Effects of caseload and wages on the child maltreatment substantiation rate

Sofia Guo*

December 2018

Abstract

This regression analysis uses average caseload and wage data and controls for state child poverty to find that changes in the average caseload of child welfare workers do not have significant effects on changes in the rate at which workers substantiate reports of child maltreatment. Meanwhile, changes in child poverty rates have an estimated -0.42% effect on changes in substantiation rates. Contrary to expected outcomes, changes in the average hourly wage of child welfare workers have a significant -0.95% effect on changes in the substantiation rate. This intriguing effect is potentially due to correlated effects of increased funding on other aspects of the state-level child welfare system and warrants more research.

*Special thanks to Professor Jill Berrick for her valuable input and reccomendations, Dr. Matty Litchenstein for inspiration, encouragement and guidance, and Garrett Therolf for extremely helpful advice and excellent sources.

I. Introduction

Since its incorporation into the federal government during the early 1900's (Stretch 1970), the American child welfare system has evolved from an institution focused on combating the effects of poverty on children to a hotly contested ideological platform where social workers' rights and abilities to determine a child's safety are called into question (Therolf 2018). While landmark reforms have been passed in the last century to refine state governments' jurisdictions regarding child protection (Edwards 2017), child welfare workers face significant ambiguity and adverse work conditions which hinder sustainable careers, produce high turnover rates, and increase burnout sentiment (Hamama 2012). Most critically, child welfare workers who investigate or assess reports of child maltreatment (Child Protective Services, abbreviated CPS), carry a heavy responsibility to decide whether a report indicates evidence of maltreatment (that it is substantiated) or does not have sufficient evidence or meet the legal definition of abuse or neglect (Children's Bureau 2013; Jones 1993).

Given large caseloads, low wages, lack of job security, and resource constraints on CPS workers (Getzel 1983), those who enter the field are often driven by ideological or personal factors such as religion (Therolf 2018), not conventionally predictive factors like earnings potential and education levels (Barth, et al. 2008) common in other careers. To further understand how to improve a system that affects the seven million children a year involved in maltreatment reports (Children's Welfare League of America 2018), more research on improving work incentives, working conditions and the quality of care by American child welfare professionals is required.

This paper aims to add to this growing literature by uncovering the purely empirical relationships between child welfare worker characteristics and the substantiated disposition rate of child maltreatment reports. By measuring child welfare worker characteristics using average hourly wage and caseload while controlling for the state child poverty rate and entity and time fixed effects, this paper strives to answer the question:

1. On average, do higher paid and less-burdened child welfare workers substantiate more or less child maltreatment reports, and why?

I use a multiple linear regression method (MLR) to estimate the coefficients on worker caseloads and their hourly wage to determine the economic and statistical significance of these factors on the substantiation rate in 49 states plus the District of Columbia from 1999-2010.[1] The caseload and substantiation rate data are from the annual Child Maltreatment reports published by the Children's Bureau, while the wage data are from the Occupational

[1] Colorado is omitted due to having no reported caseload data for the specified years; 2003 is omitted due to missing caseload and substantiation data.

Employment Statistics (OES) survey via the Bureau of Labor Statistics; the child poverty rates by state are from the Small Income Area and Poverty Estimates (SAIPE) program by the U.S. Census Bureau.

This series of 11 years was selected on the basis of data available from the Children's Bureau on worker caseloads and substantiation rates, as the other three data sources had much wider breadth of data available. I control for state child poverty rates due to the strong positive association between poverty and child maltreatment (Finck, et al. 2017), in an attempt to normalize any bias caused by states with systemically higher child poverty rates. As seen in Graph 1 below, there is not much variation in any of the explanatory variables except caseload, which fluctuates wildly in some states while displaying a very spotty presence in others.[2]

This inconsistent variation is a potential source of bias in my MLR estimation; however, the intuition behind including wages and caseload as the main regressors is supported by previous research, given that the correlation between any of the explanatory variables and the error term (which could include worker attitude) is most likely zero. This intuition is built on the aforementioned observation that child welfare workers are not motivated by wages or workload to conduct their jobs a certain way.

