Welfare-to-Work Programs in America, 1980 to 2005:

Meta-Analytic Evidence of the Importance of Job and Child Care Availability

Kevin M. Gorey, Ph.D., M.S.W.[a]

[a] Professor and Assumption University Research Chair in Canadian and American Population Health, School of Social Work, University of Windsor, 401 Sunset Avenue, Windsor, Ontario, N8N 3P4, Canada

Telephone: 519 253-3000, ext. 3085, FAX: 519 973-7036,E-mail: gorey@uwindsor.ca

Acknowledgement

This research was supported in part by a Canadian Institutes of Health Research investigator award and a grant from the Social Sciences and Humanities Research Council of Canada. The author gratefully acknowledges the scholarly assistance of Geir Smedslund who on behalf of his co-investigators and the Campbell Collaboration provided access to their original meta-analytic database. The research assistance of Nancy L. Richter is also gratefully acknowledged. An earlier version of this work was presented at the 8th annual international Campbell Collaboration Colloquium in Vancouver, BC (May, 2008).

Submitted February 8, 2008 to Social Work Research: 280 word abstract, 4,395 words of text and three tables.

Running head: WELFARE-TO-WORK META-ANALYSIS

Welfare-to-Work Programs in America, 1980 to 2005:

Meta-Analytic Evidence of the Importance of Job and Child Care Availability

Abstract

This meta-analysis refined and extended a recent Campbell Collaboration review of welfare-to-work programs (Smedslund et al., 2006). Its synthesis of 65 randomized controlled trial outcomes in America over the past generation systematically replicated a small overall intervention effect on employment (relative risk [RR] = 1.04, 95% confidence interval [CI] 1.03, 1.05), but its main analytic goal was to explain the large program variability that had been identified by Smedslund et al (2006). It found: (a) there was no long-term (4-6 years follow-up) employment effect of intervention in areas were jobs were relatively unavailable, by proxy where African American men had prevalently withdrawn from the labor market, (b) there was no long-term effect of interventions that did not provide child care, (c) programs that provided child care were significantly more effective than those that did not in the short- and long-term (respectively RR = 1.11 [1.09, 1.13] and RR = 1.11 [1.06, 1.17]), even in areas of high labor market withdrawal (RR = 1.12 [1.10, 1.12]), and (d) the effect of exemplary, well-resourced programs with child care was even larger (RR = 1.18 [1.14, 1.22]). Such child care supported interventions were estimated to be three to five times as likely to prevent chronic joblessness (9.9% to 15.3% vs. 2.9%) and to do so with lower cost/benefit ratios (needing to treat only 10 to 17 vs. 62 participants to prevent each incidence of chronic joblessness). The actual availability of jobs as well as such supports as child care that enable their access seem to be key elements of welfare-to-work programs that work. Future social policies that positively affect job availability and accessibility are likely to be far more effective than those that focus on the jobless themselves.

Welfare-to-Work Programs in America, 1980 to 2005:
Meta-Analytic Evidence of the Importance of Job and Child Care Availability

Welfare cash assistance programs have always served it seems as a central catalyst for political debates on the American welfare state. They are, after all, extraordinarily costly while their relative social, familial and personal benefits (or harms) have been legitimately debated for more than a generation, with reasonable social and behavioral scientists finding ample rational support for both conservative and liberal arguments. Passage of the Omnibus Budget Reconciliation Act in 1981 served to rouse conservative advocacy. In restricting the earnings and assets of welfare recipients so that fewer of them could remain eligible while working, it lent increasing support to the notion that such welfare programs actually serve as potent work disincentives. In response, some states began to rigorously test a variety of so-called welfare-to-work initiatives. Though their central objective was to cut costs by moving people off of welfare caseloads and into paying jobs, many scholars also studied their life space affects. The Family Support Act of 1988 further formalized welfare-to-work programs. It required states to enrol increasing proportions of their Aid to Families with Dependent Children (AFDC) recipients in various job opportunity and basic skills (JOBS) training programs and to rigorously evaluate them. The Personal Responsibility and Work Opportunity Act of 1996 then allowed states even greater freedom to experiment with various strategies (e.g., welfare time limits, work requirements, sanctions) for moving the recipients of Temporary Assistance for Needy Families (TANF), typically single women with children, from welfare to work.

It has been estimated that between two and five billion dollars have been invested in hundreds of such local, statewide and national welfare-to-work evaluations, many of which were

large randomized controlled trials (RCT) (Gorey, 2003), probably representing, in aggregate, the largest social experiment ever accomplished. Recently, teams of principally econometric analysts have begun to extract the wealth of synthetic knowledge contained in this huge national database (Bloom, Hill, & Riccio, 2003; Greenberg, Ashworth, Cebulla, & Walker, 2005). The task seems incomplete, however, as key aspects of intervention programs and study contexts have not yet been adequately addressed: job and child care availability. This study aims to extract this new knowledge from the national welfare-to-work database by means of a unique meta-analysis.

