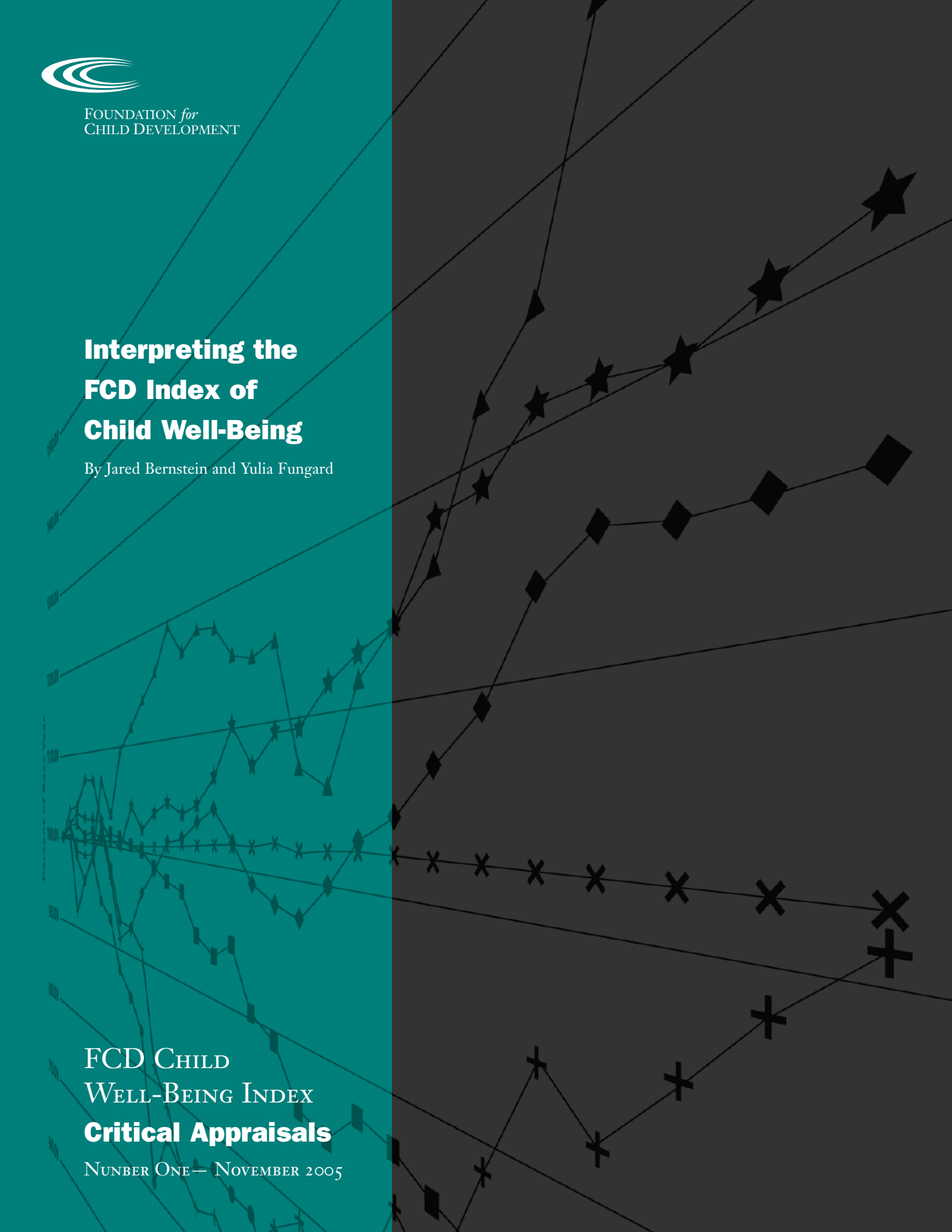


FOUNDATION *for*
CHILD DEVELOPMENT

Interpreting the FCD Index of Child Well-Being

By Jared Bernstein and Yulia Fungard

FCD CHILD
WELL-BEING INDEX
Critical Appraisals

NUNBER ONE— NOVEMBER 2005

FCD Child Well-Being Index (CWI) – Critical Appraisals

About the Series

FCD has initiated a long-term effort to develop and promote an evidence-based index of child well-being to monitor how well American children are doing over time. The FCD Child Well-Being Index (CWI) is based on seven domains of well-being comprising 28 indicators. It has been released annually since 2004.

The CWI was developed by Kenneth Land, professor of demography and sociology at Duke University, and his colleagues, Vicki Lamb and Sarah Kahler Mustillo. Land has spent the better part of his career involved in quality-of-life studies, though the CWI is his first effort based on the well-being of children.

Both FCD and Land regard the CWI as a work in progress. As a result, FCD will commission papers that reflect on both the theory and method of its construction to extend and refine it.

Interpreting the FCD Index of Child Well-Being

By Jared Bernstein and Yulia Fungard

The Index of Child Well-Being, a project of the Foundation for Child Development, is an index designed to provide information about the evolution of the quality of life of America's children. The index, hereafter CWI, is described by the project coordinator, sociologist Kenneth Land, as "…an evidence-based composite measure of trends over time in the quality of life or well-being of America's children and young people" (Land, 2005).

The components of the index are 28 series of annual data over the past 25-plus years, with each one capturing some dimension of child well-being (see Appendix Table 3). The underlying series are aggregated, by unweighted average, into seven "domains," as follows:

- Family economic well-being
- Health
- Safety/behavioral concerns
- Educational attainments
- Community connectedness
- Social relationships
- Emotional/spiritual well-being

These domains are then averaged for each year's value.

As Land writes (2003), "These seven domains of quality of life have been well-established as recurring time after time in over two decades of empirical research in numerous subjective well-being studies." They certainly have good "face validity" (i.e., they look like relevant measures), in that they reflect the key concerns of social scientists interested in child well-being.

In this working paper, we examine a number of questions raised by the CWI. First, how might we interpret the findings generated by this research? At the simplest level, one could simply accept that an up tick in the index meant that kids were doing better, but that is unsatisfying on many levels. How much better? Better in what sense? If some components are going in different directions, how does that affect our judgment? Second, given that the index's domains are unweighted, how are users to judge movements in the overall index that combine equally-weighted countervailing movements in its sub-indices?

And then there's the "yardstick," or benchmark issue: "better" compared to what? Compared to last year, the base year, the peak year, or some statistical benchmark (e.g., is the change statistically significant)? If a change in the index is statistically significant, is it socially or economically significant?

In the absence of a relevant counterfactual, the danger is that the media and others will simply report an increase (good) or a decrease (bad), without reference to other ongoing dynamics. Such dynamics, for example, a strong increase in economic growth, might be a reference point against which to judge the change in the CWI. In sum, it would be useful to think about a context within which to judge movements in the index.

Another important question asks how do movements in the index compare to those of large economic aggregates, like real gross domestic product? Should the well-being of children be expected to reflect GDP growth? The question invokes the critical issue of the impact of resource distribution, or income inequality, on child outcomes. That is, if the economy is expanding, but various factors are at play that distribute income growth towards certain groups and away from others, the linkages between economic growth and broadly shared child well-being can be broken. In fact, prior research has shown that inequality does create a wedge between economic growth and innumerable outcomes, including kids' well-being (Mishel et al, 2005, Chapter 5).

Empirically, we examine the extent to which the index and some economic series—more precisely, changes in the series—track each other. How high are the correlations and do they seem to flow from the economy to the index? Getting below the surface, do movements in the components of the CWI correlate with movements in real GDP, for example? To what extent do these correlations change over time and are there clear explanations for the changes? Is there evidence of an inequality wedge dampening the relationship between the index and the economy?