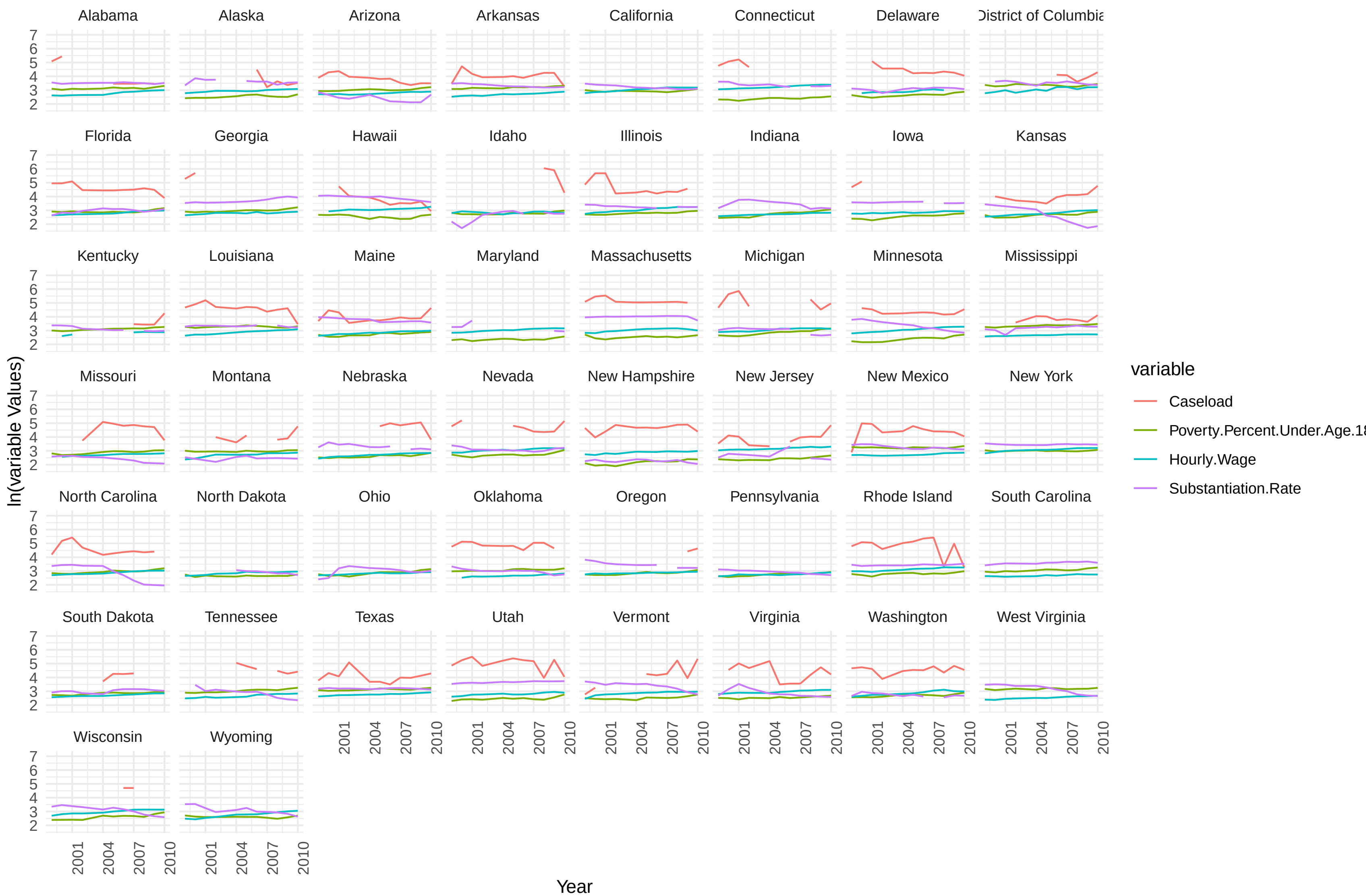
The key results indicate that changes in worker caseload are not significant determinants of changes in substantiation rates, but for every percent increase in hourly wages, the predicted change in substantiation rates decreases by 1%. This finding is most surprising in light of the assumption that given the industry's low pay, wage differentials should not be a driver of decision making among child welfare workers. Meanwhile, the child poverty rate is a statistically significant control variable that decreases the substantiation rate by 0.42% for every percent increase in the child poverty rate. This is also surprising given the evidence that higher poverty rates should increase the number of substantiated maltreatment cases. I use two robustness checks on my model, one by taking the logs of all variables after applying state and time fixed effects, and the other by omitting the five states with the highest average substantiation rates over 1999-2010 (MA, ME, GA, HI, and UT) from the original regression, both which my chosen model withstands.

In section II, I discuss the relevance of the topic in-depth, with more information on the relevance of my control variable and the existing literature regarding research on child welfare workers and its current limitations. In section III, I outline the proposed main regression model and elaborate on the details of the data sources used in the paper. Section IV drills into the mixed model results for the paper and connects sources to potential explanations for

[2] All caseload and substantiation data are state-reported to the Children's Bureau, thus inconsistencies or lack of data over substantial periods of time are common in some states.

these estimates; I also explain the robustness checks. Finally, section V concludes the paper by proposing policy implications and what further research can be done on this subject. The Appendix in section VI contains some additional graphs and regression results which may assist future research on this topic.

Graph 1: Natural logs of variable values from 1999−2010*



*Excludes 2003; data from BLS, Census.gov, Children's Bureau

II. Literature Review

In the face of growing economic uncertainty and low economic mobility for much of the population in the U.S., a large portion of the future workforce may face much higher risks of child maltreatment in the coming decade. With rates of reported child maltreatment rising slowly but steadily in recent years (Figure 1), one can begin to imagine the role an inadequately staffed child welfare system can play in magnifying these adverse effects on the nation's young. It is important to understand the context under which the current child welfare system formed before attempting to grasp the scope of child welfare worker issues today.

