FAMILY AND YOUTH CAREER DEVELOPMENT:
TOWARDS A BETTER UNDERSTANDING OF THE MECHANISMS

A Dissertation in
Human Development and Family Studies

by

Xiaoran Sun

© 2019 Xiaoran Sun

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2019
The dissertation of Xiaoran Sun was reviewed and approved* by the following:

Susan M. McHale
Distinguished Professor of Human Development and Family Studies
Professor of Demography
Dissertation Adviser
Chair of Committee

Nilam Ram
Professor of Human Development and Family Studies
Professor of Psychology

Gregory M. Fosco
Associate Professor of Human Development and Family Studies
Karl R. and Diane Wendle Fink Early Career Professor for the Study of Families

Molly A. Martin
Associate Professor of Sociology and Demography

Lisa M. Gatzke-Kopp
Professor-in-Charge of Human Development and Family Studies Program
ABSTRACT

Theory and empirical evidence have highlighted the importance of family context to youth career development across adolescence and young adulthood (Eccles, 2011; Lawson, 2018; Lent et al., 1994; Porfeli & Vondracek, 2009; Savickas, 2002; Super, 1980; Whiston & Keller, 2004). Built on a model that incorporates a variety of theoretical frameworks illuminating the role of family in youth career development, this dissertation aimed to advance understanding of mechanisms underlying the interplay between family systems and youth career development. In particular, three studies addressed mechanisms in the theoretical model that were highlighted in theory but understudied in prior research. Study 1 aimed to explain the long-term implications of mother- and father-adolescent relationship quality for career attainment in young adulthood through youth career development processes, in particular, career adaptivity. Using longitudinal data from 236 youth (53% female; age $M = 15.17$, $SD = .96$ at Time 1) and structural equation modeling, tests of a mediation model revealed the mediating role of adolescent career adaptivity (captured by academic performance, sense of control, and self-worth; one year after Time 1) in the link between mother-adolescent relationship quality (Time 1) and young adult occupational prestige at around age 26, though the effects of father-adolescent relationship quality were nonsignificant. Study 2 focused on whether and how youth career development processes, such as their early work experiences, have implications for their family relationships. Given that work is a pervasive yet understudied context for Latino youth (U.S. Bureau of Labor Statistics, 2018), this study took an ethnic homogeneous approach to test longitudinal implications of Mexican-origin youth’s work experiences (including work hours and workplace discrimination) for their relationship quality with fathers as reported by both youth and fathers. Using data with two time points across a two-year interval from 187 youth (52.4% female, 64.7% born in U.S., 50.8%
older siblings; $M = 19.33, SD = 1.78$ at Time 1) from 127 Mexican-origin families, results of
multivariate multilevel models revealed a curvilinear link between youth workplace
discrimination and fathers’ reports of relationship quality, and linear effects of youth work hours
and workplace discrimination on youth relationship reports qualified by youth gender and mother
employment. Study 3, based on the National Longitudinal Study of Adolescent Health (Add
Health; Harris & Udry, 1994-2008), took an innovative approach to analytically synthesize 101
prior studies by examining 53 family experience variables in adolescence as predictors of young
adults’ educational attainment, an important step in career development that conditions future
career opportunities and outcomes (IOM & NRC, 2015). In particular, using a machine learning
approach, this study answered three questions: (1) How accurately does this broad range of
adolescent family factors predict young adult educational attainment? (2) When examined
concurrently, which family experience factors are the best predictors of young adult educational
attainment? And (3) What complex patterns, including nonlinearities and interactions involving
this range of family factors, merit further examination? Overall, this dissertation provided
evidence for the theoretical model in illuminating the mechanisms linking family systems and
youth career development processes and their future career attainment, introduced innovative
methods to research on family and career development, provided implications for family-based
practice promoting youth achievement, and directed attention to future research efforts for a
better understanding of the mechanisms underlying family influences on youth career
development.
# TABLE OF CONTENTS

List of Tables........................................................................................................vi

List of Figures.........................................................................................................vii

Acknowledgements..............................................................................................viii

Chapter 1. GENERAL INTRODUCTION.................................................................1

   Overview of Three Dissertation Studies..........................................................5
   References.........................................................................................................10

Chapter 2. CAREER ADAPTIVITY MEDIATES LONGITUDINAL LINKS BETWEEN
PARENT-ADOLESCENT RELATIONSHIPS AND YOUNG ADULT OCCUPATIONAL
ATTAINMENT.................................................................16

   Introduction.....................................................................................................16
   Method............................................................................................................24
   Results............................................................................................................31
   Discussion......................................................................................................36
   References.....................................................................................................44

Chapter 3. IMPLICATIONS OF MEXICAN-ORIGIN YOUTH’S WORK EXPERIENCES FOR
RELATIONSHIPS WITH FATHERS.........................................................54

   Introduction.....................................................................................................54
   Method............................................................................................................65
   Results............................................................................................................70
   Discussion......................................................................................................75
   References.....................................................................................................84

Chapter 4. ADOLESCENT FAMILY EXPERIENCE PREDICT YOUNG ADULT
EDUCATIONAL ATTAINMENT: A DATA-BASED CROSS-STUDY SYNTHESIS WITH
MACHINE LEARNING.................................................................95

   Introduction.....................................................................................................95
   Method............................................................................................................106
   Results............................................................................................................114
   Discussion......................................................................................................122
   References.....................................................................................................131

Chapter 5. GENERAL DISCUSSION.................................................................178

   References.....................................................................................................186
## LIST OF TABLES

Table 2.1. Means ($M$), Standard Deviations ($SD$), and Correlations Among Study Variables Stacked Across Siblings ($N = 236$) .................................................................50

Table 3.1. An Illustration of Data Stacked Within Each Family for Analysis With Multivariate Multilevel Models ........................................................................................................91

Table 3.2. Means ($M$), Standard Deviations ($SD$), and Correlations Between Study Variables Stacked Across Siblings ($N = 187$) ...............................................................92

Table 3.3. Unstandardized Coefficients ($\gamma$) and Standard Errors ($SE$) From Two-Intercept Multivariate Multilevel Models Testing Curvilinear Associations Between Youth’s Work Experiences (Work Hours and Workplace Discrimination; Time 1) and Fathers’ and Youth’s Reports of Relationship Quality ($RQ_{Father}$ and $RQ_{Youth}$; Time 2), and Their Moderation by Youth Gender and Mothers’ Employment Status .........................93

Table 4.1. Overview of Publications ($N = 101$) Using Add Health Data That Examined Effects of Family Experience Variables at Wave I on Educational Attainment Outcomes at Waves III or IV ................................................................................................................146

Table 4.2. Family Experience Categories and Constructs for Corresponding Variables Used as Predictors in ML Models Predicting Young Adult Educational Attainment ..............168

Table 4.3. Results of Tuning and Prediction Performances of Models Predicting College Enrollment and Graduation Trained by Regularized Logistic Regression and Random Forests ........................................................................................................170
LIST OF FIGURES

Figure 1.1. An illustration of the role of family in youth career development and career attainment as proposed across career development theories ........................................2

Figure 2.1. The mediation model linking mother- and father-adolescent relationship quality to young adult occupational prestige through adolescent career adaptivity ....................51

Figure 2.2. The career adaptation process as proposed in the career construction theory ............52

Figure 2.3. Estimated model with standardized path coefficients and standard errors .............53

Figure 3.1. Curvilinear association between youth workplace discrimination and father-youth relationship quality reported by father; the turning point is indicated by the dotted line and the 95% CI, by dashed lines .................................................................94

Figure 4.1. An illustration of nested 5-fold cross-validation implemented in this study ............171

Figure 4.2. Feature importance rankings for predicting college enrollment with regularized logistic regression (left) and random forests (right) .................................................172

Figure 4.3. Recursive feature elimination results for predicting college enrollment with regularized logistic regression (upper) and random forests (lower) and the models selected in order to identify important features (i.e., predictors) ..............................173

Figure 4.4. Feature importance rankings for predicting college graduation with regularized logistic regression (left) and random forests (right) ..................................................174

Figure 4.5. Recursive feature elimination results for predicting college graduation with regularized logistic regression (upper) and random forests (lower) and the models selected in order to identify important features (i.e., predictors) ..................................175

Figure 4.6. Partial dependence plots (2D) of selected important family experience predictors in the random forests model predicting the probability of college enrollment (upper panel) and college graduation (lower panel) ..............................................................................176

Figure 4.7. Partial dependence plots (3D) of the random forests model predicting the probability of college enrollment (upper panel) from the interactions between shared dinner with parents and father education (left) and between intergenerational closure and father educational expectations (right), and model predicting the probability of college graduation (lower panel) from the interactions between father-adolescent shared activities and father occupational prestige (left) and between mother-adolescent relationship quality and father-adolescent shared activities (right) ..................................................177
ACKNOWLEDGEMENTS

First, I would like to thank the members of my committee who have been so supportive and helpful for my dissertation. Susan McHale served as my adviser/mentor continuously for all 4 years of my time in graduate school. I want to thank her for always being encouraging and patient with me, generous about her time spent in mentoring me, and willing to offer any help. I learned so much from her, including (but not limited to) in how to think critically and deeply about research, how to rigorously and innovatively conduct research, and how to write good academic papers. She has always been my role model who made me believe that women in academia can be distinguished and impactful and extremely kind at the same time. As an international student being in U.S. for graduate school, I have inevitably been through many difficulties and challenges, but I have never lost hope, because Susan was always there for me, making me feel that no matter how hard the situation was, I could always make a way out of it with her support. I am also thankful for her support for me exploring new research fields and learning new technical skills even though that could take my time away from working for her. With her support, I was able to minor in Social Data Analytics (SoDA) and take classes for big data and machine learning, which allowed me to study family and youth development in these innovative approaches.