The main findings are:

- Given the lack of weights, it is difficult to determine whether changes in the CWI accurately reflect changes in child well-being. We suggest ways in which the project directors might consider addressing this challenge.
- Putting this constraint aside, there are numerous criteria by which to interpret changes in the index. We examine statistical significance, but much more interesting is the extent to which the index correlates with movement in the economy. We find little correlation between the overall CWI and aggregate economic variables, but more so regarding the economic domain. One simple simulation suggests that the 2001 recession, a fairly mild one, cost the CWI about one percentage point.
- While we believe overall economic progress is a valid benchmark against which to judge progress in the CWI, the increase in economic inequality creates a wedge between economic growth and child well-being. Given the importance of this phenomenon, we ask whether the index should include a measure of inequality. Our assessment is probably not, since the variables in the economy domain already reflect inequality's impact on family incomes and poverty.
- As Land has pointed out, there is actually an implicit weighting scheme in the index, where the variables in domains with more components are down-weighted relative to those with fewer components. In the absence of a convincing rationale for this implicit weighting scheme, we think it would be better to weight all the underlying variables equally.
- More for fun than insight, we ran our own little survey to weight the CWI, revealing—surprise!—that economists tend to give more weight to economic variables. Our exercise did, however, suggest to us that deriving a set of weights from experts in the field of child well-being may not be too daunting a task for Land et al to undertake.
- The CWI has great potential—potential that is already being realized—to get key policy and media actors talking and thinking about these critical issues in a more holistic manner than perhaps ever before. In our era of isolated policy silos, this is one of the project's great contributions.

Interpreting Changes in the CWI

What is one to make of changes in the CWI? When the consumer price index (CPI) rises, we generally believe that prices have risen; when the Dow Jones Industrial Average falls, we know that the prices of stocks in that index have fallen. When the CWI goes up, are children better off?

At one level, this is a simple question of statistical significance, discussed below. But, though we tackle it below, this is the least compelling part of the question. After all, a tick up in the CPI can, and does, move markets without being significant. This is likely because economic actors believe the index is telling them something about where prices are headed, and therefore provides them with critical information about related policies, like interest rate movements.

Given the many and highly varied components of the CWI, and the lack of weights (each domain is treated equally in the calculation of the overall index), the extent to which it accurately reflects children's well-being is an open question. Those who think child obesity matters may find it hard to accept that child well-being is up 13% over the last 10 years, a period when the rate of overweight children rose by 29%. On the other hand, it's much less challenging to imagine that a price analyst (and we fit much more closely into this category than child experts) could accept the fact that overall prices have gone up 27% over the last decade, even though college tuition and fees are up 77%.

A salient difference here is the lack of weights in the CWI. We discuss this issue in greater depth below, but the difference in the two assessments just noted stems in large part from the fact that weighting consumer prices is far easier than weighting the CWI, since consumers can reveal their preferences (i.e., weights) by how they divide up their consumption among the market basket of goods in the CPI. But how does one determine how much higher child poverty lowers child well-being relative to greater obesity, not to mention the rate of weekly religious attendance or how much families with children have moved within the year?

For the record, we should note that there's been a great deal of controversy as to whether the CPI accurately measures price changes. However, the critique is less regarding the appropriateness of consumer weights and more whether the index appropriately captures the changes in the quality of the goods in the market basket. For our purposes, the point to keep in mind is that no index is impenetrable to criticism.[1]

Even within domains, divergent trends can create uncertainty as to whether the trend within the domain is truly reflecting improvement. The social relationships' domain has two components: the share of kids in single parent families, and the share that has moved in the past year. The former component fell 11% over the past 10 years while the latter is up about 6% (the polarity of the variables in the CWI is structured so that improvements in child well-being move up the number line and visa versa). These movements essentially cancel each other out, and the domain is up slightly over this period (the "kids who moved" index has a higher base in these years so it dominates the average for the domain).

Putting aside the question as to whether these two variables suitably capture children's "social relationships," analysts may well differ on how they weight these trends, with our priors suggesting that single parenthood would get a heavier weight than the moves in the last year. Just by way of illustration, arbitrarily weighting single parenthood at 0.8 and moving at 0.2 yields a decline of 4% in the domain over this period.

In the absence of information on the relative importance of the CWI's components, it may remain hard to know just what to make of changes in the overall index. As stressed in the conclusion, however, this does not invalidate the utility of the CWI, which has the potential, a potential that is already being realized, to bring these issues together in a way that is both unique and useful.

Another question of interpretation asks "relative to what?" That is, against which benchmarks should we judge changes in the CWI? Too often in policy measurement, the debates reflect only "first derivatives"—is a metric going up or down?—ignoring key questions of whether the change is what we might expect or hope for, given the movements of other related variables. For example, various publications cite the increase in the index of 4.5% over the base year of 1975 (FCD, 2005).[2] Yet our economy, in terms of real GDP growth, more than doubled over these years, creating a huge differential. Should we be disappointed in this achievement gap or is GDP growth an inappropriate benchmark?

We examine this question in detail below, but this is, of course, but one benchmark. There are other yardsticks to consider as well. Another economic yardstick is productivity growth, since it is a common assumption among economists that productivity growth translates fairly directly into higher living standards (the intuition here is that an X% gain in productivity means we can increase output by X% without adding hours of work, or, keep the same level of output and enjoy more leisure). Between 1995 and 2000, the CWI grew by 8% and productivity by 13%, suggesting that about two-thirds of the potential increase in living standards from higher productivity growth were reflected in increased child well-being, as measured by the CWI. Since 2000, this differential has been much greater, with productivity up 12% and the CWI up 2% (this comparison goes to 2003, since the 2004 CWI value is mostly based on forecasts).

Many reports on the index have benchmarked its movements in one-time periods against those in other periods. This seems to us a useful, internally consistent comparison, but it's still important, in our view, to identify some criteria outside the index against which to benchmark its progress.

Statistical significance

One basic question of interpretation asks how much must the index change over a given period to reach statistical significance. While this may not be the most compelling question—we don't recall it being raised at all in the extensive media coverage of the CWI—it is important in the interest of avoiding over-interpretation of small changes that may well be indistinguishable from no change at all (type I errors).

1 Sociologist Don Hernandez commented on an earlier draft, noting that to the extent that one's personal market basket diverges from the average, the CPI will be less reflective of the price changes one experiences. In the CWI context, the point is that even a widely accepted weighting scheme would be unlikely to satisfy someone who believed, e.g., that obesity was far and away the biggest problem facing kids.

2 Readers may note that some of the values for the CWI differ slightly between our paper of various FCD publications. That is due to a few minor mistakes we found in the underlying computations of the index, which, when fixed, make no substantive difference to the previously reported results.

As described below, we use a method called "bootstrapping" to determine the necessary statistics for this section. This is a common method for determining statistical significance by randomly resampling with replacement from the original distribution (in this case, from the seven domains). From each resampled series, we then calculate the relevant statistic. This yields a distribution of that particular statistic—in this case, the annual change in the CWI—from which we can derive a standard error.

The conclusion of this exercise is that annual changes in the index larger than 0.46 are statistically significant, i.e., they are unlikely to have occurred by chance. Those who are uninterested in the details of how this value was derived are encouraged to skip ahead to the next section.

Determining statistical significance in an index like the CWI is not trivial. Since the index is an average of 28 underlying series, the use of parametric statistics invokes unsupportable assumptions about the relevant distributions from which to derive significance tests. When the distribution of the test statistics is unknown or too complex to derive, researchers often turn to bootstrapping techniques.[3]

In the case of the CWI, there is a question as to what is the appropriate level at which to derive bootstrap samples. While the simplest approach is to bootstrap the final index, this is unnecessary restrictive. Intuitively, by resampling with replacement, we want the bootstrap procedure to identify the extent of variance in the sample, which in the case of the CWI is a function of the variance in the time-series that comprise the index. Since the index is an average of the domains, and not an average of the 28 series, we decided to bootstrap at the domain level. Thus, we take 1000 samples with replacement from each of the domains and for every sample we recalculate the CWI for that year.