Pooled Secondary and Meta-Analytic Analyses

A meta-analysis or quantitative research review and a secondary pooled analysis together synthesized the findings of 123 mandatory welfare-to-work experiments that were initiated during America's pre-TANF era (Bloom et al., 2003; Greenberg et al., 2005). Intervention group members typically received job search strategic assistance, time limited work experience, education and/or training and were compared with control group members who typically experienced usual welfare services. Aggregate findings were: (a) modest employment and earnings gains among the intervention group peaked after two to three years and diminished significantly after another two to three years of follow-up, (b) programs that emphasized getting clients into actual jobs first and which proportionally served more older and non-Hispanic white clients seemed more effective, and (c) only very limited, equivocal support was found for the possible supportive effects of job availability in study areas (unemployment rates used as proxies) or program provision of child care. These analyses served scholarly and policy communities well, especially through their synthesis of overall or main intervention effects. However, their regression explorations of intervention moderations by client, intervention and contextual characteristics tended to be based on small subsamples and to emphasize regression

statistics and statistical significance, rather than on more readily understood effect size indicators of practical significance.

The Campbell Collaboration systematic review. A Campbell meta-analytic review that extensively searched for studies from all parts of the world included 73 North American RCTs of mandatory and voluntary welfare-to-work interventions including studies of the post-TANF era (Smedslund et al., 2006). Its observed positive, but rather small overall intervention effects on employment and earnings replicated previous meta-analytic findings. It further validated the notion that work first or labor force attachment strategies are more effective than education and training-based or human capital development strategies. And it lent further credence to the notion that intervention effects tend to fade appreciably over time (e.g., observed an aggregate 10% relative diminishment in the risk of unemployment at 1 year follow-up [intervention versus control group], but only a 4% risk diminishment after 5 years). Regrettably, its meta-analytic plan did not account for such length of follow-up differences and so confounded all of its moderator variable explorations, many of which were observed to have no statistical effect (e.g., child care provision) or, though observed to be significant in a statistical sense, their practical significance could be fairly characterized as miniscule (e.g., unemployment rates in study areas). Smedslund and colleagues' (2006) most interesting finding by far did not concern summary intervention effects, but rather their great variability across places. "The local context in which a program was executed had a much more powerful effect on the outcome than did the program itself (italics mine)."

Need For Meta-Analytic Refinement

It seemed that the meta-analytic power of the Campbell Collaboration review might be significantly bolstered by better accounting for follow-up length and by enhancing the predictive

validity of key moderator variables. This meta-analytic study aims to do just that, and in so doing make its outcomes more relevant to social welfare policy makers. Unemployment, the economic cycle or job availability proxy used in the previous meta-analysis, is limited in a number ways. It only accounts for those who are still actively looking for work—labor force participators. It does not meaningfully assess chronic joblessness and consequent labor force withdrawal that can occur when good jobs are not available or accessible (Kain, 1992; Kasarda, 1989). The rate of labor force withdrawal would seem a much better proxy for lack of employment opportunities than the unemployment rate. Moreover, given the extant sociological evidence and this field's typical urban underclass context, the labor market withdrawal (LMW) rate of African American men would likely be the most predictive such sentinel indicator (Holzer, in press; Lichter & Oliver, 2000; Wilson 1987; 1996).

The idea that child care would bolster the effectiveness of welfare-to-work interventions among typical TANF clients, women with young children, admittedly, is not ground breaking. It is supported by ample quantitative and qualitative research (Axelsen, Friesner, Rosenman, & Snarr, 2007; Carnochan et al., 2005; Cochi Ficano, Gennetian, & Morris, 2006). So it is surprising that previous meta-analyses have generally not supported this notion. It seems possible that some error related to the diffusion of treatment in the primary studies intruded into these meta-analyses. For example, some welfare-to-work with child care interventions may not have differed significantly from their corresponding control groups on actual child care use because of effective control group case management, especially in states that provided child care subsidies. Or case managers working to support their clients with job search and/or training interventions may have also connected more of such 'intensively served' clients to child care services, even though child care was not part-and-parcel of the planned welfare-to-work

intervention. This meta-analysis will go beyond the mere conceptual definition of interventions to their empirical verifications. So regardless of how an intervention was described the following question will be asked of each study comparison: Did the intervention and control groups actually differ significantly on their use of child care services?

Key moderators in the present meta-analysis will therefore be LMW among African American men and the intervention-integral provision of child care, and the effects of these factors will be tested within length of follow-up strata. It aims for a more analytically valid and complex, so life space and policy valid, integration of this field's synthetic knowledge.

Methods

Selection of the Sample for Meta-Analysis

Original Campbell Collaboration review (Smedslund et al., 2006). The following research literature databases were searched (1963 to 2006): Campbell Collaboration Trials Register, Cochrane Controlled Trials Register, Sociological Abstracts, Social Science Citation Index, International Bibliography of the Social Sciences, ERIC, PsycINFO, Public Affairs Information Service, Cumulative Index to Nursing and Allied Health Literature, MEDLINE, EMBASE, Dissertation Abstracts International and the System for Information on Grey Literature. Detailed key word search schemes may be summarized as follows: (welfare-to-work or workfare or numerous acronymic program names [e.g., A Better Chance (ABC), Greater Avenues to Independence (GAIN), Job Opportunities and Basic Skills Training (JOBS), etc.]) and (employment or job or work or income or earnings). These searches were augmented with bibliographic reviews of retrieved manuscripts, a search of the book "The digest of social experiments" (Greenberg & Shroder, 2004), world wide web searches of policy research center and government websites, and contact with key informants within this field's scholarly network.