I want to thank Nilam Ram, who has always been willing to treat me as an independent scholar and mentally challenge me in constructive ways, making me grow intellectually and be a better researcher, and I sincerely believe that I would not have landed my upcoming job at Stanford University without his kind help. I am also especially thankful for his insights and support for my dissertation study with machine learning. It was the initial communication with him and his encouragement that made this study possible to conduct and continue.

I have always enjoyed my interactions with Greg Fosco. His class on family systems
theory and research was one of the most interesting and helpful classes that I have taken in the graduate program. He has also been helpful with my professional development from early on, when he served as my candidacy committee member. For the dissertation, I want to thank Greg for his feedback on the proposal, all of which has been very constructive towards improving my work.

I also want to thank Molly Martin, who has been willing to serve on my committee since my comprehensive exam. She offered so many helpful insights from sociological perspectives, which I believe are beneficial to not only my dissertation, but also my macro-level thinking as a family researcher. I also gained knowledge about family and career development from sociology literature thanks to the readings that she has recommended to me.

I would also like to thank other HDFS faculty members who have served on my Master’s thesis and candidacy committee—Eva Lefkowitz, Peter Molenaar, Nan Crouter, and Mayra Bámaca-Colbert, and who have taught great classes that expanded my knowledge in human development and family studies and methodologies—Douglas Teti, Chuck Geier, Dave Eggebean, Zita Oravecz, Tim Brick, Damon Jones, Eric Loken, Bethany Bray, Sara Vasilenko as well as Greg and Peter as just mentioned. Outside HDFS department, I would like to thank Dana Mitra (EDTHP) for her class on academic writing, and Jia Li (STAT), Saeed Abdullah (IST), Ashton Verdery (SoDA) and Burt Monroe (SoDA) for their classes related to big data and machine learning. My SoDA co-mentor, Guangchi Chi, has been very kind in sharing his Twitter data with me and helping me to develop my big data research projects. Outside Penn State, I want to thank Kimberly Updegraff at Arizona State University, who has collaborated multiple papers with me and kindly offered advice to my professional development.

I have special thanks to Sherry Corneal. I got to interact with her in her teaching seminar
in Fall 2017 where I learned from her about the art of teaching and the spirit of life. I will always keep her in my memory and in my heart.

I want to thank my graduate student colleagues from within and outside the department, who are also my dear friends supporting me through graduate school in a country far away from home—Lorena Aceves, Ashley Tate, Kaitlin Fronberg, Yunying Le, Xutong Zhang, Grace Mak, Lu Ou, Mengya Xia, Linying Ji, Xiao Yang, Bingyuan Liu, Zhi Chai, and many more (and please forgive me if I fail to mention your name). It was them who helped me through all the cultural shocks and in adapting to the graduate school environment, and who have always been there sharing the ups and downs in graduate school and in life. They are my family in State College.

For this dissertation, I used data from the Family Relationships Project, funded by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01- HD32336) to Susan M. McHale and Ann C. Crouter, Co-Principal Investigators and the Juntos Project, funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01HD39666) to Kimberly A. Updegraff and Adriana J. Umaña-Taylor, Principal Investigators and Susan M. McHale and Ann C. Crouter, Co-Principal Investigators. I also used data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. I received the training for big data and machine learning research with support by the National Science Foundation under IGERT Grant DGE-1144860, Big Data Social Science. Also, I want to thank the Morningside Cultural China Program at Zhejiang University who have been providing me with scholarship in the past four
years that covered my expense on trip back to China every winter break.

Finally, I thank my parents, who always love me and believe in me. In the past four years, they told me so many times how proud they were with me getting a Ph.D. in a field that I am passionate about. Since I came to U.S., My parents and I have had long video-calls every weekend by which they have consistently provided me with psychological support. Knowing that they always have my back made me endure the good and bad times. They were always genuinely happy for me with all the milestones, including major ones such as passing candidacy and comprehensive exams and minor ones such as the first time I spoke up in class and my first guest lecture.
CHAPTER 1
GENERAL INTRODUCTION

Across childhood and adolescence career development is a central task that lays the foundation for adulthood career success and for well-being, more generally (Hartung, Porfeli, & Vondracek, 2005; Savickas, 2002; Super, 1980). Youth’s career development includes building competencies and feelings of self-efficacy, forming career orientations and aspirations, pursuing education and early work experiences, and ultimately making career choices. And, in the course of development, youth’s interactions with their social contexts allow for realizing person-environment fit, which is in itself a factor in career success (Eccles, 2011; Lent, Brown, & Hackett, 1994; Porfeli & Vondracek, 2009; Savickas, 2002; Super, 1980). Among the many social contexts that impinge on youth development, family plays a key role according to both theory (Eccles, 2011; Lent et al., 1994; Porfeli & Vondracek, 2009; Savickas, 2002; Super, 1980) and empirical evidence (Bryant, Zvonkovic, & Reynolds, 2006; Lawson, 2018; Whiston & Keller, 2004). Thus, understanding how family systems are interrelated with youth career development processes and predict future career success has important implications for individual well-being across the lifespan, parenting and family socialization practices, and public policy (Lawson, 2018; Whiston & Keller, 2004).

Theoretical frameworks on life-span career development, despite the variety in their proposed mechanisms underlying career development and attainment, have all highlighted the role of family. Figure 1.1 summarizes how family systems shape youth career development and attainment as proposed in a range of career development theories and frameworks, including the life-career theory (Super, 1980), the social cognitive career theory (Lent et al., 1994), the expectancy-value model of achievement (Eccles, 2011), the developmental-contextual model of career development (Porfeli & Vondracek, 2009), and the career construction theory (Savickas,
These theories all propose that family experiences in youth’s early years, including family resources and socialization processes across childhood and adolescence, lay the foundation for career attainment in adulthood (path $a$), potentially through shaping early career development processes (path $b$) that lead to later career attainment (path $c$). Although different theories propose different mechanisms and components for youth career development, they all recognize that youth career development is a systemic process with multiple components influencing one another (path $e$). Further, the bi-directionality in person-context interactions has been emphasized across theories, and encompasses the feedback loop from youth career development processes to family systems (path $d$). Finally, family systems also include multiple components that interact with one another (path $f$) in influencing youth career development and attainment.

Figure 1.1. An illustration of the role of family in youth career development and career attainment as proposed across career development theories.

Grounded in several theories, evidence for paths $a$, $b$, and $e$ has been documented in previous empirical studies. With respect to path $a$, longitudinal research shows that family income, parental occupational status and positive parenting in childhood and/or adolescence are associated with career success in (young) adulthood, including employment, income, occupational prestige, and career satisfaction (Ashby & Schoon, 2010; Caspi, Wright, Moffitt, & Silva, 1998; Dubow, Boxer, & Huesmann, 2009; Gordon & Cui, 2015; Lawson, Crouter, &
McHale, 2015; Sun, McHale, & Updegraff, 2017; Wiesner, Vondracek, Capaldi, & Porfeli, 2003). Cross-sectional and short-term longitudinal studies have tested path b, revealing, for example, effects of parenting and parents’ occupational status on youth’s career self-efficacy (Lim & Loo, 2003), attitudes, values, preferences, motivations and aspirations (Porfeli, Wang, & Hartung, 2008; Weinshenker, 2006), exploration and adaptation and vocational interests and identity (Bryant et al., 2006; Whiston & Keller, 2004). Research also has examined interconnections among different career development components in adolescence (Diemer & Blustein, 2007; Hirschi, Herrmann, & Keller, 2015; Negru-Subtirica, Pop, Crocetti, 2015), captured by path e.

Research on paths c, d, and f, however, is more limited. With respect to path c, research has been limited on the continuity of career development spanning across different developmental periods, such as adolescence and young adulthood (Bryant et al., 2006; Whiston & Keller, 2004). Moreover, little is known about whether and how the combination of paths b and c may help to explain path a, despite existing theoretical propositions. As exceptions, prior studies involving tests of both paths b and c have focused on youth’s career orientations (e.g., ambitions, values, and aspirations) that are shaped by family experiences and linked to adulthood career attainment (Ashby & Schoon, 2010; Dubow et al., 2009), but other dimensions of youth career development have rarely been considered as potential mechanisms linking youth’s family experiences and their future career attainment. Almost nothing is known about path d, except for early research on the effects of adolescent employment status, work intensity and earnings on their family relationship quality (Shanahan, Elder, Burchinal, & Conger, 1996a; 1996b). Further, even though theoretical accounts highlight that family is a multi-component and dynamic system, few studies have examined multiple family processes and their interactions and most rely
on data from a single family member. As such, empirical evidence pertaining to path \( f \) is limited.