For the bootstrap, and throughout this review, we focus on changes in the indices, not on their levels. The main issue here is that the vast majority of the underlying series, like many time-series, are serially correlated (i.e., the value in one year is closely correlated to the value in nearby years). Bootstrapping theory suggests that such correlations should be maintained in the resampling, a task most typically accomplished through the use of moving blocks (i.e., resampling consecutive blocks of a time-series as opposed to individual observations). When dealing with many underlying series, we doubted our ability to keep track of whether we had achieved the goal of maintaining the correlations internal to the 28 series. Once differenced, however, the series were stationary and we could use normal bootstrapping resampling procedures.

The authors of the CWI tend to present and discuss changes in terms of differences in the index numbers themselves, and we present our results for these changes. It is more common to examine percent or log changes in such indices, as is done with, e.g., the government's various price and wage indices (e.g., the CPI or Employment Cost Index). But since the CWI moves very little off its base of 100 in 1975, growing only 4.2% over the full period, point changes are very nearly coincident with percent changes.

The mean of the 1000 bootstrapped samples of changes in the index was 0.175, just about equal to the mean of the change in the CWI, 0.173 (note that we used all the years of data, even though some of the values at the end of the series are forecasts; omitting these had no effect on the analysis, so we left them in). The standard deviation of the bootstrapped sample means was 0.228, which is about the same as the standard error of the mean from the CWI: 0.230, meaning that changes of about 0.4 are statistically significant at the 5% level. Thus, according to the bootstrap results, the use of the usual parametric formulas for assessing statistical significance would not be problematic.

3 The bootstrap is a method for determining statistical significance by randomly resampling with replacement from the original distribution (in this case, from the domains). From each resampled series, we then calculate the relevant statistic. This yields a distribution of that particular statistic—in this case, the annual change in the CWI—from which we can derive a standard error.

Based on these statistics, and the fact that index can register some relatively large annual changes (the standard deviation of the change in 1.26), it is not that uncommon for a single year change to reach statistical significance. However, given the "choppy" nature of the year-to-year changes, it makes sense to put more faith in changes that occur over longer periods. Commentators may have intuited this, as most media discussion of the index tends to reference longer-term trends in the index and its components. However, even when yearly changes in the overall index are not statistically significant, it is possible for movements in the internal components to move significantly in opposite directions, canceling each other out in the overall index.

Correlations and (Granger) Causality With Economic Aggregates

As noted above, one obvious benchmark for the CWI is the state of the economy. Most would agree that the trajectory of child-well being ought to be at least somewhat linked to changes in the economy. We know from cross-national comparisons, for example, that many health indicators are positively and causally linked to GDP per capita, and that richer countries are more likely to make significant public investments in various aspects of life—education, safety, environment—that are salutary for children. This section examines the relationship between some highly visible macro-variables, primarily real gross domestic product (GDP), and the CWI.

Some economic variables, such as median family income and child poverty, are embedded in the CWI. Thus, the correlation between, say, real GDP and the overall index will be a function of their relative weight in the index and the correlation between these sub-indices and GDP.

The first component of this correlation—their relative weight in the CWI—is least important. The 28 sub-indices are unweighted, although, by dint of aggregation into domains they do get an implicit weight, inverse to the number of indices in each domain (we discuss this in the next section).

Much more determinant of the extent to which the CWI, specifically the economic well-being domain of the CWI, correlates with other economic variables, like GDP, is the extent to which these variables correlate with median family income, child poverty, the rate of kids with health insurance, and parental employment.

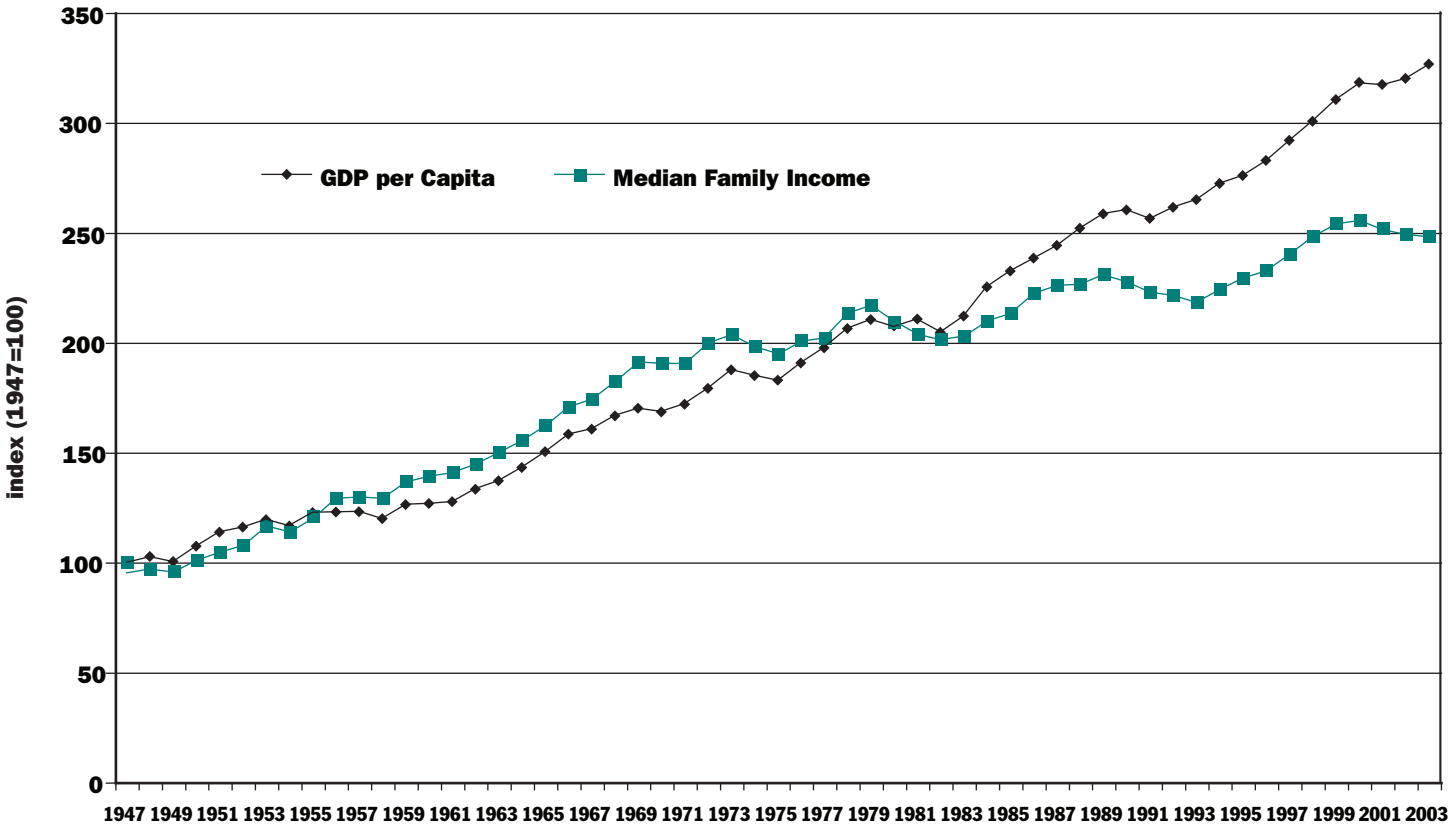
This observation is important in the following sense: if, as has been the case over the period over which the CWI is measured, economic inequality is rising, the correlation between the economic measures in the index and GDP is weakened.[4] As overall growth is increasingly concentrated at the upper end of the income scale, inequality creates a wedge between, say, growth in the median family income or movements in child poverty and growth in GDP.