Additionally, studies had to meet these inclusion criteria: (a) used a randomized controlled trial design, (b) included participants who received welfare cash assistance (Aid to Families with Dependent Children [AFDC] or Temporary Aid for Needy Families [TANF]) and/or food stamps, (c) included interventions with labor force attachment (e.g., time-limited work experience, job search assistance) and/or human capital development (e.g., remedial education, vocational training) components and (d) assessed work status among intervention and control groups. Intervention programs that were tested in different places or over different lengths of follow-up were treated as independent hypothesis tests. A total of 73 such independent study endpoints or outcomes comprised the original meta-analytic database (see Table 1 and Included References, Smedslund et al., 2006).

Sample extension and refinement. This study's sampling frame was extended to include Social Work Abstracts and Social Service Abstracts and searches were extended to January of 2008. Seven independent study outcomes that had been ongoing at the time of the original meta-analysis were added to this study's database (Bloom, Hendra, & Page, 2006; Bloom, Miller, & Azurdia, 2007; Le Blanc, Miller, Martinson, & Azurdia, 2007; Martinson, & Hendra, 2006; Navarro, van Dok, & Hendra, 2007). Fifteen study outcomes that were included in the original meta-analysis were excluded from this one. These were contextual outliers that would likely diminish meta-analytic power. They were excluded for one or more of the following reasons: not U.S. (two Canadian study outcomes), accomplished prior to 1980 (five study outcomes), sampled teens only (one outcome) or focused on men (four outcomes), were very small (two studies had less than 100 participants), followed participants for less than a year (three outcomes) or produced a single gross intervention effect estimate across numerous aggregated places (five outcomes) (Bell & Orr, 1994; Carcagno, Cecil, & Ohls, 1982; Friedman, et al., 1980; Goldman,

Friedlander, & Long, 1986; Gordon & Agodini, 1999; Lerman, et al., 1986; Masters, 1981; Maynard, Nicholson, & Rangarajan, 1993; Michalopoulos, Card, Gennetian, Harknett, & Robins, 2000; Nelson, 1984; Puma & Burstein, 1994; Quets, Robins, Pan, Michalopoulos, & Card, 2002; Wilkinson & Ross, 1972; Zambrowski, Gordon, & Berenson, 1993). A total of 65 independent study outcomes were included in this meta-analysis.

Meta-Analysis: Labor Market Withdrawal and Child Care Provision Within Follow-Up Strata

The original meta-analysis explored both fixed and random effects models whose pooled findings were essentially identical (Smedslund et al., 2006). That is, their pooled risk ratios typically differed by only a few thousandths of a decimal place. This meta-analysis, on the other hand, specifically contrasted study outcomes by key methodological (short-term vs. long-term follow-up), contextual (relatively low vs. high labor market withdrawal among African American men in the study area) and intervention program (child care provided or not) characteristics. Therefore, it used fixed effects models as it assumed substantial homogeneity of intervention effects within specific categories of interest. None of the original meta-analysis' explorations of 31 bivariate moderations of the overall employment effect by all available participant, methodological and contextual characteristics (see Tables 9, 13 and 14, Smedslund et al., 2006) accounted for length of follow-up (LFU). Because it essentially averaged the moderation of effects of such disparate studies as those that, for example, followed participants for merely one year with those that followed participants for five years, it confounded the effects of each potential moderator with follow-up length. The present multivariate meta-analysis aimed to solve these problems by focusing on two key moderators—labor market withdrawal and child care provision—and testing their intervention moderating effects within LFU strata: short-term and long-term.

Procedural definitions of job and child care availabilities. The prevalence of labor force nonparticipation among adult African American men in each study area (metropolitan area, county or state) and time period was taken from the nearest population census, 1990 or 2000 (U.S. Bureau of the Census, 2008). Regardless of each study's conceptual definition of child care, this meta-analytic study's procedural definition was based on the empirical comparison of prevalent child care use among intervention and control groups. When the study's intervention group was significantly ($p < .05$) more able to use child care, it was coded 'yes' on child care provision. Otherwise, it was coded 'no.' When such data was not available in the primary study reports, similar to previous meta-analyses, conceptual definitions were used. This review's child care definition differed from the Campbell Collaboration review's in six of 65 instances. These variables were abstracted and coded from primary study reports by two raters whose initial agreement was 88.5%. All inter-rater disagreements were then resolved through discussion of differences and recoding of any discordant study reports.

Meta-analysis. Study effects (natural logarithm of their relative risks [RR]) were weighted by their inverse variances, computed from estimated standard errors ($1 / SE^2$) so that larger, more precise studies, weight more. Such precision weighted effects were then pooled within key categories using weighted regression models. Pooled relative risks within 95% confidence intervals (CI) were calculated from regression statistics as were tests of heterogeneity within pooled groups (χ^2) and comparisons between groups (z) (Greenland, 1987; Grizzle, Starmer, & Koch, 1969). Sensitivity explorations found that the overall pooled intervention effect (RR = 1.04 [95% CI 1.03, 1.05], 65 study outcomes) differed significantly between the following two groups: followed for less than four years (RR = 1.06 [1.05, 1.07], 48 study outcomes) or for four or more years (RR = 1.03 [1.02, 1.04], 17 study outcomes), $z = 1.73$, $p <$

.05. Therefore, the moderation of intervention effects by labor market withdrawal and child care provision will be tested separately within theses follow-up strata: short-term (1-3 years) and long-term (4-6 years).