Accordingly, aiming for a more complete understanding of how family experiences shape youth career development and attainment, my dissertation research examined paths \( c \), \( d \), and \( f \) (Figure 1.1), paths that have been proposed and highlighted in theories but about which there is little or no empirical evidence.

**Study 1** targeted path \( c \), in combination with path \( b \), by examining adolescents’ career adaptivity, a fundamental construct for youth career development and attainment according to the career construction theory (Savickas, 2002; 2013). Specifically, I tested whether and how mother- and father-adolescent relationship qualities were linked to adolescents’ career adaptivity one year later, and in turn, whether parent-youth relationships and adaptivity predicted career attainment at around age 26. **Study 2** targeted path \( d \) by testing associations, over a two-year interval, between youth’s early work experiences and their relationship quality with fathers in Mexican-origin families. **Study 3** targeted path \( f \), in combination with path \( b \), by analytically synthesizing 53 family experience variables in adolescence as predictors of young adults’ educational attainment, an important step in career development that conditions future career opportunities and outcomes (Dubow et al., 2009; IOM & NRC, 2015). In particular, I applied a machine learning approach to model and reveal complex dynamics (including nonlinearities and interactions) among these family predictors (i.e., path \( f \)).

Importantly, all three studies were grounded in models of human development (Bronfenbrenner, 1979; 1986; Studies 2 and 3), family systems (Minuchin, 1985; Studies 2 and 3) and career development theories (Savickas, 2002; Study 1), complementing prior empirical research that has often been atheoretical (Whiston & Keller, 2004). Moreover, given that family socialization and career development are both gendered (Lawson et al., 2015; Maccoby, 1998),
the role of youth gender was highlighted in all three studies as a potential moderator in the interplay between family and youth career development (Studies 1 and 2) and as a factor that interacts with family system processes and resources in youth career development and attainment (Study 3).

In addition to these substantive contributions, this dissertation research incorporated a number of methodological strengths. First, all three studies used longitudinal data, and both Studies 1 and 3 used data from different developmental periods, an advance over prior studies of family and youth career development that have relied primarily on cross-sectional or short-term longitudinal data (Bryant et al., 2006; Whiston & Keller, 2004). Data analysis methods applied for these longitudinal data include structural equation modeling for the mediation model (Study 1), multivariate multilevel modeling (Study 2), and machine learning algorithms, including regularized logistic regression and random forests, for building predictive models—an innovative step in family and career development research (Study 3). In addition, samples across studies come from difference sources, including a predominantly White sample from a northeastern state (Study 1), a Mexican-origin sample from a southwestern state (Study 2), and a nationally representative sample (Study 3).

**Overview of Three Dissertation Studies**

**Study 1: Career Adaptivity Mediates Longitudinal Links Between Parent-Adolescent Relationships and Young Adult Occupational Attainment.**

Given the limited understanding of the career development processes linking adolescents’ family experiences and their future career attainment, this study assessed a mediation mechanism whereby mother- and father-adolescent relationship qualities are linked to adolescents’ career adaptivity (path b in Figure 1.1), which in turn, predicts occupational attainment in young
adulthood (path $c$ in Figure 1.1). The focus on career adaptivity was grounded in the career construction theory, which proposes that career adaptation, a multi-dimensional process including adaptive readiness, adaptability resources, adapting responses and adaptation results, underlies youth career development (Savickas, 2013). Among the four dimensions, career adaptivity has been highlighted in the theory and in empirical evidence as fundamental for other dimensions (Hirschi et al., 2015; Rudolph, Lavigne, & Zacher, 2017; Savickas, Porfeli, Hilton, & Savickas, 2018; Šverko & Babarović, 2018; Zhou, Guan, Xin, Mak, & Deng, 2016). Based on the extant literature, I captured career adaptivity as a latent construct, manifested by adolescents’ academic achievement, sense of control and self-worth. To examine its role as a mediating process, I tested the longitudinal links between parents’ and adolescents’ reports of their relationship quality, career adaptivity one year later, and occupational attainment in young adulthood. Within the mediation model, I also tested the moderating role of youth gender as well as potential differences between effects of mother- and father-adolescent relationship quality on both adolescent career adaptivity and young adult occupational attainment.

**Study 2: Implications of Mexican-Origin Youth’s Work Experiences for Relationships With Fathers.**

Youth’s early work experiences are a central component of their career development (Savickas, 2002; Zimmer-Gembeck & Mortimer, 2006). Importantly, those work experiences do not take place in isolation, but are tied to youth’s experiences in other contexts, such as their families. This study aimed to examine the work-family interface, path $d$ in Figure 1.1, by examining effects of youth work experiences on their family relationships, specifically father-youth relationship quality. Although youth’s work-family interface, especially their work’s effects on family life, has been rarely examined, this test was supported by the ecological model,
which highlights the mesosystem (i.e., interplay between different proximal contexts, such as work and family) in individual development (Bronfenbrenner, 1986), family systems theory, which highlights that families are open systems and that family members’ experiences in and outside the family are interdependent (Minuchin, 1985), and the abundant literature on work-family interface among adults (Perry-Jenkins & Wadsworth, 2017). In this study, I focused on work experiences of Mexican-origin youth given the high prevalence of employment, including high-intensity work, among Latino youth (U.S. Bureau of Labor Statistics, 2018; Hwang & Domina, 2017), and their relationship quality with fathers given fathers’ predominant breadwinning role emphasized in this sociocultural group (Pinto & Coltrane, 2009; Wildsmith, Ramos-Olazagasti, & Alvira-Hammond, 2018). In particular, I tested longitudinal effects of youth work hours and workplace discrimination in late adolescence and young adulthood on father-youth relationship quality two years later. The co-existence of both positive and negative implications of work for family life, previous findings about curvilinear associations between youth work intensity and their academic performance, adjustment and father-youth communication (Hwang & Domina, 2017; Mortimer, Finch, Ryu, Shanahan, & Call, 1996; Shanahan et al., 1996b), led me to test a curvilinear pattern in these longitudinal effects, toward identifying the optimal level, or the turning points, at which the directions of the associations changed. Further, informed by the person-context interaction tenet of the ecological model (Bronfenbrenner, 1986) and the family systems tenet of family member interdependencies (Minuchin, 1985), I also tested potential moderation of the linkages between youth work experiences and father-youth relationship quality, specifically as a function of youth gender and maternal employment.

**Study 3: Adolescent Family Experiences Predict Young Adult Educational Attainment: A**
Data-Based, Cross-Study Synthesis With Machine Learning.

Frameworks such as family systems theory (Minuchin, 1985) and the developmental-contextual model of career development (Porfeli & Vondracek, 2009) have highlighted family as a complex system wherein processes and resources are multifaceted and interdependent. A key step in illuminating the role of family in youth career development is to unpack the family system by examining how multiple family processes and resources function together (path $f$ in Figure 1.1), as these relate to youth career development processes (path $b$). Accordingly, this study examined a wide range of adolescent family experiences as predictors of young adult educational attainment outcomes, specifically college enrollment and graduation given their important roles in later job opportunities and choices (IOM & NRC, 2015). I took the novel step of analytically synthesizing results from prior studies, using a machine learning (ML) approach, to address three questions (1) How accurately does this broad range of adolescent family factors predict young adult educational attainment? (2) When examined concurrently, which family experience factors are the best predictors of young adult educational attainment? And (3) What complex patterns, including nonlinearities and interactions involving this range of family factors, merit further examination? Based on a review of prior studies ($N = 101$ publications) that used data from the National Longitudinal Study of Adolescent Health (Add Health; Harris & Udry, 1994-2008) to investigate links between adolescent family experiences and young adult educational attainment, I identified 53 family experience variables to include in building comprehensive models predicting college enrollment and graduation. To do so I applied an ML-based approach due to its capacity to maximize predictive power, its functions of feature importance estimation and feature selection, which can identify the strongest predictors of an outcome, and using some algorithms (e.g., random forests), its capacity to freely estimate
nonlinear and interactive effects among predictors (Breiman, 2001; Brick, Koffer, Gerstorf, & Ram, 2017; Yarkoni & Westfall, 2017). Results in response to the three major study questions would advance understanding about the role of the multi-dimensional, multi-component family system in educational attainment, and illuminate the complexities among family processes and resources that worth further examinations.

In sum, my dissertation papers collectively extended prior research to advance understanding of the links between family system processes and resources in adolescence and youth career development and attainment. Each study addressed one or more of the paths depicted in the model shown in Figure 1.1, which was grounded in theories of career development and developmental and family frameworks. By testing paths b and c in Study 1, path d in Study 2, and paths f and b in Study 3, my dissertation examined processes highlighted in theoretical work but not yet the focus of empirical studies. Together these dissertation papers were designed to reveal the complex and systemic nature of family and youth career development.
References


CHAPTER 2
CAREER ADAPTIVITY MEDIATES LONGITUDINAL LINKS BETWEEN PARENT-ADOLESCENT RELATIONSHIPS AND YOUNG ADULT OCCUPATIONAL ATTAINMENT

Introduction

Occupational attainment in young adulthood is critical to well-being throughout adulthood (IOM & NRC, 2015), and adolescence is an important period during which youth develop competence and readiness for future occupational choices and achievements (Clausen, 1991; Fouad, 2007; Savickas, 2002). In adolescence, family experiences play an important role in shaping youth’s career development, including establishing career interests, values, and aspirations and in career exploration and adaptation processes (Bryant, Zvonkovic, & Reynolds, 2006; Lawson, 2018; Whiston & Keller, 2004). For example, although longitudinal data are limited, adolescents’ relationships with parents are one domain of family experience that has been linked to occupational attainment in young adulthood (Gordon & Cui, 2015; X. Sun, McHale, & Updegraff, 2017). In addition to limited longitudinal research on the links between adolescents’ family experiences and their later occupational attainment, we also know little about the career development processes in adolescence that may help to explain how family experiences affect young adulthood attainment (Gordon & Cui, 2015; Whiston & Keller, 2004).