4 See Blank (1993) for an analysis of this phenomenon re: poverty.

This relationship is clearly shown in Figure 1, which plots per capita GDP against median family income.[5] Between 1947 and the early 1980's the two variables moved pretty much in sync; both about doubled in real terms over these years, implying that the benefits of the expanding economy were reaching middle-income families. Yet, over the past few decades, this relationship has failed to hold. As economic growth has become increasingly concentrated among upper income families, the living standards of those in the middle of the income scale have risen much less in step with GDP per capita. This dynamic will dampen correlations between overall economic indicators, like GDP and productivity, and the economic components of the CWI.

Quantifying these correlations is straightforward, but first we must decide whether to compare changes or levels. The fact that most of these variables have strong trend components suggests we learn little from the level comparisons. Simply by dint of the underlying trends, level comparison will produce spurious correlations that tell us, for example, that two variables are both going up over time. This is particularly relevant when GDP is involved, since, outside of relatively rare downturns, GDP grows each year (as shown in the previous figure). Much more revealing is the extent to which their changes correlate.

Figure 1: **Real GDP per capita and real median family Income, 1947-2003**



5 As noted in Bernstein (http://www.epi.org/content.cfm/webfeatures_snapshots_09082004), inequality is by no means the only factor in play here. Others include the shift to single parent families and the fact that the GDP deflator has grown less quickly than the income deflator (here, the CPI-RS). The inequality effect, however, dominates these other explanations.

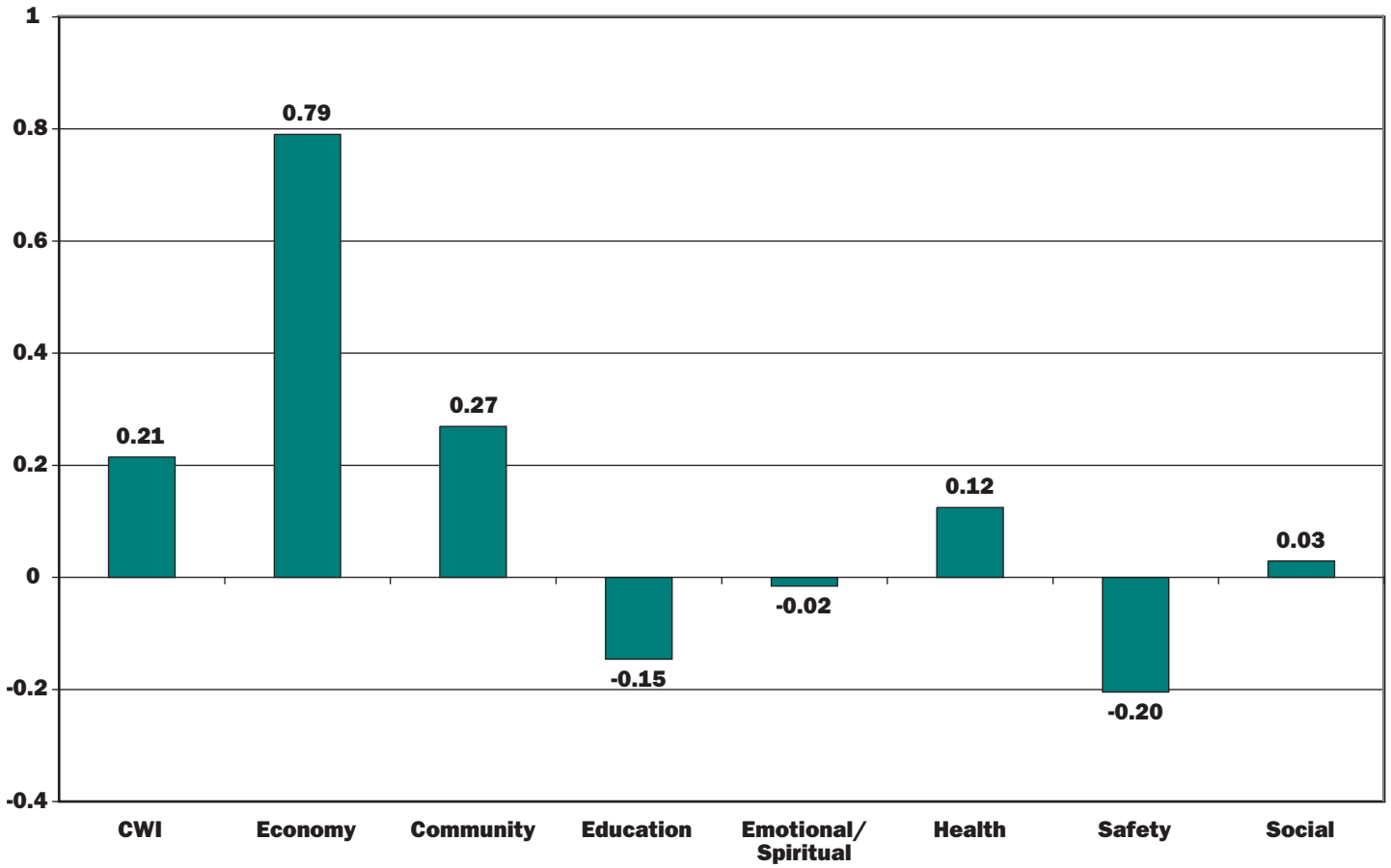
Figure 2 presents a bar graph of correlations of the changes in the CWI and the seven domains with those in real GDP.[6] Other aggregate measures of interest were either far less correlated than GDP (e.g., productivity), or were virtually the same (e.g., per capita GDP) so we stick with log changes in GDP throughout the analysis.

The results are revealing in that the only correlation of any magnitude is with the economy domain (r=0.79). Changes in the CWI are not very highly correlated with those of real GDP, yielding a coefficient of 0.215. The next figure (Figure 3) shows the correlations between GDP and the economy domain components. With the exception of rate of kids with health insurance, these correlations are also relatively high.

It is not unexpected that the economy domain correlates with GDP, but we might have expected a greater correlation between changes in the other domains and the overall index. We discuss this expectation in greater detail below.

Correlation is, of course, not causation, and there are simply far too many components comprising both the economy and the CWI to construct a detailed, causal model, i.e., one that would convincingly prove or disprove that changes in the economy were driving changes in the CWI. We can, however, examine "Granger causality." This is a simple and intuitive way of examining whether one variable contains information helping to explain the trajectory of another, or visa versa. That is, with this test, we can see whether changes in the economy