Practical policy significance indices. It is well known among experienced meta-analysts that such aggregations of relatively large numbers of studies with relatively large numbers of participants typically render the assessment of statistical significance nearly moot. The overall intervention effect estimate, even if practically miniscule, is bound to be deemed significant in a statistical sense (i.e., 95% CIs will very likely not include the null RR value of 1.00). So in addition to pooled RR estimates, a number of other intervention effect size indices were calculated to shed as much light as possible on the practical policy significance of this study's findings. First, preventive fractions (PF) were calculated from each pooled RR (PF = [RR – 1] / RR, Miettinen, 1974). In this study's context, this may be defined as the proportion of undesirable outcomes (chronic joblessness) likely prevented by welfare-to-work interventions. Next, the number needed to treat (NNT) was calculated. The NNT is the inverse of the absolute risk reduction (ARR) where the ARR is the difference between the intervention and control group event rates (Laupacis, Sackett, & Roberts, 1988). In this instance it is an estimate of the number of intervention participants needed to produce one employed participant. For example, if closely corresponding to this field's overall pooled effect, a given welfare-to-work intervention resulted in a 55% employment rate among its participants, one would need to enrol 50 initially unemployed participants to produce one additional employed participant (NNT = 1 / [.55- .53], 53% was the aggregate event or employment rate among the control group at follow-up). This metric can be helpful in thinking about intervention cost/benefit ratios. And finally, average annual intervention-control group earnings differences in 2006-equivalent dollars were

calculated (Officer & Williamson, 2007). The practical life space significance of this metric among typically low wage earners attempting to exit welfare seems obvious.

Results

Sample Description

This review of the practical impacts of labor market forces and child care on the short- and long-term effectiveness of welfare-to-work interventions in America is based on 65 study outcomes of employment initiated between 1982 and 2002, followed until 2005 (see Table 1). Most of this research was accomplished in urban areas. Three-quarters of the samples were city-based (75.4%) with the majority of the statewide samples (23.1%) also representative of urban areas. In typical study areas that were dominated by northeastern, western and mid-western metropolitan areas, approximately a third of all African American men had withdrawn from the labor force (Mdn = 34.2%). And boding well for this meta-analysis' power, such labor force nonparticipation varied quite widely, from less than 10% in certain study areas to more than half in some others.

As for the welfare-to-work interventions, nearly all were mandated (87.7%) with some employment component (98.5%, typically job search or time-limited job experience). Sanctions were enacted in the majority of the studies (61.5%), typically affecting one of every five program participants. Time limits and financial incentives were incorporated less often (27.7%), more typically used in post-TANF (1996) experiments. And nearly a third of the studied interventions provided tangible child care assistance (30.8%), either in-house, through additional subsidies and/or active case management and referral to existing resources. Again, such child care categorical variance served to ensure adequate meta-analytic subsamples for analysis. Finally, consistent with AFDC and TANF service missions, most of the nearly 350,000 people who

participated in these experiments were women (93.0%, typically with two children). Two-thirds of the aggregate review sample was comprised of racial or ethnic minority people, predominantly African Americans and Hispanic people, who in fact, were the majority in most of the studies (84.1%). Clearly, this study's aggregate sample seems highly representative of the urban underclass who received welfare assistance during an extended era of social welfare transitions in America.

Moderating Effects of Job and Child Care Availability

A significant meta-analytic interaction—LMW by LFU—is depicted in the top of table 2. The moderating effect of LMW differed by LFU strata. Modest intervention effects were consistently observed for short-term evaluations whether labor market withdrawal among African American men in study areas was relatively low or high, respectively, relative risk (RR) = 1.06 (95% CI 1.05, 1.07) and RR = 1.05 (1.03, 1.07), $z = 0.40$, NS. This effect was only maintained in long-term evaluations in study areas where LMW was relatively low (RR = 1.07 [1.04, 1.10]), whereas, there was no longer evidence of any significant employment effect of welfare-to-work interventions in high LMW study areas (RR = 0.98 [0.96, 1.01]), $z = 4.68$, $p < .05$. When the focus was on an arguably well-resourced homogeneous program such as GAIN, that incidentally provided child care support, a similar intervention effect moderation by LMW was observed even in the short-term: low LMW (RR = 1.18 [1.14, 1.22]) and high LMW (RR = 1.09 [1.02, 1.16]), $z = 2.18$, $p < .05$. And though long-term GAIN evaluations were not reported, the aggregate evidence suggests that GAIN's effects have probably been maintained in the majority of the areas studied (three-quarters of the study areas were low LMW areas).

