.214

Activities in Clusters	
Men	
1	Interior Cleaning, Working Main Job, Grocery Shopping
2	Food Prep, Shopping, Socializing with Others
3	Kitchen Cleanup, Physical Care of Children, Talking with Children, Relaxing/Doing Nothing Alone, Watching TV, Rest of Relaxing and Leisure
4	Financial management, Picking Up Children, Rest of Caring for Children, Education Related Child Caregiving, Eating and Drinking
5	Playing with children, Sports and Exercise
Women	
1	Relaxing/Doing Nothing Alone, Financial Management
2	Interior Cleaning, Working Main Job, Grocery Shopping, Kitchen Cleanup
3	Food Prep, Physical Care of Children, Picking Up Children, Watching TV, Rest of Relaxing and Leisure,
4	Talking with Children, Education Related Child Care, Shopping, Socializing with Others, Eating and Drinking
5	Playing with Children, Rest of Caring for Children, Sports and Exercise

Note: Sample is restricted to those men and women who had some minutes of child caregiving during their diary day. Hierarchal cluster method with Ward linkage in Stata used to determine clusters. Clusters determined on a sample of average activity emotion scores weighted by proportion of men's and women's time spent in the activity.

Table 6: Activities in Gender-Differentiated Clusters

	Men's Cluster number	Women's Cluster number
Interior cleaning	1	2
Working main job	1	2
Grocery shopping	1	2
Food prep	2	3
Shopping not groceries	2	4
Socializing with others	2	4
Kitchen cleanup	3	2
Physical care of children	3	3
Talking with children	3	4
Relaxing/doing nothing alone	3	1
Watching TV	3	3
Rest of relaxing and leisure	3	3
Financial management	4	1
Picking up or dropping off children	4	3
Rest of caring for children	4	5
Education related child caregiving	4	4
Eating/drinking	4	4
Playing with children	5	5
All of participating in sports, exercise and recreation	5	5

Note: Same as Table 5.

Table 7: Average person-level and activity level U-indices

	U-index*	
	Men	Women
Average person-level U-indices		
All 19 included activities	22.1%	24.6%
Child caregiving activities	9.9%	18.8%
Average activity-level U-indices		
Interior cleaning	47.5%	30.1%
Food prep	7.6%	18.1%
Kitchen cleanup	6.7%	41.7%
Financial management	10.3%	89.0%
Physical care of children	11.6%	20.0%
Playing with children	13.9%	3.9%
Talking with children	5.0%	4.2%
Picking up or dropping off children	9.5%	46.3%
Rest of caring for children	1.9%	17.8%
Education related child caregiving	14.6%	15.8%
Working main job	30.7%	32.5%
Grocery shopping	51.6%	21.2%
Shopping not groceries	8.4%	9.3%
Eating/drinking	17.5%	15.3%
Socializing with others	16.9%	32.3%
Relaxing/doing nothing alone	17.0%	41.0%
Watching TV	26.2%	25.1%
Rest of relaxing and leisure	15.7%	20.2%
All of participating in sports, exercise and recreation	2.9%	11.6%

*This is the average percentage of time that is experienced as unpleasant.
Note: Sample is restricted to those men and women who had some minutes of child caregiving during their diary day.

Table 8: Actual and counter-factual average person-level U-index

		Actual U-index for men	Counter-factual U-index for men if they used their time like women	Actual U-index for women	Counter-factual U-index for women if they used their time like men	Percent of gender-gap in actual U-index attributable to differences in gendered activity-level U-indices
Using women's time use as the reference group	All 19 included activities	22.1%	20.9%	24.6%	NA	148%
	All 6 included child caregiving activities	9.9%	10.0%	18.8%	NA	99%
Using men's time use as the reference group	All 19 included activities	22.1%	NA	24.6%	25.2%	124%
	All 6 included child caregiving activities	9.9%	NA	18.8%	17.8%	88%

Note: Sample is restricted to those men and women who had some minutes of child caregiving during their diary day.