Accordingly, the current study was designed to examine career development processes in adolescence and test whether they serve as a mechanism through which adolescents’ family experiences, specifically, relationships with mothers and fathers, are linked to their occupational attainment in young adulthood.

According to career construction theory, career development in adolescence is marked by career adaptation (Savickas, 2002; 2013), a process whereby youth accommodate to current and
future career-related tasks and transitions in order to achieve career success. The career construction theory highlights that family experiences shape adolescents’ career development processes, including career adaptation, and that adolescent career adaptation is fundamental to and predictive of future occupational attainment. Together, these tenets suggest a process whereby family experiences lead to occupational attainment through their effects on career adaptation. However, I could find no previous studies testing such processes using longitudinal data spanning adolescence and young adulthood. In order to better understand not only whether, but how family experiences in adolescence may influence career attainment, an important step is to illuminate the role of adolescent career adaptation processes in the longitudinal effects of family experiences for attainment in young adulthood. Thus, in this study, I tested the hypothesis that the effects of mother- and father-adolescent relationships on career attainment in young adulthood were mediated by career adaptivity (Figure 2.1)—a fundamental component in adolescent career adaptation development. Consistent with career construction theory, adaptivity was operationalized in terms of readiness, or the personal characteristics that prepare youth to accommodate to career development tasks (Hirschi, Herrmann, & Keller, 2015; Savickas, Porfeli, Hilton, & Savickas, 2018).

This study also took the important step of examining the role of gender in these processes. Although women in the U.S. have outperformed men since 1982 in educational attainment, including college graduation (Buchmann, DiPrete, & McDaniel, 2008), occupational gender segregation and an often corresponding gender wage gap favoring men have been persistent (Hegewisch, Liepman, Hayes, & Hartmann, 2010). In the face of gender differences in career development and occupational attainment (Eccles, 2011; Maccoby, 1998) and theory on the relative importance of mothers and fathers in young women’s and men’s career development
(Lim & Loo, 2003; Mortimer & Kumka, 1982; Schultheiss, Kress, Manzi, & Glasscock, 2001; Steele & Barling, 1996; Weinshenker, 2006), we still know very little about mothers’ versus fathers’ potential influences on their daughters’ and sons’ career development (Gordon & Cui, 2015; X. Sun et al., 2017). Thus, the second aim of this study was to examine differences in the effects of mother- and father-adolescent relationships and test youth gender as a potential moderator of the links between parent-adolescent relationships, adolescent career adaptivity and young adult occupational attainment.

**Adolescents’ Relationship with Parents and Occupational Attainment in Young Adulthood**

From a life course perspective, family experiences are important to youth’s concurrent development, and they continue to exert their influence over the long term (Elder, 1998). Thus, this perspective has served as a foundation for longitudinal studies on youth career development that examine whether and how adolescent experiences, including those with family members, predict occupational attainment in (young) adulthood (Ashby & Schoon, 2010; Clausen, 1991; Dubow, Boxer, & Huesmann, 2009; Gordon & Cui, 2015). More specifically, social capital theory highlights positive relationship with parents as one component of family’s social capital that fosters youth’s education and occupation achievement (Coleman, 1988). And, from among the multiple aspects of adolescents’ family experiences, the quality of relationships with parents has been a focus of research on youth career development and attainment (Caspi, Wright, Moffitt, & Silva, 1998; Dubow et al., 2009; Gordon & Cui, 2015; Lim & Loo, 2003; X. Sun et al., 2017; Whiston & Keller, 2004).

In addition to life course and social capital perspectives, the expectancy-value achievement model proposes to explain the role of parents in youth achievement (Eccles, 2011). This model targets parents’ and youth’s achievement orientations, expectations, and subjective task values as
key factors in youth’s educational and occupational attainments. Consistent with the expectancy-achievement model, one study found that both career aspirations and attainment values mediated the links between adolescents’ family experiences, including family socioeconomic status and parental expectations, and their occupational attainment in young adulthood (Ashby & Schoon, 2010). In other words, family experiences had their effects on attainment at least in part, through career development processes. Building on this work, in the current study I examined the mediating role of adolescent career adaptivity in the longitudinal association between parent-adolescent relationship quality and young adult occupational attainment.

**Adolescent Career Adaptivity and Occupational Attainment**

As noted, according to career construction theory, a central process in career development involves career adaptation, whereby youth accommodate to developmental and ecological conditions and transitions (i.e., changes in self and context) toward achieving an optimal person-environment fit (Savickas, 2002). As illustrated in Figure 2.2, career adaptation has four components: adaptive readiness, adaptability resources, adapting responses, and adaptation results (Savickas, 2013). In this study, I focused on adaptive readiness, in most literature and hereafter referred to as career adaptivity, given that it is fundamental to and conditions the three other components of career adaptation (Hirschi et al., 2015; Savickas, 2013).

Adaptivity was initially defined as “personal characteristics of flexibility or willingness to meet career tasks, transitions, and traumas with fitting responses” (Savickas, 2013, p. 157). Unlike the later career adaptation dimensions that have been defined and assessed by career tasks and transitions, adaptivity captures generic personal characteristics that prepare individuals to accommodate to career development tasks. In more recent research, adaptivity has been operationalized as a multi-component construct including cognitive ability and academic
performance, self-evaluations (e.g., self-worth, esteem, and confidence), sense of control, and open, extraverted, conscientious, and proactive personality traits (Hirschi, et al., 2015; Rudolph, Lavigne, & Zacher, 2017; Savickas et al., 2018; Šverko & Babarović, 2018; Zhou, Guan, Xin, Mak, & Deng, 2016). Prior studies have examined these individual components to determine the effects of adaptivity on other career adaptation processes and career adjustments. For example, Šverko and Babarović (2018) found that high school-aged youth’s academic performance, extraversion, conscientiousness, openness to experience, and self-evaluations were each linked across one semester to career adaptability resources. Other studies documented links from self-evaluation, proactive personality and sense of control to concurrent career adaptability resources and subsequent career adjustment outcomes over the short term, including career decision-making self-efficacy one month later (Zhou et al., 2016) and career planning six months later (Hirschi et al., 2015). Further supporting its significance, the operationalization of career adaptivity overlaps with that of adolescent “planful competence” within a life course perspective, a construct that incorporates self-confidence, dependability, and intellectual involvement and that is predictive of occupational status in middle to late adulthood (Clausen, 1991). Accordingly, grounded in the career construction theory and building on such findings, in this study I examined adaptivity as a latent construct with multiple components, specifically, academic performance, sense of control, and general self-worth. I also extended the longitudinal scope of research on career adaptation processes, which has primarily involved cross-sectional and short-term longitudinal designs, to examine the role of adaptivity in adolescence for career attainment about ten years later, in young adulthood.

As noted, I also extended research on the career construction theory to incorporate the role of family experiences, specifically relationships with parents. Although this theory highlights the
role of family experiences in career adaptation processes (Savickas, 2002), prior work has rarely included measures of family dynamics. A range of studies documents that parent-youth relationship qualities predict youth characteristics that are central to career adaptivity, including academic performance, sense of control, and self-worth (Garber, Robinson, & Valentiner, 1997; Steinberg, 2001; Y. Sun, 2001). Further, adults retrospectively attributed their ‘planful competence’ to good parenting practices (Clausen, 1991). Thus, I built on prior studies to test whether mother- and father-youth relationships predicted youth career adaptivity characteristics in adolescence, and whether these, in turn, predicted career attainment in young adulthood. Specifically, as illustrated in Figure 2.1, career adaptivity was a latent construct manifested by academic performance, sense of control, and self-worth, and was expected to explain potential links between mother- and father-adolescent relationship quality and young adult occupational prestige.

The Role of Gender in Parent-Adolescent Relationship-Adaptivity-Occupational Attainment Linkages

I also built on prior literature to test whether and how mother- and father-adolescent relationships differ in their effects on adolescents’ career adaptivity and, in turn, occupational attainment, as well as the role of youth gender in these processes. Previous research on the role of family experiences in career development has rarely included measures of both mother- and father-adolescent relationships to document their potential unique and relative implications. The limited body of literature on maternal versus paternal influences relies primarily on cross-sectional data, and findings have been mixed. One set of studies highlights mothers’ stronger influences on youth career development relative to fathers’ based on findings that mothers spend more time with youth and are more involved with their career development in childhood and
adolescence. Retrospective, qualitative data have indicated that young adults perceived their mothers as the most influential family member in their career exploration, as compared to fathers and siblings (Schultheiss et al., 2001). And, a cross-sectional study examining the roles of mother and father in the same model found that mother-adolescent relationship experiences were more strongly linked than those of fathers to young women’s and men’s self-efficacy and work attitudes (Lim & Loo, 2003).