A significant main meta-analytic effect of child care is depicted on the bottom of table 2. It did not appear to interact with LFU as its protective effects were similar in the short- and long-

term. Welfare-to-work interventions that provided child care support were significantly more effective than those that did not provide such support less than 4 years (RR = 1.11 [1.09, 1.13] vs. RR = 1.03 [1.02, 1.04]), $z = 7.54$, $p < .05$) and up to six years later (RR = 1.11 [1.06, 1.17] vs. RR = 1.02 [1.00, 1.04]), $z = 3.26$, $p < .05$). As for practical significance, they were three times as likely to prevent chronic unemployment (9.9% vs. 2.9%) and to likely do so with lower cost/benefit ratios. It seems that programs that do not provide child care support need to enrol nearly four times as many participants as do programs with child care to prevent each incidence of chronic joblessness (62 vs. 17 participants) and that such child care-supported welfare recipients end up earning significantly more when they exit welfare to a job.

A child care by LMW interaction was not observed for short-term evaluations (Table 3). Child care had a similar protective impact in study areas where the prevalence of LMW among African American men was relatively low (RR = 1.12 [1.10, 1.14] vs. RR = 1.03 [1.02, 1.04]), $z = 6.41$, $p < .05$) or high (RR = 1.12 [1.09, 1.15] vs. RR = 1.02 [1.01, 1.03]), $z = 4.60$, $p < .05$). Of the four effect distributions displayed in table 3, only one, child care not provided within high LMW study areas, was significantly heterogeneous, χ^2 (9) = 37.14, $p < .05$. Explorations found a significantly larger intervention effect among studies of older participants (two study outcomes, mean age 32 and older, RR = 1.11 [1.07, 1.15]) as compared with those of younger participants (eight study outcomes, mean age less than 32, RR = 1.01 [1.00, 1.02]), $z = 2.49$, $p < .05$. No other participant, intervention program, contextual or study methodological characteristic was significantly related to this outcome variability.

Discussion

This meta-analysis of 65 large RCT outcomes of welfare-to-work interventions in America systematically replicated the overall precision weighted finding of a previous one

(Smedslund et al., 2006). Both estimated that, on average, such interventions probably diminish the risk of chronic unemployment among welfare recipients by only four percent. The previous Campbell Collaboration review provocatively observed that local study contexts seemed to influence outcomes more than did the programs themselves, but could not adequately account for such study context variability in its analyses. Improving key variable validities and extending the sample of studies well into the post-TANF era, this review generated evidence strongly suggesting that the availability of jobs in the areas studied and the provision of child care support with welfare-to-work interventions accounts for most such contextual outcome variability. In fact, this review clearly demonstrated that such job and child care resources were not only more influential than the welfare-to-work interventions themselves, but that they probably accounted for nearly all apparent program successes. For example, aggregate long-term intervention effects in areas of hypothesized low job availability (i.e., high LMW areas) or of programs that did not provide any additional child care support were null, and their combined short-term effects (i.e., high LMW and no child care) were nil, that is, they had a statistically significant though practically near meaningless impact on employment.

Interpretations of the generally small effects of welfare-to-work interventions by this and previous of this field's meta-analyses ought to consider their typically, quite conservative, comparator; that is, aggregate control group members who typically received usual welfare program—AFDC or TANF—services. Recall that more than half of this meta-analysis' control group (53%) had exited welfare to a job by the time of aggregate study end-points. During a period of economic growth, the majority of welfare recipients found jobs without benefit of welfare-to-work program mandates, sanctions or typically meagre training and work experience. This finding not only provided a resounding refutation of the welfare dependency myth, but also

a reasonable defense for use of a usual welfare service control group. Given this meta-analysis's focus—the effects of environmental, social-economic, forces—it seems to have great policy utility. It essentially provided precise control for the personal behavioural characteristics of welfare recipients themselves, which were, after all, the target of many welfare reform policies. This study's central findings suggested that they are, in fact, of near inconsequence. Whereas, social forces seem clearly to be of great consequence. Social policies, the economy and even the physical contexts of contemporary urban environments shape the availability and accessibility of employment and other opportunities (Coulton, 2003; Kain, 1992; Kasarda, 1989; Wilson, 1987; 1996; Ziliak, Figlio, Davis, Connolly, 2000). This study's findings are consistent with the notion that welfare recipients are as likely as other Americans to avail themselves of all such opportunities that are truly available to them. Policy makers grappling with challenges in the future would better serve all Americans by focusing on so shaping social, economic and physical environments, rather than the people that live in them.

Possible Meta-Analytic Review Limitations

An asymmetrical "trimmed and filled" funnel plot (Duval & Tweedie, 2000) of study standard errors by the natural logarithm of their employment relative risks provided some indication of publication bias in the Campbell Collaboration review (Smedslund et al., 2006). The same was true of this review. So it is likely that a few of the smaller studies with larger, statistically significant, effects were published, whereas, some null findings may not have been. It should be noted though that any such bias that may have intruded would serve to make this study's aggregate small effect (e.g., RR = 1.04), if anything, an overestimate of the truth (de Smidt & Gorey, 1997; Grenier & Gorey, 1998). None of the primary studies of the employment effectiveness of welfare-to-work interventions, however, were designed to test the specific effect