Figure 2: **Correlations Between Changes in CWI Domains and Real GDP**



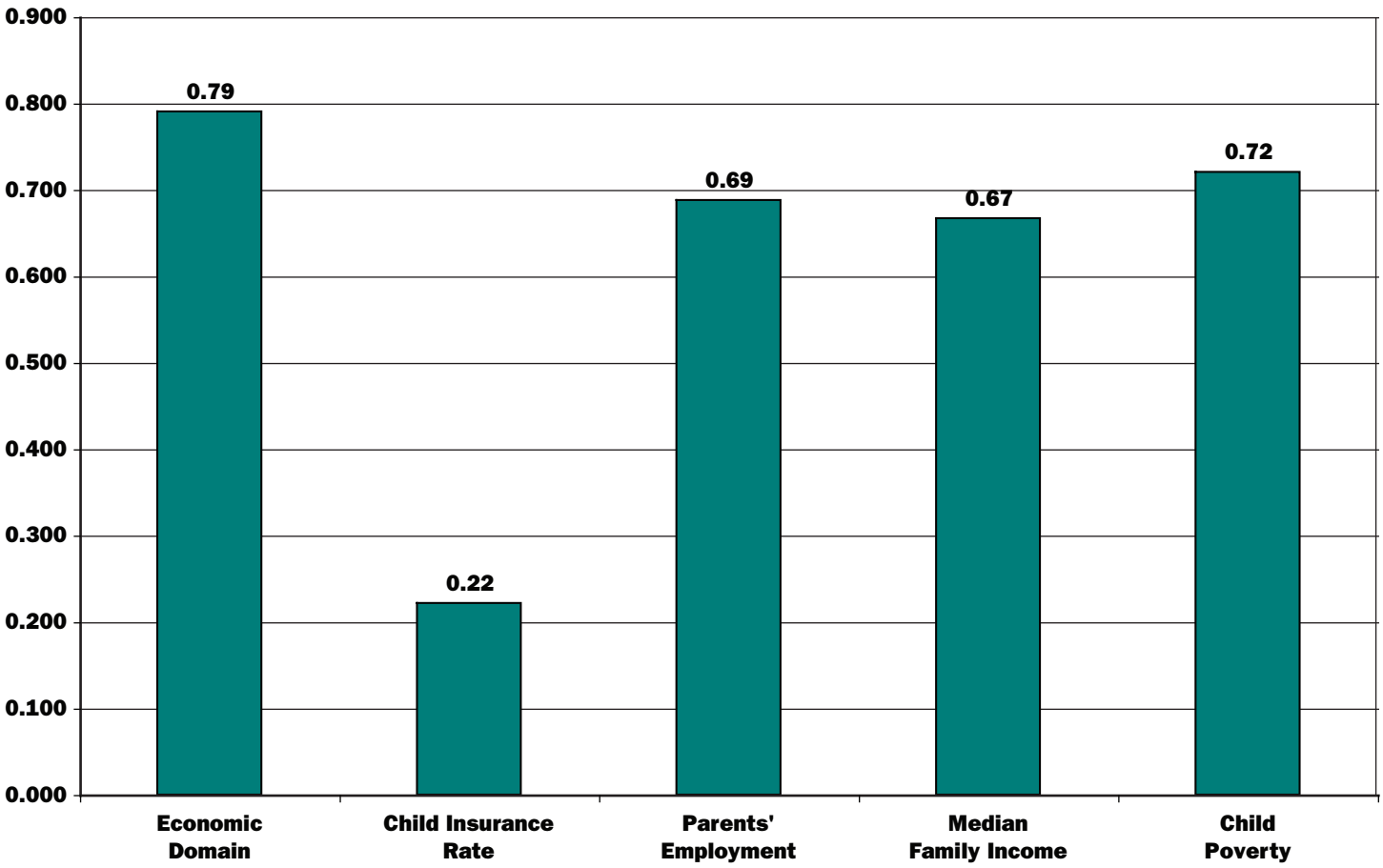
6 Changes in GDP and related values are in logs. Change in the index is not typically logged since, and thus represent percentage-point changes from year-to-year. Since most domains don't end up too far from their base of 100 in 1973, this choice makes little difference to the statistical analysis, e.g., a change from 103 to 104 is equal to one percentage point, or, in logs, 1%.

precede changes in the CWI or its components. If changes in GDP tend to occur before those in the CWI, we say that GDP "Granger causes" the CWI, in the sense that changes in one variable (GDP) precedes the other.[7]

The results of these tests (relegated to Appendix Table 1) reveal no Granger causality in these variables (we use two-year lags in all tests; longer lags were insignificant). To take one example, in the language of Granger tests, we cannot reject the null hypothesis that dGDP does not cause dCWI or dECON, i.e., changes in real GDP do not precede changes in the index. The hypothesis noted above, and shown in Figure 1, is that inequality gets in between overall growth and the CWI or its component domains. However, controls for inequality—we tried the change in the family income Gini coefficient and the ratio of the top 5% and bottom 20%—have little effect on the outcome. But these are fairly crude controls in a relatively short-time series.

Another approach, also subject to sample limitations, is to run the test over a period when inequality was growing more slowly, which, for these data is 1992 forward, a period when inequality between the middle and low-end of the income scale was actually compressing (see Mishel et al, 2004). These

Figure 3: **Correlations Between Changes in Economy Domain Components and Real GDP**



7 Mechanically, the test is comprised of two regressions. One regresses changes in the CWI on its own lags and changes in lagged dGDP; the other regresses dGDP on its own lags and dCWI. In each case, the Granger test asks whether the lags of the independent variable (e.g., dGDP in the dCWI regression) are jointly significant.

results, though based on too few observations to be very confident in the findings, suggest that Granger causality does in fact run from GDP growth to both the CWI (though only at the <0.10 significance level) and the ECON domain (at the <0.05 level). Based on these results, we later ask whether it may be worth more directly incorporating the impact of inequality into the CWI.

Since the Granger test is largely interested in precedence, it leaves out contemporary variables. Appendix Table 2 shows a few regressions with contemporary GDP as an independent variable. Note that changes in the log of real GDP are strongly correlated with the economy domain (note much higher adjusted R-squared and significant t-stats) though not with the CWI. Again, this is not evidence of causality. We are simply tapping the relatively strong correlation shown above between changes in GDP and those in the economy domain, or, more bluntly, the fact that in downturns, real GDP falls along with median family income and parental employment, and poverty rises. The elasticity for the economy domain is statistically indistinguishable from one, suggesting that a one percent increase in the growth of real GDP adds one percentage point to the change in the economy domain, which in turn adds 1/7 of a point to the growth in the CWI.

This gives some sense of the extent to periods of weak GDP growth correlate with worse outcomes in the overall index. For example, we can use the regression to simulate the level of the economy domain, and that of the CWI, if instead of falling in 2001-2003, the growth rate of real GDP had held steady at its 2000 level. Under this simulation, instead of falling 3.5 points, 2000-2004, the economy domain would have gone up 4.5 percentage points, for a net increase of about eight points in the level of the 2004 economy domain over that of 2000. Averaging through with the rest of the domains left at their original levels, this would have lifted the 2004 CWI by more than a percentage point.

The analysis reveals that over the life of the CWI, overall economic growth, as measured by the growth of real GDP is not very intimately related to the index. The correlation between economic growth and the CWI is small, and Granger tests provide no evidence of temporal precedence: changes in GDP do not help explain the future path of the CWI.

For economic determinists, this may be puzzling or unsettling. That is, those who believe that economic conditions are important determinants of everything from birth weight, to educational attainment, to the share of single parents, might expect these economic effects to be picked up by the search for correlations undertaken above. And as noted, cross-national comparisons do show that richer countries do better on many, if not most, of the indicators in the CWI.

But such international comparisons are typically comparing very large level differences in wealth and outcomes, quite a different exercise from the above analysis. In this sense, the lack of correlation may not be so surprising. We might not expect a good economic year or two, in terms of GDP growth, to have much impact either way on, for example, child suicide, which (and we are not experts) seems to us much less related to broad economic conditions than, say, a variable like child poverty.

In that regard, it is notable that we find strong correlations between GDP and the economy domain. Again, we could not establish Granger causality—changes in GDP do not (Granger) cause changes in the economy domain—over the 1975-2004 period. However, we hypothesized that one reason for this is the fact that rising inequality over much of this period weakened the relationship between the economy and the components of the economy domain. Though direct controls for this did not help much, when we isolated the analysis to years when inequality growth was dampened—years in the 1990s when poverty and median family income responded much more closely to the economy's growth—we were able to show GDP growth pushing up the economy domain.

We take from this analysis, and much more so from reams of other research, that economic inequality is an important determinant of the economic well-being of children, particularly kids in less advantaged families. This research has established that overall economic growth by itself is a necessary condition for lifting the economic conditions of lower-end families. But unless said growth is equitably distributed, either by the labor market or through redistributional fiscal mechanisms, it will not be sufficient.

Does this insight imply that a measure of inequality should play more of a direct role in the index? While we can see a rationale for its inclusion, it is, in a sense, "lurking" throughout, in the very sense we have been stressing. That is, given that child poverty and median family income are already in the index, and that inequality intimately affects their outcomes, it's not obvious that entering inequality explicitly adds much. In fact, we would push harder for its inclusion if those who built the CWI included a measure like GDP, which can hide the extent to which unequal distribution is leading to less optimal outcomes than the trend implied by overall economic growth (which may be why they left it out).