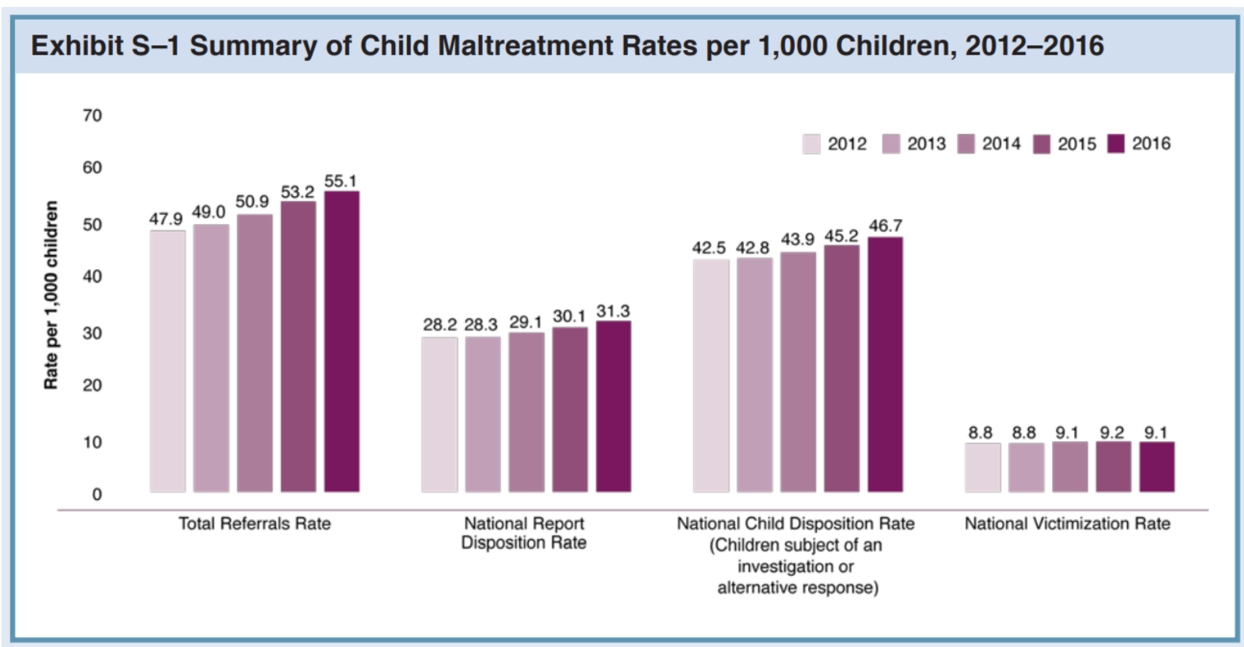


Figure 1: Child Maltreatment Report from "Child Maltreatment 2016", Children's Bureau

The concept of child welfare originated in the late 19th century, when destitute children were placed into orphanages because their parents could not provide basic needs like food and shelter (Finck, et al. 2017). Most of these orphanages were tended to by religiously motivated groups, who believed that they had a moral responsibility to take care of the poor. The federal government did not step in until 1909 when President Theodore Roosevelt held the first White House Conference on Children (Stretch 1970), which established key institutions such as the Federal Children's Bureau and gave state governments the right to regulate all agencies caring for "dependent children", or those not under the care of their parents likely due to economic reasons. Over the next few decades, a scaffolding was built on federal funding via Social Security and state block grants, which have culminated into a complicated national system of child welfare today, largely dictated at the state level with varying approaches to regulating families.

Most interestingly, while this regulatory system originally designed to address child poverty has grown in size and complexity, the child poverty rate has not diminished accordingly; from 1959 to 2012, there was only a 5% decrease in the average estimated national childhood poverty rate (Finck, et al. 2017). The focus on poverty as a driving factor for increased risk of child maltreatment is backed by many studies; children who grow up in financially unstable households are likelier to experience immediate and long-term psychological challenges, in addition to elevated risk of family violence and maltreatment (Pryor, et al. 2018; Finck, et al. 2017; Page, Spetz and Millar 2005; Garbarino 1989). In order to successfully isolate the effects of worker characteristics on subststantiation rates, my reasoning for including child poverty as a control variable for the substantiation rate is thus supported.

Since the child welfare system is deeply rooted in attempting to regulate poor children and their families, the question of worker bias in deciding the fate of children becomes unavoidable. Tying to the original question of what determines how a child welfare worker evaluates the merits of a maltreatment report, the judgement of workers who have particular ideas of what families should look or act like under "normal" conditions becomes extremely subjective, especially if they associate images of poverty with moral judgement; this ideological aspect of decision making can lead to anything from separating young children from their families without compelling reasons to the death of a child from lack of intervention (Therolf 2018).

While some research points out that poor families are truly in need of more services than better-off families (Jonson-Reid, Drake and Kohl 2008), other studies point to racial bias as a major source of increased propensity to substantiate maltreatment cases which would otherwise be diverted to other forms of service such as financial aid programs (Detlaff, et al. 2011). The complexity of factors driving a child welfare worker's decision making makes it all the more important to search for empirical evidence to clarify the drivers of this bias.

Literature on the detailed decision making process of CPS workers points to many ambiguities that are left to the worker's personal judgement to substantiate a case. Oftentimes, workers must extrapolate from incomplete evidence in a report, whether it be lack of a parental admission to abuse or no report of severity of injury to the child; this causes workers to be influenced by other factors like parental physical appearance and language fluency in their decision to substantiate, (Jones 1993) which in any other context of decision making may be considered extremely biased.

Workers also face major ambiguities at a fundamental level: the criteria which they are told to follow when substantiating child abuse or neglect. A persistent example of this is the vague nature of state-level inconsistencies in definitions of non-visible harm; based off the general definitions that they must substantiate cases if there is evidence of emotional abuse to the child, workers must make the judgement of emotional or mental injury of the child themselves with no additional information than what impression the child gives, if any (Cohen and Sussman 1975). Compounding this problem is the fact that many traumatized children are unable or afraid to accurately recall what happened to them; this is due to the often very young ages of victims limiting their ability to communicate, or fear of their abusers retaliating, who are often parents or close relatives (Jones 1993).