modifiers that this meta-analysis did: labor market withdrawal and child care provision. For this and the following reasons publication bias seems highly unlikely to be a potent alternative explanation for this review's findings: (a) sample sizes were quite large (only one was less than 500) and did not differ significantly between key strata of meta-analytic variables (LFU, LMW or child care), (b) sample sizes were not significantly associated with effect sizes within such key strata, and (c) exemplary fail-safe Ns at $p < .05$ for two key findings (13 study outcomes of child care provision in low LMW areas and 5 such outcomes in high LMW areas, Table 3), were found to be 198 and 34, respectively (Rosenthal, 1979). These are the estimated number of studies with null findings that would have to exist in file drawers to change this review's central conclusions. These fail-safe Ns were approximately 15 and 7 times the number of respective study outcomes included in this review, so it seems highly resistant to the potential impact of unretrieved null results.

Conclusion

Child care clearly seems the key element of welfare-to-work programs that work. It ought to be integral in any future interventions to assist welfare recipients or other families at similar risk of joblessness. And it seems equally clear that it is the availability of good jobs that ultimately matters most. Proactive social economic policies that positively affect such job availability will always be more effective than reactive social welfare policies that aim to affect the jobless themselves.

References

Axelsen, D., Friesner, D., Rosenman, R., & Snarr, H. (2007). Welfare recipient work choice and in-kind benefits in Washington state. Applied Economics, 39, 1021-1036.

Bell, S. H., & Orr, L. L. (1994). Is subsidized employment cost-effective for welfare recipients? Experimental evidence from seven states. Journal of Human Resources, 29, 42-61.

Bloom, D., Hendra, R., & Page, J. (2006). The Employment Retention and Advancement project: Results from the Chicago ERA site. Washington: U.S. Department of Health and Human Services, and the Manpower Demonstration Research Corporation.

Bloom, D., Miller, C., & Azurdia, G. (2007). The Employment Retention and Advancement project: Results from the Personal Roads to Individual Development and Employment (PRIDE) program in New York City. Washington: U.S. Department of Health and Human Services, and the Manpower Demonstration Research Corporation.

Bloom, H. S., Hill, C. J., & Riccio, J. A. (2003). Linking program implementation and effectiveness: Lessons from a pooled sample of welfare-to-work experiments. Journal of Policy Analysis and Management, 22, 551-575.

Carcagno, G. J., Cecil, R., & Ohls, J. C. (1982). Using private agencies to place public assistance clients in jobs. Journal of Human Resources, 17, 132-143.

Carnochan, S., Ketch, V., De Marco, A., Taylor, S., Abramson, A., & Austin, M. J. (2005). Assessing the initial impact of welfare reform: A synthesis of research studies (1998-2002). The Social Policy Journal, 4, 3-31.

Cochi Ficano, C. K., Gennetian, L. A., & Morris, P. A. (2006). Child care subsidies and employment behavior among very-low-income populations in three states. Review of Policy Research, 23, 681-698.

Coulton, C. J. (2003). Metropolitan inequalities and the ecology of work: Implications for welfare reform. Social Service Review, 77, 159-190.

de Smidt, G. A., & Gorey, K. M. (1997). Unpublished social work research: Systematic replication of a recent meta-analysis of published intervention effectiveness research. Social Work Research, 21, 58-62.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta analysis. Biometrics, 56, 455-463.

Friedman, B., Davenport, B., Evans, R., Hahn, A., Hausman, L., & Papirno, C. (1980). An evaluation of the Massachusetts Work Experience program. Washington: U.S. Department of Labor, ManPower Administration.

Goldman, B., Friedlander, D., & Long, D. (1986). Final report on the San Diego Job Search and Work Experience demonstration. New York: Manpower Demonstration Research Corporation.

Gordon, A., & Agodini, R. (1999). Early impacts of the Virginia Independence program. Final report submitted to the Virginia Department of Social Services. Princeton, NJ: Mathematica Policy Research, Inc.

Gorey, K. M. (November, 2003). Caseload and lifespace effects of Canadian and American welfare reforms. Paper presented at the 131st annual meeting of the American Public Health Association, Social Work Section. San Francisco.

Greenberg, D., Ashworth, K., Cebulla, A., & Walker, R. (2005). When welfare-to-work programs seem to work well: Explaining why Riverside and Portland shine so brightly. Industrial and Labor Relations Review, 59, 34-50.

Greenberg, D. H., & Shroder, M. (2004). The digest of social experiments (3rd ed.). Washington, DC: Urban Institute Press.

Greenland, S. (1987). Quantitative methods in the review of epidemiologic literature. Epidemiologic Reviews, 9, 1-30.

Grenier, A. M., & Gorey, K. M. (1998). Effectiveness of social work with older people and their families: A meta-analysis of conference proceedings. Social Work Research, 22, 60-64.

Grizzle, J. E., Starmer, C. F., & Koch, G. G. (1969). Analysis of categorical data by linear models. Biometrics, 25, 489-504.

Holzer, H. J. (in press). The labor market and young black men: Updating Moynihan's perspective. Annals of the American Academy of Political and Social Science.

Kain, J. K. (1992). The spatial mismatch hypothesis: Three decades later. Housing Policy Debate, 3, 371-462.