A second group of studies, however, targeted father-adolescent relationships as the more powerful influence on youth career development and attainment (Hoffman, Hofacker, & Goldsmith, 1992; Zhao, Lim, & Teo, 2012) given that mothers’ warmth and involvement are more scripted and thus less salient to youth (Daly, 1996), and that fathers’ family role prioritizes connecting youth to the world outside family, including school and work (Maccoby, 1998). Consistent with these arguments, a retrospective study among college students in business majors showed that both young men and women more often reported father than mother as the primary influence on their career choices (Hoffman et al., 1992). Further, a cross-sectional study, which included information only on fathers, showed that fathers’ career-related parenting practices were concurrently associated with both young women’s and men’s career self-efficacy (Zhao et al., 2012). To address this mixture of findings, in the present study, I included measures of relationships with both mothers and fathers in a single model (Figure 2.1) and compared the effects of mother- and father-adolescent relationship quality on adolescents’ career adaptivity and subsequently, occupational attainment.

A surprisingly small set of studies has considered the role of youth gender in the links between experiences with mothers and fathers and career development. According to the social learning tenet that youth primarily model their same-gender parents (Maccoby, 1998), however,
relationship quality with mother should be more salient for girls’ whereas relationship quality with father should be more salient for boys’ career development. Consistent with this perspective, some previous studies have either focused on mothers’ influences on daughters’ career development (Steele & Barling, 1996; Weinshenker, 2006), or fathers’ influences on sons, such as in early studies of father-to-son career inheritance (Mortimer & Kumka, 1982). However, I was unable to find any prior study that compared the effects of both mother- and father-relationships on career development in adolescent girls and boys or young adult women and men. Therefore, in an effort to advance understanding of the role of gender in family socialization of career development, I tested the moderating role of youth gender in mother and father-adolescent relationship — career adaptivity — young adult occupational attainment linkages, specifically, the social learning hypothesis that linkages within same-gender dyads would be stronger than those of mixed-gender dyad linkages.

The Present Study

In sum, this study examined: (1) the effects of mother- and father-adolescent relationship quality on adolescents’ career adaptivity and young adult occupational attainment and the role of adaptivity as a mediator of these longitudinal linkages, and (2) the roles of parent and youth gender in these processes. The study’s hypothesized model, shown in Figure 2.1, was grounded primarily in the career construction theory (Savickas, 2002; 2013) and informed by the life course perspective (Elder, 1998), the social capital theory (Coleman, 1988) and the expectancy-value achievement model (Eccles, 2011) as well as previous findings about the effects of parenting on adolescent career development (Whiston & Keller, 2004) and research on the role of career adaptivity characteristics in career development (Hirschi et al., 2015; Šverko & Babarović, 2010; Zhou et al., 2016). To capture parent-adolescent relationship quality, I included
both parents’ and adolescents’ reports of relationship positivity in middle adolescence in order to reflect their shared relationship experiences, and separate reports on mother-youth and father-youth relationships to capture their potentially unique roles in youth career development (Cook, 2001). As indices of the latent construct, career adaptivity, I included academic achievement (report card grades) and adolescents’ reports of their sense of control and self-worth assessed one year following collection of the parent-youth relationship reports. Finally, following prior research (Ashby & Schoon, 2010; Dubow et al., 2009; X. Sun et al., 2017), young adult occupational attainment was indexed by occupational prestige, measured at about age 26. For covariates, I included mothers’ and fathers’ education and adolescents’ and young adults’ age to account for potential confounding effects of parents’ achievement (Caspi et al., 1998; Dubow et al., 2009; Gordon & Cui, 2015; X. Sun et al., 2017) and potential developmental differences in career adaptivity and attainment (Savickas, 2002), respectively.

Expanding upon prior research, I also tested the effects of mother- and father-adolescent relationships in the same model as a means of addressing alternative arguments and inconsistent findings about the relative influences of mothers and fathers on youth career development. Further, based on theories and research that have highlighted the salient role of same-gender parents in youth career development, I tested the moderation effect of youth gender to illuminate potential differences in the effects of mother- and father-adolescent relationships on their daughters’ and sons’ career adaptivity in adolescence and in turn, occupational attainment in young adulthood.

Method

Participants

I used data from the Phases 6, 7, and 11 and 12 (referred to Times 1, 2, 3, and 4 hereafter)
of a longitudinal project on youth development and family relationships, when the measures of interest were collected. This project started by sending recruitment letters to families in 16 school districts of a northeastern state; interested and eligible families returned postcards to the project office. Eligibility criteria were: (1) mothers and fathers were always-married and employed and (2) the family included a firstborn in the fourth or fifth grade and a secondborn, 1-4 years younger than the firstborn. We did not have data on the number of families that fulfilled the criteria but failed to respond, but over 90% of eligible families that returned postcards agreed to participate (total \( N = 203 \) families at Phase 1). There was a one year interval between Times 1 and 2 when youth were in middle to late adolescence, a two-year interval between Times 3 and 4 when youth were young adults; and a 9-year interval between Time 2 and 3.

The present study included families \(( N = 147 \) who completed interviews conducted at Time 1 and/or 2 as well as surveys at Times 3 and/or 4. \( T \)-tests and chi-squared analyses revealed that families included in this study did not differ from those not included on baseline demographic characteristics (mother and father education, family income, family size, youth gender and age) except for parent age. At baseline, compared to participating families, parents in nonparticipating families were significantly younger, \( M = 35.32 \ (SD = 3.63) \) v. \( M = 37.16 \ (SD = 3.95) \), \( t(201) = -2.58, p = .01 \) and \( M = 37.43 \ (SD = 3.79) \) v. \( M = 39.43 \ (SD = 5.30) \), \( t(201) = -2.99, p = .003 \), for mothers and fathers, respectively, and thus parent ages were controlled in following analyses. A total of 236 youth (53% female) were included in the analyses, 122 firstborns (52% female) and 114 secondborns (54% female). Consistent with previous studies on young adult achievement using the same sample (author citation), this study focused on firstborns’ occupational attainment at Time 3 and secondborns’ at Time 4, when both siblings were around 26 years old and the majority had completed their educations and were in relatively
stable jobs. Thus, among the participating youth from the 147 families, 254 young adults who reported a job title at Time 3 (for firstborns) or Time 4 (for secondborns) were first considered, but I excluded 18 who were at school and reported a part-time job unrelated to their major area of study (e.g., a psychology undergraduate student reporting a job as a bartender), consistent with previous research with this sample (author citation). Firstborns’ average age was 16.43 ($SD = 0.79$) at Time 1 and 26.28 ($SD = 0.81$) at Time 3. Secondborns’ average age was 13.83 ($SD = 1.12$) at Time 1 and 26.09 ($SD = 1.10$) at Time 4.

The sample included almost exclusively European American families living in small cities, towns, and rural communities, reflecting the racial/ethnic background of families of the region of the state where the study was conducted (85% European American; U.S. Census Bureau, 2000). Moreover, reflecting the educational (> 80% of adults completed high school) of the targeted population (U.S. Census Bureau, 2000), at Time 1 (2000), the average education level was 14.88 (some post high school training or college; $SD = 2.05$, range = 12 - 19) for mothers and 14.94 years ($SD = 2.35$, range = 11 - 20) for fathers. Family income ($74,600; SD = $39,925, range = $26,000 - $260,000) was higher than the state’s (median income = $55,714, for married-couple families), likely because this was a sample of dual-earner, midlife parents who were established in the labor force. Despite the variability in parental education and family income, most families were working to middle-class.

**Procedure**

Data were collected from mothers, fathers, and the two siblings via two methods. At Times 1 and 2, trained interviewers conducted home interviews to obtain parents’ and adolescents’ reports of their relationship experiences and personal characteristics. Informed consent/assent procedures were implemented at the beginning of home interviews, and family members were
then interviewed separately. Families were given a $200 honorarium for each home interview. At Times 3 and 4, phone interviews were conducted with young adults who reported on their current experiences, including their occupations. Consent was audio-recorded, and young adults received $100. All procedures of the larger study, were approved by the University’s Institution Review Board ([masked protocol number]).

**Measures**

**Parent-adolescent relationship quality.** At Time 1, both parents and adolescents reported on their relationship quality in home interviews. To assess *parental reports*, mothers and fathers completed the Acceptance Scale of the Parents’ Version of the Child’s Report of Parental Behavior Inventory (CRPBI; Schwarz, Barton-Henry, & Pruzinsky, 1985) separately for their experiences with firstborns and secondborns. This 5-point (*1 = not at all, 5 = very much*) scale includes 24 items, such as “I am a person who gives my child a lot of care and attention.” Item scores were averaged to create a mean score, with higher scores reflecting higher relationship quality, and Cronbach’s alpha range from .92 to .93 across mothers’ and fathers’ reports with two siblings. *Adolescent reports* of relationships with their mothers and fathers were based on an 8-item, 5-point scale (*1 = not at all, 5 = very much*; Blyth & Foster-Clark, 1987) that included items such as, “How much do you go to your mother/father for advice/support?” Item scores were averaged to create a mean score with higher scores reflecting higher relationship quality, and Cronbach’s alphas range from .81 to .86 across siblings’ reports with mothers and fathers.