However, the possibility remains that two of the measures in the economy domain—child poverty and median family income—will fail to reflect public policy measures intended to offset the growth of inequality, such as redistributive tax policy. For example, an expansion in the Earned Income Tax Credit, a significant wage subsidy for low-income working families that lowers the post-tax child poverty rate, will not be directly picked up in the CWI, since the poverty and income measures are pre-tax. In the context of the above discussion of the inequality wedge, this is a potentially important omission, since such measures can be viewed as very conscious policy attempts to offset the wedge (or visa-versa: regressive tax changes will widen the inequality wedge). Thus, though such measures are harder to come by and might be considered less reliable, Land et al might consider examining post-tax poverty and income measures. On the other hand, to the extent the improvements in kids' living standards result from such redistributive measures, they would presumably be reflected in some of the other components in the index. Note, for example, that increases in publicly provided child health insurance coverage will show up in that series.

Weights

Probably the greatest challenge in index construction is how to weight the index's component parts. In rare cases, the weights are fairly obvious. For example, the prices of the items in the consumer price index are weighted by average consumption shares. The most recent CPI release shows that food represents 14.3% of the market basket of goods and services that comprise the index and that is its weight.[8] This is a classic example of "revealed weights," where those to whom the index applies, in this case urban consumers, "vote with their feet" to reveal the researchers with the appropriate weight. Is there a lesson here for the CWI?

The CWI gives each domain equal weight. For a given year, the value of the index is simply the average of each domain's value for that year. This raises two weighting issues: a) are there a set of weights that might be applied to the index, and b) is the current implicit weighting scheme justified? The latter issue is conceptually much simpler so we discuss it first.

As discussed above, and shown by Land in various places, it is not the case that each of the 28 underlying series gets an equal weight. Instead, each domain is equally weighted. Thus, series in relatively populous domains, like health (comprised of six series) are down-weighted relative to a domain like educational attainments, comprised of two series. This means that we can change the value and trajectory of the index by moving its component series around.

Arithmetically, the implicit weighting scheme works like this: Since each domain is weighted 1/7 in the overall index, each component in the domain is weighted by 1/[number of variables in the domain]*1/7. Thus, the variable "share of kids that moved in the last year," from the "social relationships" domain, which is comprised of only two variables, gets a weight of 1/14 (1/2*1/7). But the variable "infant mortality" from the health domain (with six variables), gets a much smaller weight of 1/42 in the overall index (1/6*1/7).

For example, by moving "rate of voting by 18-20 year olds in presidential elections" from the community connectedness domain to the social relationships domain changes both of

8 See Table 1, http://www.bls.gov/news.release/pdf/cpi.pdf.

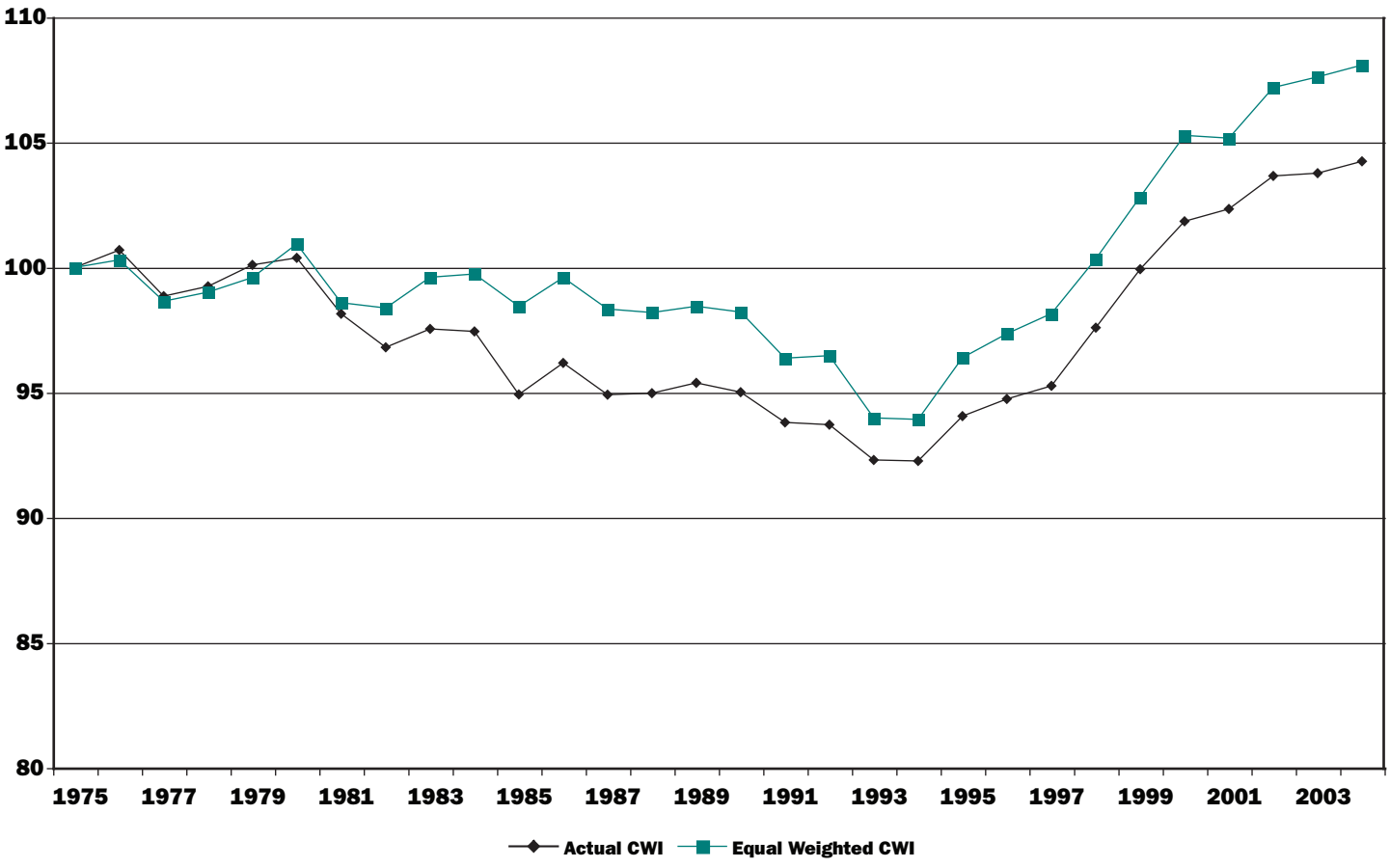
the domains (the community domain grows 10% faster without voting over the life of the index; the altered social domain grows faster in the mid-80s and falls faster in the 1990s), such that the 2004 value of the overall index rises to 105.9 from 104.2, just by dint of this one change.

Figure 4 plots the CWI as is against a version with equal weights applied to each of the 28 underlying series (Land has plotted these same figures, i.e., the CWI team is well aware of this point). Note that the version which gives equal weights to each series rises more quickly than the domain-weighted series, implying that series showing more improvement in child well-being tend to be grouped in more populous domains, and thus get down-weighted relative to those growing more slowly.

There is a rationale for domain weights. In cases where the numbers of series in the domains are very different, domain weighting might have more appeal. For example, if domain X had 50 series and domain Y had three, equal weighting would lead to the "domain-theme" in X playing a much larger role than the theme of Y, relative to an approach wherein X and Y were equally weighted domains.

Even considering this possibility, we still view domain weighting as a potential shortcoming of the CWI. While we are not sociologists steeped in the theoretical and empirical rationales for grouping the sub-indices as they currently stand, we are confident that reasonable experts could disagree on whether, say, child suicide belongs in the health, safety/behavioral, or spiritual well-being domain (it's

Figure 4: **Domain Weights vs. Equal Weights**



currently in the latter). Domain-weighting imposes a set of "Land-weights" on the index—weights made by choices of Land's team as to how they think the series should be grouped into the domains. There is, of course, no more knowledgeable experts on the CWI than Land and his team, but it seems to us that whatever is gained by such an implicit weighting scheme is not worth the critique that it engenders. We therefore think equal weighting dominates domain-weighting. None of this would preclude the researchers from the enlightening discussions of the trends in the domains.