Kasarda, J. D. (1989). Urban industrial transition and the underclass. Annals of the American Academy of Political and Social Science, 501, 26-47.

Laupacis, A., Sackett D. L., & Roberts, R. S. (1988). An assessment of clinically useful measures of the consequences of treatment. New England Journal of Medicine, 318, 1728-1733.

Le Blanc, A., Miller, C., Martinson, K., & Azurdia, G. (2007). The Employment Retention and Advancement project: Results from Minnesota's tier 2 program. Washington: U.S. Department of Health and Human Services and Manpower Demonstration Research Corporation.

Lerman, R., Friedman, B., Ajemian, S., Fairchild, C. K., Jastrzab, J., Kulik, J. et al. (1986). Food Stamp Work Registration and Job Search demonstration. Unpublished manuscript.

Lichter, M. I., & Oliver, M. L. (2000). Racial differences in labor force participation and long-term joblessness among less-educated men. In L. D. Bobo, M. L. Oliver, J. H. Johnson, Jr.,

& A. Valenzuela (Eds.), Prismatic metropolis: Inequality in Los Angeles (pp. 220-248). New York: Russell Sage Foundation.

Martinson, K., & Hendra, R. (2006). The Employment Retention and Advancement project: Results from the Texas ERA site. Washington: U.S. Department of Health and Human Services, and the Manpower Demonstration Research Corporation.

Masters, S. (1981). The effects of supported work on the earnings and transfer payments of its AFDC target group. Journal of Human Resources, 16, 600-636.

Maynard, R. A., Nicholson, W., & Rangarajan, A. (1993). Breaking the cycle of poverty: The effectiveness of mandatory services for welfare-dependent teenage parents. Princeton, NJ: Mathematica Policy Research, Inc.

Michalopoulos, C., Card, D., Gennetian, L., Harknett, K., & Robins, P. K. (2000). The Self-Sufficiency Project at 36 months. Ottawa: Social Research and Demonstration Corporation.

Miettinen, O. S. (1974). Proportion of disease caused or prevented by a given exposure, trait or intervention. American Journal of Epidemiology, 99, 325-332.

Navarro, D., van Dok, M., & Hendra, R. (2007). The Employment Retention and Advancement project: Results from the Post-Assistance Self-Sufficiency (PASS) program in Riverside, California. Washington: U.S. Department of Health and Human Services, and the Manpower Demonstration Research Corporation.

Nelson, H. (1984). Evaluation of the Community Work Experience program. Olympia, WA: Washington State Department of Social and Health Services.

Officer, L. H., & Williamson, S. H. (2007). Purchasing power of money in the United States, 1790-2006. MeasuringWorth.com

Puma, M. J., & Burstein, N. R. (1994). The national evaluation of the Food Stamp Employment and Training program. Journal of Policy Analysis and Management, 13, 311-330.

Quets, G., Robins, P., Pan, E., Michalopoulos, C., & Card, D. (2002). Does SSP Plus increase employment? The effect of adding services to the self-sufficiency project's financial incentives. Ottawa: Social Research and Demonstration Corporation.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. Psychological Bulletin, 86, 638-641.

Smedslund, G., Hagen, K. B., Steiro, A., Johme, T., Dalsbø, T. K., & Rud, M. G. (2006). Work programmes for welfare recipients. Campbell Collaboration Review.

U.S. Bureau of the Census (2008). Census of population and housing, 1990 and 2000 (summary tape file 3 and summary file 4 online). Washington, DC: Data User Services Division.

Wilkinson, K. P., & Ross, P. J. (1972). Evaluation of the Mississippi AFDC experiment. Social Service Review, 46, 363-377.

Wilson, W. J. (1987). The truly disadvantaged: The inner city, the underclass, and public policy. Chicago: The University of Chicago Press.

Wilson, W. J. (1996). When work disappears: The world of the new urban poor. New York: Alfred A. Knopf.

Zambrowski, A., Gordon, A., & Berenson, L. (1993). Evaluation of the Minority Female Single Parent Demonstration: Fifth-year impacts at Centers for Employment Training (CET). Princeton, NJ: Mathematica Policy Research, Inc.

Ziliak, J. P., Figlio, D. N., Davis, E. E., & Connolly, L. S. (2000). Accounting for the decline in AFDC caseloads: Welfare reform or the economy? Journal of Human Resources, 35, 570-586.

Table 1
Primary Study Sample Descriptive Profiles: Percentage Distributions Across 65 Study Outcomes