More recent research on exactly which CPS worker characteristics cause higher rates of substantiation points to several factors; younger, less experienced and childless workers who have a history of abuse are more likely to substantiate cases holding all other factors constant. This becomes an issue when one considers that in 2017, the top placement field of bachelor's in social work graduates, often the youngest and least experienced, was child welfare over any other social welfare sector (CSWE 2017). Most importantly, researchers find that the impact of workers' personal attitudes on risk assessments for substantiation is universal across different countries and welfare systems, and that workers in Western countries hold biases against child removal and support family-preservationist ideals due to the cultural history of child welfare in the country (Benbenishty, et al. 2015).

This selective judgement in making CPS decisions was observed early in the history of child welfare and is part of a cultural "group think" attitude that workers participate in today (Therolf 2018). A 1959 study done by Nettler (1958-1959) on child welfare agency staff in Texas revealed that workers were already driven by the concept of "occupational selection and occupational interest", or that their attitudes were reinforced by those around them instead of educational facts or observations. Even decades after Nettler's study, only 49% of the public child welfare workforce held a social work degree in 1981 prior to job entry (Vinokur-Kaplan 1987); the majority of child welfare workers hold a bachelor's or master's degree in other fields, or no degree at all. By today's standards, it is hard to envision more than half a workforce not professionally trained to conduct such demanding work like substantiating cases of maltreatment. While the workforce may be more uniform in its current state, the legacy of prior lack of qualifications may underlie best practices that are still in use due to natural inertia in the field.

Under these circumstances, one begins to understand the why group think, personal ideologies, and reliance on the opinions of higher-ups, instead of more rigorous methods of judgement provided through higher education, has persisted in the child welfare practice well into the 21st century. In addition, child welfare workers report higher levels of psychological stress and anxiety than any other group in social services (Tham and Meagher 2009), leading to a high turnover and burnout rates (Hamama 2012).

There is emerging evidence that this child welfare worker "attitude" factor is the primary driver of the decision-making process in substantiation, instead of factors like the mother's wishes for the child's future having any effect on CPS decisions (Benbenishty, et al. 2015). A recent trans-national study done on CPS workers in California, England, Finland and Norway found that U.S. CPS workers rely heavily on their supervisors, managers and lawyers to approve their decisions regarding cases (Berrick, et al. 2018) as opposed to the needs of families or other lateral sources which are equally available.

These recent findings make crucial headway into the nebulous question of what really drives CPS workers' decisions, but none apply a straightforward economic empirical approach to examining wage and caseload effects. Not only has the research on child welfare workers just begun to grow, it is most often conducted through very specific, survey-based studies done in foreign countries with more organized social worker information than the United States. The most economics-focused work in this area has been conducted on the social worker labor market (M. C. Barth 2003) on uncovering why social workers' wages are depressed; the author finds that low wages are caused by the low proportion of workers who have social work degrees in the field, frequent hiring of non-degreed workers and the effects of steady labor supply from those determined to pursue a career in social work who constitute a large part of the workforce. However, this research does not answer the child welfare specific question of how varying pay and workloads effect substantiation rates. My paper aims to fill the gap in this knowledge on American child welfare worker behavior through an empirical lens, using state-level data to focus on nationwide trends in the 2000's. While this method of research grossly simplifies the complex input factors for substantiation, it also provides an attempt to apply econometrics to a field traditionally dominated by very specific and complex qualitative survey work.

III. Model Outline

This paper's identification strategy uses the Multiple Linear Regression (MLR) method to isolate the effects of changes in caseload and worker wages on changes in the child maltreatment report substantiation rate. By controlling for changes in the rate of child poverty in each state, the proposed model below removes omitted variable bias caused by the effects of increased levels of child poverty on increased numbers of substantiated maltreatment cases through the strong positive correlation of poverty and maltreatment. Additionally, this identification controls for variation among states which is constant over time (such as systematically higher substantiation rates in particular states) and variation over time which is constant across states[3].

Some weaknesses remain with this identification strategy, particularly with the lack of average worker education levels, experience, and personal characteristics (race, gender, socioeconomic background) as main regressors. The argument for including these omitted variables is that they may each be affecting changes in the substantiation rate through the error term and causing the estimated coefficients to become biased and inconsistent. However, evidence shows that this omitted variable bias (OVB) may not be a serious issue. According to a child welfare worker job satisfaction study conducted by Barth, et al. (2008), each of the omitted variables mentioned above have no statistically significant effect on reported job satisfaction; only the quality of supervision received by workers (measured in hours per week) was statistically and economically significant. If one assumes that increased job satisfaction drives more time and effort put into work, which in turn drives a systemic willingness to substantiate more or less reports, there is an argument to be made in favor of omitting these variables due to their lacking effects on workers' decision-making process. Without this assumption, however, it is very much possible that these omitted variables are causing bias in my estimates, a factor to consider when interpreting the results.

The main equation to be estimated is:

$$\begin{aligned}\ln(SubstantiationRate)_{i,t} = \alpha_{i,t} + \beta_{1i,t}\ln(Caseload)_{i,t} + \beta_{2i,t}\ln(HourlyWage)_{i,t} \\ +\beta_{3i,t}\ln(ChildPoverty)_{i,t} + \gamma_i + \lambda_t + u_{i,t}\end{aligned} \tag{1}$$

where γ_i and λ_t are the state and time fixed effects, respectively. The *SubstantiationRate* refers to the number of reported cases that were ruled substantiated by caseworkers divided by total number of reports processed in state t in year i. The $\alpha_{i,t}$ variable is the average substantiation rate in each state and year given zero changes in all the control variables, while the $\beta_{1i,t}$ coefficient measures the percent change in the substantiation rate caused by changes in the average number of cases per year and state assigned to each CPS worker. $\beta_{2i,t}$ measures the percent change in the substantiation rate caused by changes in average hourly wages of child, family and school social workers, and $\beta_{3i,t}$ measures the percent change in the substantiation rate caused by percent changes in the child poverty rate of each state per year.