Should Land and FCD go much further than this, trying to derive a set of relative weights for the 28 series in the index? This is a tougher call. Hagerty and Land (2004) convincingly argue that in the absence of "revealed weights," their best bet, in the sense of generating the least disagreement about what matters most, is equal weighting. But this is Solomonic wisdom: since I can't identify the mother, I'll just cut the baby in half! Or, more to the point, since Land et al have no weights, they have no good rationale as to how to weight components.

One critical review of the CWI stressed the problem of equal weighting as follows (from Hulbert, 2004):

"Because the index weights all the data and domains equally, there's a limit to how revealing, or reliable, the cumulative figure can be. How does the bad news (a rise in obesity, for example) stack up against the good news (a drop in juvenile crime and teenage pregnancy)? In calculating kids' overall welfare, should the domain of ''Emotional/Spiritual Well-Being'' (which plunged in the 1980s) be considered on a par with ''Material Well-Being'' (which has been gradually rising)? Never mind the theological debate; it's a methodological quandary: it's much harder to measure the former. Surveys that ask 12th graders to rate the importance of religion -- one of three indicators of ''Emotional/Spiritual Well-Being'' (along with church attendance and the suicide rate) -- aren't nearly as solid as figures on annual income. (And such data obviously don't tell you much about the souls and moods of, say, toddlers or preteenagers.)"

These are good points, but further reflection suggests—to us, at least—that the best way to derive a convincing weighting scheme for an index like the CWI is not at all obvious. One line of thought on this topic of weighting indexes believes such weights should come "from the people," and putting aside resource constraints, it might be feasible to ask the public to rank the components in some hierarchical manner. Another line of thought suggests that experts in the field, in this case, child well-being, have the knowledge required to reveal which components should receive higher relative weights such that the overall index would provide information with greater validity.

In other words, "revealed weights" depend on whom we are looking to for revelation. In the CPI, the answer is straightforward: urban consumers. But what is the analogous audience for the CWI? The description on the FCD website states that a goal of the index is to "…help describe and monitor the condition of children over time to national policymakers and to the American public," two large and disparate groups.[9]

Asking the "American public" to rank the components in the index strikes us a touch ambitious. There are, however, many experts in child well-being, in the fields such as psychology, sociology, education, and health. Perhaps it would be feasible to survey experts in these fields to derive an experimental set of weights for the index. We believe that those who follow the index would be very interested in how a weighted CWI would differ from the unweighted version.

More for fun than for insight, we designed a simple, Likert-based survey, asking 33 people to rank, on a scale of one to five, the importance of the 28 variables in the CWI.[10] This is but one way, and a convenient but not entirely satisfactory one, to ask people to reveal weights for an index like this. A better way, but one we thought was too time-consuming for our experiment, is to force responders to rank preferences. That is, with the Likert format, there's nothing stopping respondents from giving 'fives' to each component. A better method might be to restrain their choices by giving them a certain, fixed number of points that they can distribute. If they "spend" a five on child poverty, they have less to "spend" on parental employment.

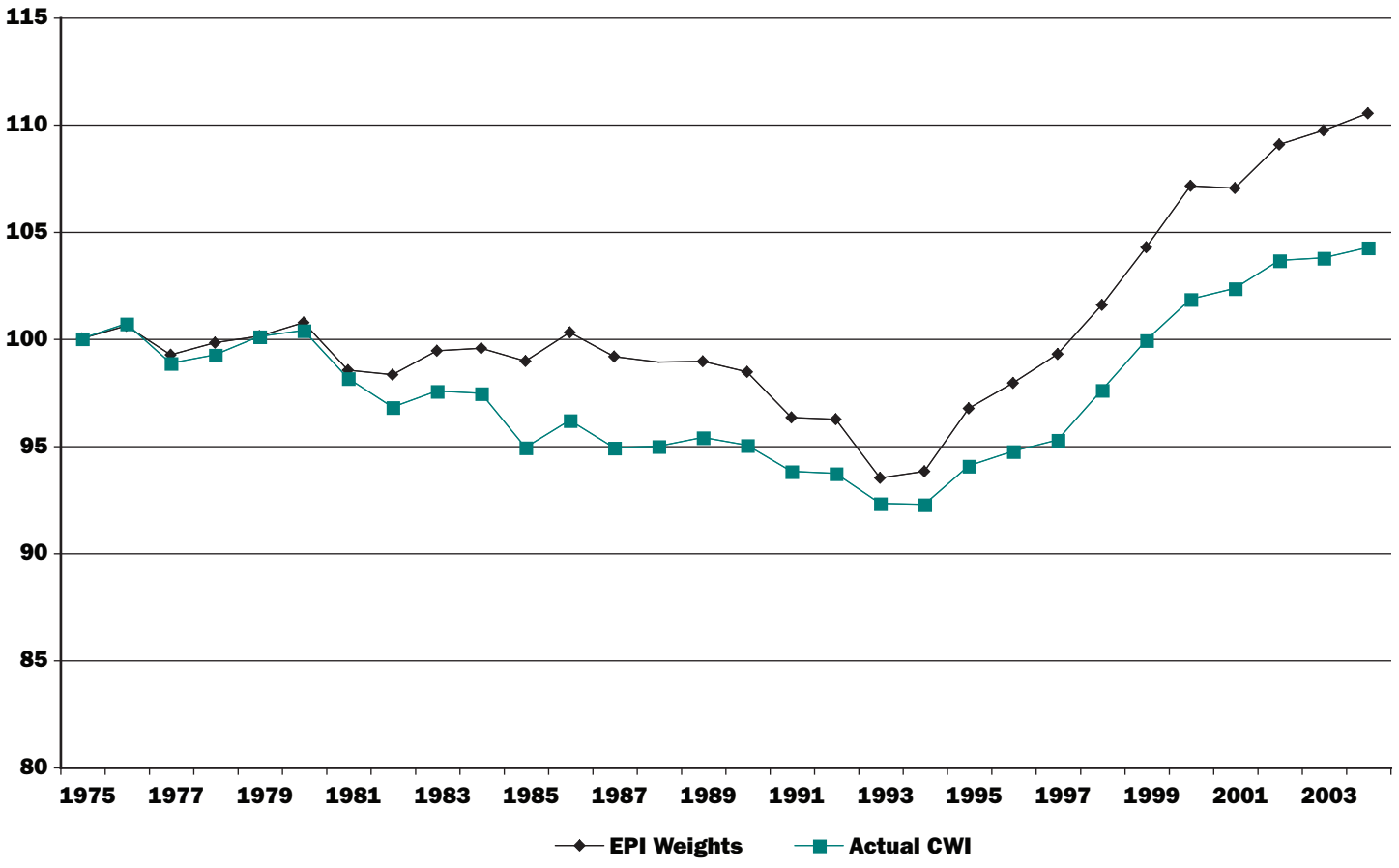
9 http://www.fcd-us.org/ourwork/k-index.html.

10 Here is the wording of the e-mail we sent out: If you can spare five minutes, please rank—***on a scale of 1 to 5, going from not important to very important***—the following components of an index of child well-being. For example, if you think one of the following measures is very important to the well-being of kids, put a '5' in the line next to the indicator. If you think it's of no importance, put '1', medium importance, put '3', etc.

We stress that our weighting exercise does not meet even the weakest criterion for representativeness and we only tried it to see if a set of weights drawn from a bunch of employees at EPI (and some of their friends) would make any difference to the outcome. The results of the EPI weights are shown in Figure 5. The major departure from the unweighted CWI is in the 1990s when the economy was particularly strong and growth was far more balanced than at any other time over the life of the CWI. Again, this probably reveals more about what happens when you survey solely EPI-type economists than anything else.