Contexts	%	Interventions[b] and Participants	%
Year Begun[a]		Child Care Provided	30.8
1982 to 1989	29.2	Mandated	87.7
1990 to 1999	60.0	Sanctions enacted	61.5
2000 to 2002	10.8	Time limits	27.7
Mdn = 1992, M = 1992, SD = 4.85		Financial incentives	27.7
Maximum Length of Follow-Up (Years)		Intervention Focus	
1	15.4	Employment alone	41.5
2 to 4	61.5	Employment and education	57.0
5 to 6	23.1	Education/training alone	1.5
Mdn = 2.00, M = 2.79, SD = 1.42			
Region of the United States		Intervention and Control Sample	
Mid-west	29.2	444 to 999	10.8
West and northwest	26.3	1,000 to 2,999	27.7
Northeast	18.5	3,000 to 4,999	32.3
Southwest	13.5	5,000 to 9,999	13.8
South and southeast	12.3	10,000 to 29,795	15.4
		Mdn = 3,939, M = 5,373, SD = 5264	
Urbanity		Sample Percentage Female	
City-based	75.4	90 to 100%	68.3
Statewide	23.1	80 to 89%	25.4
Rural	1.5	66 to 79%	6.3
		Mdn = 93.00, M = 91.80, SD = 7.13	
African American Men 16 Years of Age And Over Not in Labor Force		Sample Percentage Minority	
8.7 to 33.2%	36.9	75 to 100%[c]	38.5
33.3 to 39.9%	47.7	50 to 74%[c]	44.6
40.0 to 63.3%	15.4	2 to 49%	16.9
Mdn = 34.20, M = 34.37, SD = 7.56		Mdn = 65.00, M = 64.43, SD = 21.30	

[a] Twelve study samples (18.5%) were during the post-TANF era (i.e., 1996 or later).
[b] Thirty-three study outcomes were from three welfare-to-work programs: Greater Avenues to Independence (GAIN, 7), Job Opportunities and Basic Skills Training/National Evaluation of Welfare-to-Work Strategies (JOBS/NEWWS, 19) and the Employment Retention and Advancement (ERA, 7) Project. The remaining 32 outcomes were from 25 other programs.
[c] Of 54 such study samples, 39 were predominantly African American and 15 Hispanic.

Table 2
Summary of Welfare-to-Work Intervention Effects Moderated By Length of Follow-up And By Key Labor Market and Intervention Characteristics

	Followed 1-3 Years		Followed 4-6 Years	
	Prevalence of African American Men Not in Labor Force[a]			
	Low*	High*	Low	High
Study outcomes	33	15	14	3
Total participants	156,481	84,628	63,272	44,837
Risk Ratio (95% CI)	1.06 (1.05,1.07)	1.05 (1.03,1.07)	1.07 (1.04,1.10)	0.98 (0.96,1.01)
Preventive fraction	5.6%	4.8%	6.5%	NA
Number needed to treat	31	37	27	NA
Earnings difference (N)[b]	$660 (30)	$730 (14)	$935 (7)	NA
	GAIN: Greater Avenues to Independence Program			
	Prevalence African American Men Not in Labor Force[c]			
	Low	High		
Study outcomes	5	2		
Total participants	32,128	6,410		
Risk Ratio (95% CI)	1.18 (1.14,1.22)	1.09 (1.02,1.16)		
Preventive fraction	15.3%	8.3%		
Number needed to treat	10	21		
Earnings difference (N)	$1,985 (5)	$870 (2)		
	Child Care Provided as Part of the Intervention			
	Yes*	No*	Yes	No
Study outcomes	18	30	2	15
Total participants	72,372	168,737	7,573	100,536
Risk Ratio (95% CI)	1.11 (1.09,1.13)	1.03 (1.02,1.04)	1.11 (1.06,1.17)	1.02 (1.00,1.04)
Preventive fraction	9.9%	2.9%	9.9%	NA
Number needed to treat	17	62	17	NA
Earnings difference (N)	$1,130 (16)	$425 (28)	$1,055 (2)	NA

[a] Highest quartile (high [38% or more not in labor force]) compared with the lower three quartiles (low [less than 38% not in labor force]).
[b] Annual intervention-control group earnings difference in 2006-equivalent dollars (N = number of studies reporting earnings).
[c] Highest quartile (high [40% or more not in labor force]) compared with the lower three quartiles (low [less than 40% not in labor force]).
* Distribution significantly heterogeneous (χ^2 statistic), $p < .05$.

Table 3
Summary of Welfare-to-Work Intervention Effects Moderated By Child Care Provision Within Levels of Labor Market Nonparticipation: Followed Less Than Four Years

	Prevalence of African American Men Not in Labor Force[a]			
	Low		High	
	Child Care Provided as Part of the Intervention			
	Yes[b]	No[b]	Yes	No[*]
Study outcomes	13	20	5	10
Total participants	56,751	99,730	15,621	69,007
Risk Ratio (95% CI)	1.12 (1.10,1.14)	1.03 (1.02,1.04)	1.12 (1.09,1.15)	1.02 (1.01,1.03)
Preventive fraction	10.7%	2.9%	10.7%	2.0%
Number needed to treat	16	62	16	94
Earnings difference (N)[c]	$1,090 (11)	$400 (19)	$1,225 (5)	$450 (9)

Note. There was not enough meta-analytic power to test the 2-way child care by LMW interaction among long-term studies or the 3-way child care by LMW by LFU interaction. One analytic cell was empty (followed 4-6 years, high LMW and no child care).

[a] Highest quartile (high [38% or more not in labor force]) compared with the lower three quartiles (low [less than 38% not in labor force]).

[b] Excluding one study outcome, not significantly heterogeneous.

[c] Annual intervention-control group earnings difference in 2006-equivalent dollars (N = number of studies reporting earnings).

[*] Distribution significantly heterogeneous (χ^2 statistic), $p < .05$.