The methodology for this equation is to use OLS through MLR because it is the best linear

[3] As seen in Graph 1, there is some variation among states that shows steady decreases in the substantiation rate particularly over 2004-2010, which time fixed effects controls for. In addition, almost all states show a slow increase in poverty rates and average hourly wages over the decade; the time fixed effects certainly control for this steady increase to better isolate the coefficients of interest.

unbiased estimator, as long as the MLR assumptions hold true. In this case, we assume that there is no severe OVB as argued above and that the controls and data transformation into logs (countering heteroskedasticity) diminish any bias caused by MLR assumption violations.

I expect the caseload and wage variables to drive changes in the substantiation rate, due to the assumption that heavily loaded workers who are not paid substantially high wages will change the rate at which they decide to pursue cases. This assumption is based on the idea that substantiating cases creates immediate increases in workload down the line for CPS workers; thus, I expect that workers who are paid less and have higher workloads will substantiate less cases to divert additional work off their plates into other channels of the child welfare system. An alternative hypothesis is that neither of these main regressors should have any statistically significant effect on the substantiation rate, given the arguments for worker attitude being the primary driver of substantiation decisions. As controls, the child poverty rate and the state and fixed effects are expected to have statistically significant and positive (for poverty) effects on the substantiation rate, as higher rates of child poverty are strongly associated with higher rates of maltreatment, thus more maltreatment reports in volume should correspond with more substantiated reports holding all else constant.

IIIa. Data

Table 1: Summary statistics (2 outliers removed)

Statistic	Caseload	Poverty.Percent.Under.Age.18	Hourly.Wage	Substantiation.Rate
N	342	548	538	524
Mean	99.61	17.39	17.91	26.01
St. Dev.	85.14	5.11	3.54	10.85
Min	16.00	6.60	10.70	5.54
Pctl(25)	53.00	13.50	15.28	18.71
Pctl(75)	123.00	20.80	19.89	32.64
Max	1,225.00	32.40	29.66	58.76

The data sources used for this regression include (1) substantiation rates and caseload data from the U.S. Children's Bureau, (2) Child, School and Family Social Worker Hourly Wages from the Occupational Employment Statistics from the U.S. Bureau of Labor Statistics, and (3) state child poverty rates from the Small Area Income and Poverty Estimates (SAIPE) data from the U.S. Census Bureau.

The sample size is bounded by the smallest amount of observations, 342, for the approximated worker caseload over the time period 1999-2010 (omitting 2003 for when there is no substantiation data available online) or over 11 years. The dataset is of panel nature, where there are multiple entities (49 states plus the District of Columbia, excluding Colorado for lack of any caseload data) being tracked over multiple periods of time (11 years, excluding 2003, from 1999-2010) and multiple variables (caseload, hourly wage, state child poverty rate).

The caseload data are taken by the reported average number of cases per child protectives services worker each year from the Children's Bureau; the hourly wage is estimated by BLS for "Child, family and school social workers" as a group; and state child poverty rate estimated from the Census Bureau via the SAIPE survey is a percent of the state's children aged 0-17/18 who are under the poverty line. The substantiation rates were calculated by diving the total number of substantiated reports by the total number of maltreatment reports received for each state that year.

Data shortcomings and challenges not already noted include the obvious missing data on caseloads - this is due to the "self reporting" nature of the Child Maltreatment reports published by the Children's Bureau, where the Bureau asks states who have data to report instead of requiring responses[4]. In addition, the original span of data was planned for 1999-2016; however, upon closer investigation the Children's Bureau modified the way substantiation data was represented in their reports from 2011 forward; thus in the interest of time, I decided to limit the scope. There were significant time challenges in cleaning, aggregating and extracting all the data from different sources, including manually extracting tables from PDF's for the maltreatment reports, downloading and filtering all the BLS and Census data, and discovering cases of missing data after using RStudio to compile the final dataset.

IV. Model Results

$$\begin{aligned}\ln(SubstantiationRate)_{i,t} = \alpha_{i,t} + \beta_{1i,t}\ln(Caseload)_{i,t} + \beta_{2i,t}\ln(HourlyWage)_{i,t} \\ +\beta_{3i,t}\ln(ChildPoverty)_{i,t} + \gamma_i + \lambda_t + u_{i,t}\end{aligned} \tag{2}$$

Initially, the primary hypothesis predicted that workers who are paid less and have higher workloads will substantiate less cases to divert additional work off their plates into other channels of the child welfare system. The alternative hypothesis stated that neither of these main regressors should have any statistically significant effect on the substantiation

[4]High-quality, consistent and standard national and state level reporting of child maltreatment cases has been virtually nonexistent since the founding of the Children's Bureau in the early 1900's. As the field has only recently become more standardized regarding maltreatment data reporting, the lack of concrete national standards continually allows states to self-report under no consistent mandate. The Children's Bureau is one of the only sources of aggregated data of this nature, which is why I have used this source.

rate, given the arguments for worker attitude being the primary driver of substantiation decisions. As controls, the child poverty rate and the state and fixed effects were expected to have statistically significant and positive (for poverty) effects on the substantiation rate, as higher rates of child poverty are strongly associated with higher rates of maltreatment, thus more maltreatment reports in volume should correspond with more substantiated reports holding all else constant.