**Career adaptivity.** I used adolescent career adaptivity indices measured at Time 2, specifically, academic performance, sense of control, and self-worth. Adolescents’ *academic performance* was assessed via grade point average (GPA), using grades from adolescents’ report cards provided by parents during the home interview. GPA was calculated from letter grades in
math, science, social studies, and language arts that were recoded into numerical scores ($A = 4, B = 3, C = 2, D = 1, E = 0$), such that high scores indicating higher academic performance. Sense of control was measured using the short version of the Nowicki-Strickland Internal-External Control Scale (Nowicki & Duke, 1983) where adolescents responded “yes” (coded as 1) or “no” (coded as 2) to 21 items such as, “Do you feel that most of the time it doesn’t pay to try hard because things never turn out right anyway?” Item scores were summed into a total score, with higher scores reflecting higher levels of perceived control, and Cronbach’s alphas are .72 and .74 for firstborns and secondborns, respectively. Self-worth was measured using the 5-item general self-worth subscale of the Perceived Competence Scale (Harter, 1982). Items were in a structured alternative format, such as, “Some teenagers are often disappointed with themselves BUT Other teenagers are pretty pleased with themselves.” Adolescents first indicated whether the statement on the left or the right side they were more like and then decided if that statement was “sort of true” or “really true” about them. The response was coded into a 4-point scale where $1 = \text{really true}$ and $2 = \text{sort of true}$ as described on the left side, and $3 = \text{sort of true}$ and $4 = \text{really true}$ as described on the right side. Item scores were averaged into a mean score, with higher scores reflecting higher self-worth, and Cronbach’s alphas are .89 and .85 for firstborns and secondborns, respectively. As noted, in this study, career adaptivity is a latent construct represented by the three observed indices.

**Occupational prestige.** At Times 3 and 4, young adults were asked about their employment status and the title of their current job in phone interviews. The jobs that firstborns reported at Time 3 and secondborns reported at Time 4 were coded using the Occupational Prestige Ratings from the 1989 General Social Survey (ICPSR 9593; Davis, Smith, Hodge, Nakao, & Treas, 1991). Young adults’ occupational prestige scores ranged from 20.05 (e.g.,
maids and housemen) to 73.70 (e.g., computer scientists).

**Covariates** include mothers’ and fathers’ education and age (in years) at Time 1, adolescents’ age at Time 2, and young adults’ age concurrent with their occupational attainment reports. For education, mothers and fathers reported on a scale where 12 = high school graduate, 13 = high school graduate plus vocational/technical/job training, 14 = some college but no degree, 15 = associate’s degree, 16 = bachelor’s degree, 17 = some education after undergraduate degree but no advanced degree, 18 = master’s degree, 19 = professional degree (e.g., Law, Medicine), 20 = Ph.D.

**Analytic Strategy**

For the analyses, I used Structural Equation Modeling (SEM) to test the mediation model in Figure 2.1. All study variables were standardized to obtain standardized path coefficients indicative of effect sizes. To index mother- and father-adolescent relationship quality, both parents’ and adolescents’ reports (observed variables) were loaded onto each dyadic relationship quality (latent variable). The two factor loadings for mothers’ and fathers’ reports, and the two for adolescents’ reports on mothers and fathers were respectively constrained as equal for measurement invariance between mother- and father-adolescent relationship quality. Further, adolescents’ academic performance, sense of control, and self-worth (observed variables) were loaded onto career adaptivity (latent variable). To free up the estimation of factor loadings on the observed variables, I fixed the variances of all latent variables (i.e., mother- and father-adolescent relationship quality, and career adaptivity) to 1.0 (Brown, 2006). Given the clustered nature of the data (i.e., siblings within families), I estimated the model with the Satorra-Bentler correction that can adjust the point and variance estimations and fit indices according to the non-independent, clustering structure of the data (Satorra & Bentler, 1994), using the ‘lavaan.survey’
package (Oberski, 2014) in R 3.5.1 (R Core Team, 2018). Consistent with Hu and Bentler’s (1999) criteria, model indices indicating goodness-of-fit included: (1) a non-significant result of the chi-squared test; (2) root mean square error of approximation (RMSEA) value close to or below .06; (3) root mean squared residual (SRMR) close to or below .08; and (4) comparative fit index (CFI) and the Tucker-Lewis index (TLI) close to or above .95. On the condition of the model’s goodness-of-fit, I evaluated mediation using the Monte Carlo method with 20,000 repetitions for the simulation and a 95% confidence interval (CI, Selig & Preacher, 2008) to test the indirect effects of mother- and father-adolescent relationship quality on young adult occupational prestige through career adaptivity.

To address the aim of examining differences in effects of mother- and father-adolescent relationship quality, I first conducted a model comparison using the Satorra-Bentler scaled chi-squared difference test (Satorra & Bentler, 2010) between the model wherein the path coefficients of mother- and father-adolescent relationship quality to career adaptivity were freely estimated and the model wherein these two coefficients were constrained as equal. That the model fits differ significantly would indicate a significant difference in mother- versus father-youth relationships’ direct effects on career adaptivity. Further, I compared the indirect effects of mother- and father-adolescent relationship quality on occupational prestige through career adaptivity using the Monte Carlo method with 20,000 repetitions. Finally, to test the moderating role of adolescent/young adult gender, I compared the model wherein all the coefficients were freely estimated across young women and men with the model wherein the coefficients were constrained as equal across gender. Significant chi-squared differences between the models indicate significant moderation by gender.

To handle missing values (1.2% to 7.6%) in the study variables measured at Times 1 and
2, I used augmented multiple multivariate imputation using chained equations (MICE) with 5 imputations by the R ‘mice’ package (van Buuren & Groothuis-Oudshoorn, 2011), for which all the study variables were included in the estimation equations. Mothers’ and fathers’ education and age at Time 1, adolescents’ age at Time 2 and young adults’ age were initially included as covariates, and nonsignificant covariates were removed from the final model.

Further, given that this may be the first study wherein career adaptivity was estimated as a latent construct with multiple observed indices, despite strong theoretical foundations, I conducted follow-up robustness checks. Specifically, I estimated three models to determine whether the significant mediation effect of career adaptivity resulted from a single, predominant index that explained the parent-youth relationship-prestige linkage as well or better than the career adaptivity latent construct. In these models, the three indices that comprised the career adaptivity latent variable, academic performance, sense of control, and self-worth, were treated as single mediators, replacing the career adaptivity latent construct in the original model. Results showing that the mediation effects of these individual indices were not as strong as that of the career adaptivity latent variable would indicate that none of the indices could replace the latent variable, and thereby provide evidence that this multi-index construct best explains observed relationship quality—occupational prestige linkages.

**Results**

**Preliminary Analyses**

Descriptive data for study variables (before standardization) are shown in Table 2.1. On average, both mother- and father-adolescent relationship quality, as reported by adolescents and parents, were high—well above the midpoint of the 5-point scale. Each of the three indices of career adaptivity, that is, academic performance, sense of control and self-worth, also had means
above the midpoint of the scale. Young adults’ average job prestige was 51.16, a score around that of administrators and officials (51.23) and managers (50.64).