In sum, we clearly believe that FCD and the CWI team should pursue these weighting issues, possibly considering a test survey of child well-being experts. As sociologist Don Hernandez commented on an earlier draft of this paper, the important question regarding weights and the CWI is whether they would change the outcomes of the domain-weighted index. He's right, of course, but the answer is unknowable without trolling for some weights. We realize it's not a trivial endeavor, but we're sure that a relevant group of experts is connected to the Internet, and our guess is they'd be willing to help out by answering an e-questionnaire.

Figure 5: **Applying EPI Weights (Economists Up-weight the economy!)**



11 For example, one commenter on an earlier draft noted that the index would benefit from the 1) inclusion of information from the new American Time Use Survey on the number of parental hours devoted to activities with children (unfortunately, there are no historical data from this survey), or 2) the amount of time children spent watching TV.

Conclusion

Perhaps the only thing we learn from this home-spun weighting exercise, which took us but a few hours, is that it might be feasible for the CWI team to field an e-mail survey to experts in the field. But it also made us think more about the pitfalls of index construction. Given that the CWI has already shown itself to be a tool that generates much needed attention in the critical area of child-well being, our hope is that answering some of the questions we've raised can improve an already useful innovation.

Especially without the weights—either from experts or from the public—it's hard to know precisely how to judge the progress of the CWI. As Hulbert (2004) noted, the index gives you "…a number and a media-friendly acronym, to be sure, but you may not have a clear notion of whether kids are better or worse off." If suicide falls and income goes up by a similar number of percentage points, do they cancel each other out? What if fewer kids smoked but more were victims of crimes? While Hagerty and Land (2004) may have proved that in the absence of better information, equal weighting generates the least disagreement between people, it seems impossible to resolve these choices in a way that would best inform consumers of the index.

In the meantime, unless there is a very compelling rationale, switching from domain weights to equal-component weights seems like an improvement.

However, there is another side to the CWI, or at least to how it has played out so far in the echo chamber of policy debates and discussions. While movements in the overall index inevitably get the headlines, virtually every media article or scholarly discussion of the index has quickly gotten below the surface and looked at the conflicting or complementary trends. Much like the way economic analysts never fail to discuss the components of GDP growth or the "core CPI" (price movements leaving out volatile components), in its short lifetime, the CWI has already generated a great deal of discussion regarding its component parts.

In this regard, the CWI makes at least two important contributions. First, it gets people—journalists, policy makers, the public—talking about child well-being, surely one of the most important things to be concerned about in any society. Second, it hopefully gets them talking and thinking in more holistic manner, appreciating that all the components of the index matter (and surely many more variables that are left out).[11] Though we often talk about children as economic agents (as in poverty discussions), health consumers, students, patients (obesity discussions), and so on, they are, of course, all of the above and much more. To the extent that the index generates public discourse guided by this insight, it has the potential to elevate our understanding of the challenges facing children and the urgency about meeting those challenges.

The authors thank Ken Land and Sarah Meadows for kindly providing extensive underlying data, and FCD for support. We thank our EPI colleagues for responding to our questionnaire. We also thank the readers of an earlier draft for their helpful comments. Any mistakes are our own.

Bibliography

Blank, Rebecca. 1993. "Why Were Poverty Rates So High in the 1980s?" In *Poverty and Prosperity in the Late Twentieth Century*, ed. Dimitri Papadimitriou and Edward Wolff. New York: St. Martin's Press.

Hagerty, Michael R. and Kenneth C. Land. 2004. "Constructing Summary Indices of Social Well-Being: A Model for the Effect of Heterogeneous Importance Weights." Revision of a paper presented at the annual meeting of the American Sociological Association, August 16-19, 2002, Chicago, IL. (This and many other papers about the CWI are available at: http://www.soc.duke.edu/~cwi/.)

Hulbert, Ann. 2004. "Are the Kids All Right?" *New York Times Magazine*, April 11, 2004, pg. 9.

Land, Kenneth C. 2005. "The Foundation for Child Development Index of Child Well-Being (CWI), 1975-2003 with Projections for 2004." New York: Foundation for Child Development.

Land, Kenneth C. 2004. "An Evidence-Based Approach to the Construction of Summary Quality-of-Life Indices." In *Challenges for Quality of Life in the Contemporary World*, ed. Wolfgang Glatzer, Susanne von Below, and Matthias Stoffregen. New York: Kluwer.

Mishel, Lawrence, Jared Bernstein, and Sylvia Allegretto. 2005. *The State of Working America*, 2004-05, Ithaca, NY: Cornell University Press, 2001.

Appendix Table 1:

Granger Causality Tests:
Do Changes in Real GDP Proceed Changes in the CWI or the Economy Domain?
(Two lags are used for each test)

Full Sample

Null Hypothesis:	Obs	F-Statistic	Probability
DLGDP does not Granger Cause DCWI	27	0.85433	0.43921
DLGDP does not Granger Cause DECON_DOM		0.17656	0.83933

1992-2004 Sample

Null Hypothesis:	Obs	F-Statistic	Probability
DLGDP does not Granger Cause DCWI	13	3.56972	0.07797
DLGDP does not Granger Cause DECON_DOM		4.72397	0.04420

Appendix Table 2:

Regression of Changes in CWI and Economy Domain on Changes in Real GDP.
Sample (adjusted for lags): 1977-2004

Dependent Var	Constant	Dlog GDP	Lag of DV
D_CWI	-0.003	0.133	0.202
t-stat	-0.641	0.973	1.053
Adj R Sq	0.008		
DW Stat	1.877		
D_Econ Domain	-0.027	0.928	0.355
t-stat	-5.100	6.006	3.310
Adj R Sq	0.728		
DW Stat	1.843		

Appendix Table 3:

The domains and their components.

1	Family Economic Well-Being Domain	1. Poverty Rate (All Families with Children) 2. Secure Parental Employment Rate 3. Median Annual Income (All Families with Children) 4. Rate of Children with Health Insurance
2	Health Domain	1. Infant Mortality Rate 2. Low Birth Weight Rate 3. Mortality Rate (Ages 1-19) 4. Rate of Children with Very Good or Excellent Health (as reported by parents) 5. Rate of Children with Activity Limitations (as reported by parents) 6. Rate of Overweight Children and Adolescents (Ages 6-17)
3	Safety/Behavioral Domain	1. Teenage Birth Rate (Ages 10-17) 2. Rate of Violent Crime Victimization (Ages 12-17) 3. Rate of Violent Crime Offenders (Ages 12-17) 4. Rate of Cigarette Smoking (Grade 12) 5. Rate of Alcohol Drinking (Grade 12) 6. Rate of Illicit Drug Use (Grade 12)
4	Educational Attainments Domain	1. Reading Test Scores (Ages 9, 13, and 17) 2. Mathematics Test Scores (Ages 9, 13, and 17)
5	Community Connectedness Domain:	1. Rate of Preschool Enrollment (Ages 3-4) 2. Rate of Persons who have Received a High School Diploma (Ages 18-24) 3. Rate of Youths Not Working and Not in School (Ages 16-19) 4. Rate of Persons who have Received a Bachelor's Degree (Ages 25-29) 5. Rate of Voting in Presidential Elections (Ages 18-20)
6	Social Relationships Domain	1. Rate of Children in Families Headed by a Single Parent 2. Rate of Children who have Moved within the Last Year
7	Emotional/Spiritual Well-Being Domain	1. Suicide Rate (Ages 10-19) 2. Rate of Weekly Religious Attendance (Grade 12) 3. Percent who report Religion as Being Very Important (Grade 12)

The authors thank Ken Land and Sarah Meadows for kindly providing extensive underlying data, and FCD for support. We thank our EPI colleagues for responding to our questionnaire. We also thank the readers of an earlier draft for their helpful comments. Any mistakes are our own.