Table 2: MLR Regression Results with logs and Fixed Effects

	Dependent variable:	
	(Substantiation.Rate)	log(Substantiation.Rate)
	(1)	(2)
Caseload	0.011	
	(0.007)	
Hourly.Wage	0.241	
	(0.193)	
Poverty.Percent.Under.Age.18	−0.199	
	(0.122)	
log(Caseload)		−0.043
		(0.027)
log(Hourly.Wage)		-0.947^{***}
		(0.254)
log(Poverty.Percent.Under.Age.18)		-0.421^{**}
		(0.178)
Constant	23.320^{***}	7.510^{***}
	(4.868)	(0.822)
State fixed effects?	No	Yes
Time fixed effects?	No	Yes
Observations	332	332
R^2	0.025	0.863
Adjusted R^2	0.016	0.831
Residual Std. Error	11.422 (df = 328)	0.199 (df = 269)
F Statistic	2.846^{**} (df = 3; 328)	27.240^{***} (df = 62; 269)
Note:		*p<0.1; **p<0.05; ***p<0.01
		*Table omits results for state and fixed effects.

The regression results from the main specification above (regression 2) are mixed; the estimated coefficient on $\ln(Caseload)$ is not statistically significant, even at 10%, but the coefficient on $\ln(Wage)$ is highly statistically and economically significant. The insignificance of the caseload variable is surprising but reflects a possible strength of the child welfare workforce, where changes in workers' propensity to substantiate reports are immune to changes in their workloads. This effect is interesting because child social worker burnout and decline in service quality due to high workloads is a known relationship in the field (Hamama 2012). However, a potential explanation for the insignificance is that the worker responsible for substantiating cases is not necessarily the one responsible for executing the subsequent services needed (Jones 1993). This disconnect in work incentives would plausibly undermine the assumed motivation for heavily-burdened workers to substantiate less cases.

The wage coefficient estimated at -0.947 with a p-value less than 0.01 indicates that percent changes in the substantiation rate decrease almost a full percentage point for every percent increase in the hourly wage. This statistically significant result is expected according to the primary hypothesis, although the negative correlation and the 1:1 magnitude to changes in substantiation rate is surprising given the expected positive effect and the small variation in hourly wages for these workers. A potential explanation for both these findings is that higher paid workers have more motivation to conduct quality evaluations of each case. With more time and effort spent examining the merits of a report, a worker would be better able to distinguish truly substantiated cases from unsubstantiated ones, thus their rate of substantiation is lower. Evidence from M. C. Barth's (2003) study on wage factors in the child welfare field supports the alternative hypothesis that wages should not have such a strong effect on worker behavior, which is clearly refuted with this regression result. These contradicting findings are hard to untangle given a lack of evidence to be drawn on regarding wage effects on substantiation rates; this particular finding warrants more extensive study.

The coefficient on the main control variable, $\ln(ChildPoverty)$, is statistically significant at the 5% level and estimates a moderate 0.42% decrease in the substantiation rate for every percent increase in the child poverty rate. The negative and small effect of poverty on substantiation rate is surprising given known evidence that higher poverty rates are correlated with higher instances of child maltreatment, which should cause higher rates of substantiated cases through increased volumes of reports. In addition, given the results for the other two coefficients, a potential explanation for the small effect is that the majority of initial child maltreatment reports come from poverty-stricken areas. Thus, the relationship between whether these reports are substantiated and the state's overall child poverty rate would not be significant because the sample pool is mostly impoverished. This is supported by evidence that poor children are overrepresented in child welfare caseloads (Jonson-Reid, Drake and Kohl 2008) due to the increased risks for maltreatment they are exposed to. The negative relationship remains puzzling because of the overwhelming empirical evidence for positive effects of poverty on maltreatment rates; this may be caused by measurement error or other misspecifications but warrants further study.

Overall, the variables specified in this regression are jointly significant at the 1% level with a reported F-stat of 27.24. Thus, we can reject the null hypothesis that caseload and wages do not jointly influence the substantiation rate. However, we conclude that the estimated primary drivers of changes in the substantiation rate are changes in the hourly wages of child, family and school social workers, not changes in caseload. This conclusion partially supports both the primary and alternative hypotheses because the primary hypothesis expects significant coefficients on both main regressors while the second expects little to no significance.

Based on the reported adjusted R^2 of 0.83 in the final regression (3) from Table 2, we conclude that the current model specification already has strong explanatory power for variations in substantiation rate, or that approximately 83% of the variation in the substantiation rate is explained by the variation in the wage, caseload and control variables. Regarding this unclear conclusion, it is possible that the shortcomings of this regression model and data are causing these mixed results. Another potential interpretation is that the high goodness of fit from this regression does not support the secondary hypothesis about the importance of workers' attitudes.

Potential issues with endogeneity involve the OVB discussed in the model specification section, specifically surrounding worker characteristics like race, attitude and background affecting substantiation rates. In addition, missing data especially for caseload may be causing the insignificant regression coefficient, as there may not be enough consistent variation in this sample to predict an accurate relationship. A potential solution is to find an instrumental variable which provides consistent data than caseload, but still measures the effect of workload pressure on substantiation rates.