At the bivariate level, positive correlations between mothers’ and adolescents’ reports of mother-adolescent relationships and between fathers’ and adolescents’ reports of father-adolescent relationships were at moderate levels (i.e., $r = .35$ and $r = .34$), which indicated that although parents and adolescents had somewhat different perspectives on their relationship quality, their reports can be captured by a latent dyadic relationship construct. Further, correlations between the three indices of career adaptivity were all significant, ranging from low ($r = .14$) to moderate ($r = .47$) levels and indicating that, although these were somewhat different characteristics, a latent construct, that is, career adaptivity in this study, may underlie them. As Table 2.1 also shows, adolescents’ reports of their relationships with mothers and fathers were significantly positively correlated with all the adaptivity indices, and fathers’ reports were positively correlated with adolescent academic performance. This pattern of stronger bivariate correlations within adolescents’ reports of family relationship experiences and their adaptivity characteristics may be due to the self-report biases and may lead to higher loadings on adolescents’ reports from the relationship quality latent construct in the final model. Also, consistent with previous findings about long-term effects of parent-adolescent relationship quality and young adult occupational attainment, adolescents’ reports of relationships with mothers and fathers’ reports of relationships with adolescents were both significantly and positively correlated with young adult occupational prestige. Finally, the correlations between the adolescent career adaptivity indices and young adult occupational prestige were all significant, suggesting that career adaptivity may be an important precursor of occupational attainment that merits further analysis.
The final mediation model is presented in Figure 2.3. Fit indices documented that the model fit the data well: $\chi^2(21) = 28.87, p = .12$; RMSEA = .04; SRMR = .05; CFI = .96; and TLI = .94. Preliminary tests indicated that father education at Time 1 was a significant covariate, including its significant association with career adaptivity and covariance with father-adolescent relationship quality, and thus was included in the final model; other covariates, including mother education, mother and father age, and adolescent and young adult age were nonsignificant and thus were removed. Both the relationship quality and career adaptivity latent constructs had moderate to high levels of factor loadings on the observed indices. Although adolescent reports obtained higher factor loadings than parents’ reports of relationship quality, both loadings were significant ($p < .001$), indicating that both adolescents’ and parents’ reports contributed significantly to the dyadic relationship quality construct. Residuals of mothers’ and fathers’ reports on their relationship quality with adolescents covaried at trend level, indicating a potential partner effect that differed from the dyadic relationship quality construct and that has been observed in multi-reporter, multi-member family relationship models (Cook, 2001). Further, the positive covariance between residuals of sense of control and self-worth was significant, indicating that some of the covariation between these two indices was due to source, such as self-report bias on Likert scales, rather than the career adaptivity latent construct (Brown, 2006). Thus, robustness checks, as mentioned above, that separately examine the mediational role of each of the three career adaptivity indices are merited to rule out the possibility that only one of the indices, such as academic performance, accounted for the mediating role of career adaptivity.
As shown in Figure 2.3, the path from mother-adolescent relationship quality to adolescent career adaptivity about one year later was both positive and significant, and so was the path from adolescent career adaptivity to occupational prestige in young adulthood, about 10 years later. Further, the Monte Carlo simulation procedure with 20,000 repetitions revealed a significant indirect effect from mother-adolescent relationship quality to occupational prestige through career adaptivity, \( B = .24, p = .027, 95\% CI = [.02, .52] \). This result was consistent with the hypothesis that career adaptivity in adolescence mediated the effects of mother-adolescent relationship quality on young adult occupational attainment. The nonsignificant direct path from mother-adolescent relationship quality to young adult occupational prestige further indicated that this was a full mediation mechanism. In particular, the size of this indirect effect was equivalent to the size of the direct effect of mother-adolescent relationship quality on occupational prestige, \( B = .24, SE = .12, p = .041 \), estimated in a model without any mediator.

In contrast, the path from father-adolescent relationship quality to adolescent career adaptivity was nonsignificant—as was the direct effect of father-adolescent relationship quality on young adult occupational prestige. Though as a covariate and not a focus of study hypotheses, father education had a significant and positive direct effect on career adaptivity and through adaptivity, an indirect effect on young adult occupational prestige, \( B = .15, p = .017, 95\% CI = [.03, .32] \).

**Differences Between Mother- and Father-Adolescent Relationship Effects on Daughters’ and Sons’ Career Development**

According to the final model, for the whole sample, the effect of mother-adolescent relationship quality on career adaptivity and in turn, occupational prestige was significant, whereas the effect of father-adolescent relationship quality was nonsignificant. The Satorra-
Bentler scaled chi-squared difference test, however, revealed that constraining the paths from mother- and father-adolescent relationship quality to career adaptivity did not significantly degrade the model fit, $\chi^2(1) = .26, p = .60$, and the indirect effects of relationship quality on occupational prestige through career adaptivity also did not differ significantly between mother and father, $B_{diff} = .10, p = .296$, 95% CI = [-.27, .49]. This result indicated that although the hypothesized mediation effect was significant for mothers and not for fathers, the difference between the mother and father effects was not statistically significant.

To test whether mother and father effects differed by youth gender, a multi-group analysis was conducted with youth gender as the group membership factor. The measurement models for the latent constructs, relationship quality and career adaptivity, did not differ by gender, $\chi^2(2) = 2.79, p = .25$ and $\chi^2(3) = 1.24, p = .74$. Further, model comparisons revealed that both the paths from mother- and from father-youth relationship quality to career adaptivity did not differ by adolescent gender, $\chi^2(1) = -.00, p = .99$ and $\chi^2(1) = .50, p = .48$, even though in multi-group analysis the effect of father-youth relationship quality on career adaptivity was significant for boys, $B = .46, SE = .22, p = .024$ but nonsignificant for girls, $B = .23, SE = .30, p = .450$. Nor did the path from career adaptivity to occupational prestige differ by gender, $\chi^2(1) = .94, p = .33$. In total these results suggest that adolescent gender was not a significant moderator for the mediation model.

**Robustness Checks: The Mediating Role of the Multi-Component, Latent Career Adaptivity Construct**

As mentioned above, to test whether any of the three career adaptivity indices could replace the full mediation role of the latent career adaptivity construct, three separate follow-up models were tested wherein academic performance, sense of control, and self-worth were single
mediators. Results showed that across the three models, none of the single observed indices fully mediated the link between mother-adolescent relationship quality and young adult occupational prestige. First, in contrast to the significant indirect effect through career adaptivity, indirect effects through the single indices were all at trend level: $B = .05, p = .055, 95\% CI = [-.01, .13]$ for academic performance, $B = .02, p = .091, 95\% CI = [-.00, .07]$ for sense of control, and $B = .02, p = .092, 95\% CI = [-.00, .06]$ for self-worth. Second, in contrast to the nonsignificant direct effect of mother-adolescent relationship quality on occupational prestige in the model with the latent career adaptivity variable (Figure 2.3), this direct effect was at trend level in each of the three models: $B = .18, SE = .11, p = .096$ in model with academic performance, $B = .19, SE = .10, p = .055$ in model with sense of control, and $B = .20, SE = .10, p = .051$ in model with self-worth. Together these results suggest that, despite the positive covariance between sense of control and self-worth, none of the three individual indices replaced the fully mediational role that career adaptivity played in the longitudinal links between parent-adolescent relationships and young adult occupational attainment.

**Discussion**

Both theory and empirical evidence have highlighted the role of adolescents’ family experiences in their career development and occupational attainment in young adulthood. Prior research has focused on either short-term family influences on youth career development or the main effects of early family experiences and (young) adults’ occupational attainment. In this study I extended prior research by using longitudinal data, spanning up to 12 years, to test a mediation mechanism whereby mother- and father-adolescent relationship qualities were linked to occupational prestige in young adulthood through adolescent career adaptivity, a fundamental component in youth career development as highlighted in the career construction theory.
(Savickas, 2002; 2013). Beyond the career construction theory, the life course perspective (Elder, 1998), the social capital theory (Coleman, 1988) and the expectancy-value achievement model (Eccles, 2011) also provided conceptual grounding for the current study.

Consistent with tenets in these frameworks, particularly career construction theory, results revealed that career adaptivity fully mediated the link between mother-adolescent relationship quality and young adult occupational prestige. Neither the direct nor the indirect effect through adaptivity was significant for father-adolescent relationship quality, however. Analyses focused on youth gender failed to reveal its moderation effect, though the pattern of results suggested potential gendered patterns that merit further examinations. With these findings, the current study contributed to explanations of the long-term implications of adolescents’ family relationships for young women’s and men’s occupational prestige, specifically through shaping career adaptivity in adolescence—an important developmental period for career development.

**Longitudinal Linkages Between Parent-Adolescent Relationship Quality and Young Adult Occupational Prestige Through Career Adaptivity**

By testing the hypothesized mediation model in Figure 2.1, this study advanced understanding of the career development mechanisms that underlie longitudinal links between parent-adolescent relationships and young adult occupational attainment. In particular, I found that career adaptivity, captured as a latent construct with academic performance, sense of control, and self-worth, fully mediated the effect of mother-adolescent relationship quality on young adult occupational prestige. This result is consistent with the career construction theory tenet, as yet not well-tested, that family relationships shape youth career development, especially career adaptation processes, and that these, in turn lay the foundation for adulthood career establishment (Savickas, 2002; 2013). It also is consistent with the important role of adolescents’
‘planful competence’, shaped by family experiences, in their future career success (Clausen, 1991). This finding adds to the extant literature by moving beyond prior studies of career aspirations and attainment value mediators to highlight career adaptivity as an important career development dimension that explains effects of adolescent family experiences on young adult occupational attainment. In other words, beyond their subjective orientations specifically toward career development, adolescents’ personal characteristics and competence play a significant role in connecting their family experiences, particularly their relationships with their mothers, to future attainment. Further, given that previous studies on family and youth career development have relied on cross-sectional or short-term longitudinal data within a single developmental period such as adolescence (Whiston & Keller, 2004; Bryant et al., 2006), the current study, with longitudinal data spanning across middle adolescence to young adulthood, documented that family influences on adolescent career development processes may extend, over a longer term into young adulthood, and manifest as effects on young adults’ actual occupational attainment.

More generally, this study provided evidence in support of the career construction theory and informs future research on career adaptation processes, especially career adaptivity. It builds on prior tests of the career construction theory which have focused on short-term effects of youth’s career adaptivity on other career development processes (Hirschi, et al., 2015; Šverko & Babarović, 2018; Zhou et al., 2016) to show that, beyond such short-term effects, career adaptivity in adolescence extended over an approximately ten-year period to predict young adults’ occupational prestige. In addition, prior studies have treated career adaptivity as an umbrella concept and tested effects of its components separately, and a literature search revealed no other study that captured career adaptivity as a latent construct with multiple indices (Rudolph et al., 2017). The study result that the career adaptivity latent variable, as compared to each
individual index, better explained the path from relationship quality to occupational prestige, provided support for the career adaptivity construct proposed in the career construction theory and highlighted the importance of examining career adaptivity as a latent construct in future studies. Nevertheless, it should be noted that beyond academic performance, sense of control and self-worth, other personal characteristics have been examined in research on career adaptivity, such as proactive and extraverted personality and future orientation (Rudolph et al., 2017). Thus an important further step is to build a more complete measurement model that includes the range of personal characteristics hypothesized to comprise career adaptivity.