IVa. Policy Implications & Limitations of Study

The main policy implication from one of the main regressors, caseload, is that states should not worry about caseload as a major factor affecting substantiation rates. Instead, states should focus on paying their caseworkers proportionally more each year to obtain larger decreases in their expected maltreatment substantiation rates, ceteris paribus. If states desire to systemically decrease their rates of substantiation, they should consider increasing the wages of their child welfare workers proportionally to aid this effort. One major clarification should be emphasized in this policy implication; this study's findings apply to the wage and

Table 3: Avg. hourly wage, substantiation rate and child poverty rate for the 10 highest-paying states over 1999-2010 (excl. 2003)

State	Avg Hourly Wage ($)	Avg Substantiation Rate (%)	Avg Child Pov. Rate (%)
Connecticut	25.56	29.56	10.98
New Jersey	23.90	15.30	11.53
Rhode Island	22.78	31.24	16.64
Hawaii	21.85	50.14	13.01
Minnesota	21.75	30.46	11.08
New York	21.60	31.88	20.09
Illinois	21.37	26.54	16.48
Michigan	21.36	20.32	17.60
Nevada	21.35	22.70	15.39
California	21.20	25.34	18.95

Table 4: Avg. hourly wage, substantiation rate and child poverty rate for the 10 lowest-paying states over 1999-2010 (excl. 2003)

State	Avg Hourly Wage ($)	Avg Substantiation Rate (%)	Avg Child Pov. Rate (%)
Missouri	15.25	11.23	18.05
South Dakota	15.16	20.65	17.13
Montana	14.98	11.65	19.57
Arkansas	14.96	27.81	24.25
Nebraska	14.93	27.60	13.84
South Carolina	14.59	35.93	21.45
Oklahoma	14.53	20.30	21.64
Tennessee	14.46	18.31	21.08
Mississippi	14.41	24.33	28.65
West Virginia	12.66	24.72	23.94

substantiation rate differential that states have, not the aboslute level. This is clear in Tables 3 and 4, which show interesting evidence that higher paying states on average have some of the highest absolute magnitudes of average substantiation rates.

This is likely due to higher "starting points" of substantiation rates and hourly wages, not any causal effects or relationship between the magnitudes of these measures. It is important to distinguish average absolute magnitudes of untransformed data from percent changes used in the main regression. The reported negative wage differential effect on *changes* in the substantiation case is different from the conclusion that states with higher wages have lower substantiation rates; the result just indicates that proportional increases in wages will cause proportional decreases in the substantiantion rate and does not take into account actual levels in that state.

There is little to no existing literature on this wage and substantiation rate relationship, likely due to the multitude of confounding factors that drive workers' propensity to substantiate. There is also no clear direction on whether states should strive for lower substantiation rates solely on the basis of larger wage increases for their social workers. As mentioned before, the primary explanation for such an effect is that workers who are paid proportionally more have more resources and motivation to scrutinize cases better, thus that they do not substantiate cases as often. However, there is no literature exploring the validity of this explanation and requires more qualitative study to clarify. Overall, this clear but inconclusive policy implication validates the need to study what other unmeasured factors have a significant effect on reducing substantiation rates.

The limitations of this study include the obvious omission of a key variable, worker attitude, for which there is no reliable existing data source for. In addition, missing data in a key regressor, caseload, adds to the weakness of the study's results. Form misspecification is another potential weakness, where the true functional relationship between wages, caseload and poverty rates may not be correct in this model, leading to inaccurate conclusions. Overall, this study requires much more detailed work in refining the model specification and providing enough proxies for missing data to become extremely robust.

IVb. Robustness Checks

While addressing these major limitations are beyond the scope of this project, there are several robustness checks for the proposed model which demonstrate its strengths.

1. The application of the natural log to explanatory variables once state and time fixed effects are added

Table 5: Avg. hourly wage, substantiation rate and child poverty rate for the 5 highest substantiating states over 1999-2010 (excl. 2003)

State	Avg Hourly Wage ($)	Avg Substantiation Rate (%)	Avg Child Pov. Rate (%)
Massachusetts	20.82	54.47	12.78
Hawaii	21.85	50.14	13.01
Maine	17.23	42.98	15.50
Georgia	16.42	41.43	19.96
Utah	16.43	38.97	11.88

2. The removal of the top 5 states with the highest average substantiation rate during 1999-2010 from the sample

Table 6 indicates that the significance of the wage coefficient is robust to the natural log transformation applied to the fixed effect model. Neither caseload nor poverty rate are robust to this transformation, but this robustness of the wage coefficient remaining around -0.9 and statistically significant demonstrates validity in interpreting its implications.

Table 7 illustrates the top four states with the highest average substantiation rates- omitting these states and running the regression shows that the model is robust, even when states who most likely are the primary drivers behind the dependent variable are omitted. The coefficients remain statistically significant and similar to the original results.

Table 6: MLR Regression Results with Fixed Effects

	Dependent variable:	
	(Substantiation.Rate)	log(Substantiation.Rate)
	(1)	(2)
Caseload	−0.006 (0.004)	
Hourly.Wage	-0.981^{***} (0.266)	
Poverty.Percent.Under.Age.18	−0.348 (0.247)	
log(Caseload)		−0.043 (0.027)
log(Hourly.Wage)		-0.947^{***} (0.254)
log(Poverty.Percent.Under.Age.18)		-0.421^{**} (0.178)
Constant	56.565^{***} (6.647)	7.510^{***} (0.822)
State fixed effects?	Yes	Yes
Time fixed effects?	Yes	Yes
Robust SE?	No	No
Observations	332	332
R^2	0.892	0.863
Adjusted R^2	0.868	0.831
Residual Std. Error (df = 269)	4.190	0.199
F Statistic (df = 62; 269)	36.010^{***}	27.240^{***}
Note:		*p<0.1; **p<0.05; ***p<0.01
		*Table omits results for state and fixed effects.

Table 7: MLR Regression Results; omitting MA, HI, ME, GA, and UT with Robust SE

	Dependent variable:	
	log(Substantiation.Rate)	log(Substantiation.Rate)
	(1)	(2)
log(Caseload)	−0.043 (0.027)	−0.040 (0.029)
log(Hourly.Wage)	-0.947^{***} (0.254)	-0.920^{***} (0.270)
log(Poverty.Percent.Under.Age.18)	-0.421^{**} (0.178)	-0.591^{***} (0.205)
Constant	7.510^{***} (0.822)	7.958^{***} (0.925)
State fixed effects?	Yes	Yes
Time fixed effects?	Yes	Yes
Omitted States?	No	Yes
Observations	332	298
R^2	0.863	0.839
Adjusted R^2	0.831	0.800
Residual Std. Error	0.199 (df = 269)	0.204 (df = 239)
F Statistic	27.240^{***} (df = 62; 269)	21.503^{***} (df = 58; 239)
Note:		*p<0.1; **p<0.05; ***p<0.01
		*Table omits results for state and fixed effects.

V. Conclusion

This paper tackles the question of whether increased caseloads and decreased wages would predict systemically lower or higher substantiation rates. The main idea behind such an attempt was to chip away at the big unknown drivers behind child welfare workers' decision making process, especially regarding how they decide what constitutes substantiated child abuse and neglect. The paper used an MLR model to regress average hourly wages and annual caseload on the substantiation rate, controlling for state child poverty rates and including time and state fixed effects.

The key results indicate that increases in child, family and school social workers' wages have a negative proportional effect on the substantiation rate, while changes in caseload have no statistically significant effect on changes in substantiation rates. In addition, higher rates of child poverty have a more mild negative effect on the substantiation rate, a surprising finding given the mountain of existing evidence for strong positive effects of poverty on child maltreatment rates (and implicitly substantiation rates).

All of these key findings, which have unclear policy implications due to the lack of literature clarifying wage effects on child welfare worker decision making, point to the complexity of factors behind child welfare workers' decision making process and validates the need for further research to be done on these factors. Researchers should further study child welfare workers' wages and unique characteristics like worker attitudes or ideologies to grasp a better idea of how to conduct successful policy in reducing child maltreatment rates.

VI. Appendix


Graph 2: ln(substantiation rate) on ln(explanatory vars)
Caseload
Poverty.Percent.Under.Age.18
Hourly.Wage
ln(substantiation Rate)
4.0
3.5
3.0
2.5
2.0
2
3
4
5
6
7
ln(variable values)
variable
Caseload
Poverty.Percent.Under.Age.18
Hourly.Wage
Data from BLS, Census.gov, Children's Bureau

Table 8: MLR Regression Results with Various Fixed Effects

	Dependent variable:		
	(Substantiation.Rate)	(Substantiation.Rate)	(Substantiation.Rate)
	(1)	(2)	(3)
Caseload	0.011	-0.005	0.007
	(0.007)	(0.004)	(0.007)
Hourly.Wage	0.241	-0.960^{***}	0.916^{***}
	(0.193)	(0.149)	(0.238)
Poverty.Percent.Under.Age.18	-0.199	-0.316^{*}	0.101
	(0.122)	(0.174)	(0.135)
Constant	23.320^{***}	57.191^{***}	10.761^{*}
	(4.868)	(3.697)	(5.665)
State fixed effects?	No	Yes	No
Time fixed effects?	No	No	Yes
Observations	332	332	332
R^2	0.025	0.888	0.097
Adjusted R^2	0.016	0.868	0.060
Residual Std. Error	11.422 (df = 328)	4.191 (df = 279)	11.166 (df = 318)
F Statistic	2.846^{**} (df = 3; 328)	42.704^{***} (df = 52; 279)	2.627^{***} (df = 13; 318)

Note: *p<0.1; **p<0.05; ***p<0.01

*Table omits results for state and fixed effects.

Table 9: MLR Regression Results with Various Fixed Effects and log Transformations

	Dependent variable:		
	(Substantiation.Rate)	log(Substantiation.Rate)	log(Substantiation.Rate)
	(1)	(2)	(3)
Caseload	−0.006 (0.004)	−0.0003 (0.0002)	
Hourly.Wage	-0.981^{***} (0.266)		
log(Caseload)			−0.043 (0.027)
log(Hourly.Wage)		-0.973^{***} (0.253)	-0.947^{***} (0.254)
Poverty.Percent.Under.Age.18	−0.348 (0.247)	-0.027^{**} (0.012)	
log(Poverty.Percent.Under.Age.18)			-0.421^{**} (0.178)
Constant	56.565^{***} (6.647)	6.743^{***} (0.706)	7.510^{***} (0.822)
State fixed effects?	Yes	Yes	Yes
Time fixed effects?	Yes	Yes	Yes
Observations	332	332	332
R^2	0.892	0.862	0.863
Adjusted R^2	0.868	0.830	0.831
Residual Std. Error (df = 269)	4.190	0.200	0.199
F Statistic (df = 62; 269)	36.010^{***}	27.000^{***}	27.240^{***}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

*Table omits results for state and fixed effects.