Mothers’ Versus Fathers’ Roles in Career Adaptivity and Occupational Prestige

The findings also advance understanding of mothers’ and fathers’ roles in daughters’ and sons’ career development and attainment. As noted, research has rarely examined both mothers and fathers in the same models to reveal their potentially unique influences on youth career development and young adult occupational attainment. In addition, studies that have examined mothers’ versus fathers’ effects on youth career development and attainment have generally been limited to cross-sectional or retrospective, qualitative data, with mixed findings (Hoffman et al., 1992; Lim & Loo, 2003; Schultheiss et al., 2001). In this study, I was able to examine both mother- and father-adolescent relationship quality in the same model. The result showed that mother-adolescent relationship quality had a positive and close to medium size (standardized coefficient $B = .46$) direct effect on adolescent career adaptivity, and a positive and small ($B = .24$) indirect effect on young adult occupational prestige via career adaptivity. In contrast, these effects were nonsignificant for father-adolescent relationship quality, with its direct effect on career adaptivity only about 60% of the size of the mother-adolescent relationship quality direct effect. This pattern suggests that, in providing support for fostering career development with an
important type of the family’s social capital—positive and close relationships, mothers may play a more salient role, possibly due to their greater involvement in youth career development (Schultheiss et al., 2001). This result is also consistent with that from another cross-sectional study which showed that adolescents’ relationship experiences with mothers but not fathers were significantly linked to their self-efficacy and work attitudes (Lim & Loo, 2003).

Nevertheless, despite the difference in statistical significance of the path coefficients linking mother- versus father-relationship quality to career adaptivity, the model comparison and indirect effect difference test results showed that neither the direct nor the indirect effects were significantly different between mothers and fathers, and thus caution is needed in drawing conclusions about mothers’ and fathers’ distinctive roles in youth career development and attainment. Further, although not a focus of the study hypotheses, father’s education emerged as a significant covariate with a positive effect on adolescent career adaptivity, whereas mother’s education was nonsignificant (and thus omitted in the final model). This result indicated that, in comparison to the provision of social capital provided by mothers, fathers may matter more in the domains of human capital, as manifested by their education levels, and financial capital with their income that is usually tied to their education levels (Coleman, 1988). Future research should move away from focusing on the question of who plays a more important role, to identify what roles mothers and fathers play in their children’s career development.

Turning to results pertaining to youth gender, findings failed to support the hypothesis that same-gender parents would have stronger effects than other-gender parents on young women’s and men’s career development: In the face of gender differences in family socialization and youth career development and attainment (Eccles, 2011; Hegewisch et al., 2010; Maccoby, 1998), mother-adolescent relationship quality mattered for both daughters’ and sons’
occupational attainment in young adulthood through effects on their career adaptivity in adolescence. Nonetheless, a gendered pattern did emerge, with respect to the effect of father-adolescent relationship quality on adolescent boys’ career adaptivity, in keeping with the social learning hypothesis. Given that the overall youth gender moderation effect was nonsignificant, this result must, however, be interpreted cautiously. Further, these nonsignificant effects involving parent and youth gender may be due to a lack of power, given the sample size of the current study, and thus, a future direction is to test these effects with a larger sample.

**Limitations and Future Directions**

In the face of its contributions, this study’s limitations inform directions for future research. First, with the larger study’s goal of obtaining a homogeneous sample and ruling out some third variable explanations, the study sample was almost exclusively European American, two-parent and primarily dual-earner families, which limited the generalizability of findings. Family relationship quality may have implications of a different magnitude for youth occupational attainment and also may involve different mechanisms for youth from other racial/ethnic groups and family structures who are faced with additional challenges in career development and achievement (Diemer, 2007; Rocheleau, 2015). And, the family roles of mothers and fathers may differ from those of youth in this sample (Bosco & Bianco, 2005).

Further, adolescent girls and boys from different social classes can differ in their access to and utilization of social capital, including resources both in and outside family, which are important to their career choices and attainment (Hardie, 2015). Second, young adults in the sample were around 26 years old, and thus later bloomers and even longer term attainment were not examined. Future directions include further follow-up of these young adults into middle adulthood to examine whether and how family influences through adolescent career adaptivity
maintain. Third, despite theoretical foundation of the fundamental role of career adaptivity in the career adaptation process, other dimensions of career adaptation were not assessed and thus could not be tested. Thus, an important next step is longitudinal research that incorporates the four dimensions of the career adaptation process and the multiple indices of each to fully test the mediating role of career adaptation in family relationship—career attainment linkages. Relatedly, future research should move beyond the this study’s single-item, objective measure of occupational prestige to capture the multi-dimensional nature of career attainment such as salary and prestige as well as subjective experiences such as career satisfaction (Gordon & Cui, 2015; Ng, Eby, Sorensen, & Feldman, 2005). Finally, causal conclusions cannot be drawn from the current study given its correlational design. With this design, although theories guided testing of the mediation mechanism with paths from relationship quality to occupational prestige through career adaptivity, an alternative model with paths from adaptivity to prestige through relationship quality also is possible. Because career adaptivity was not fully measured at Time 1, cross-lagged panel analysis could not be conducted and this alternative direction of effect could not be studied here. Future longitudinal studies are needed that start early in development and repeatedly measure both career adaptivity and relationship quality, to test direction of effects as well as when the linkages observed in this study start to emerge. Intervention studies aimed at promoting family relationship quality and adolescents’ career adaptivity are best suited to illuminate their causal effects on youth career attainment.

Nonetheless, as noted, using longitudinal data over up to 12 years, this study added to the literature on how parent-adolescents may have long-term implications for young adult occupational attainment. Findings documented that the effect of mother-adolescent relationship quality on both young women’s and men’s occupational prestige was fully mediated by
adolescent career adaptivity, supporting the career construction theory and extending research on family and youth career adaptation processes. Further, inclusion of both mothers and fathers illuminated mothers’ powerful role in providing social capital resources within the family for both daughters’ and sons’ career development. These findings have practical implications for parent education and intervention programs targeted at youth career development and attainment.

In conclusion, this study contributed to theory and research on adolescent career development by testing tenets emphasizing career adaptation in youth career development to illuminate its role as a mechanism through which parent-adolescent relationships have implications for occupational attainment in young adulthood. Adolescence is an important period for career development wherein family experiences can have both short-term effects on career development processes such as adaptivity, and long-term implications for young adult occupational attainment—a key developmental outcome that is central to individuals’ well-being throughout adulthood (IOM & NRC, 2015).
References


Daly, K. J. (1996). Spending time with the kids: Meanings of family time for fathers. *Family...*
Relations, 45, 466–476.


<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mothers’ relationship reports with adolescents (Time 1)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Adolescents’ relationship reports with mothers (Time 1)</td>
<td>.35**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Fathers’ relationship reports with adolescents (Time 1)</td>
<td>.19**</td>
<td>.24**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Adolescents’ relationship reports with fathers (Time 1)</td>
<td>.15*</td>
<td>.51**</td>
<td>.34**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Academic performance (Time 2)</td>
<td>-.01</td>
<td>.18**</td>
<td>.18**</td>
<td>.16*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Sense of control (Time 2)</td>
<td>.03</td>
<td>.28**</td>
<td>.09</td>
<td>.23**</td>
<td>.14*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Self-worth (Time 2)</td>
<td>.01</td>
<td>.21**</td>
<td>.01</td>
<td>.22**</td>
<td>.17**</td>
<td>.47**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Young adult occupational prestige</td>
<td>.03</td>
<td>.20**</td>
<td>.15*</td>
<td>.08</td>
<td>.38**</td>
<td>.15*</td>
<td>.19**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Mother education (Time 1)</td>
<td>.01</td>
<td>.07</td>
<td>.18**</td>
<td>.08</td>
<td>.20**</td>
<td>.13*</td>
<td>.11†</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Father education (Time 1)</td>
<td>-.08</td>
<td>.01</td>
<td>.17*</td>
<td>.12†</td>
<td>.22**</td>
<td>.07</td>
<td>.11†</td>
<td>.08</td>
<td>.41**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Mother age (Time 1)</td>
<td>.05</td>
<td>.11†</td>
<td>.06</td>
<td>.03</td>
<td>.12†</td>
<td>-.00</td>
<td>.03</td>
<td>.08</td>
<td>.38**</td>
<td>.46**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Father age (Time 1)</td>
<td>-.01</td>
<td>.02</td>
<td>.05</td>
<td>-.01</td>
<td>.00</td>
<td>.02</td>
<td>.00</td>
<td>.11†</td>
<td>.31**</td>
<td>.30**</td>
<td>.66**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Adolescent age (Time 2)</td>
<td>-.17*</td>
<td>-0.04</td>
<td>-.09</td>
<td>-.09</td>
<td>-.01</td>
<td>.23**</td>
<td>.08</td>
<td>.02</td>
<td>.05</td>
<td>-.01</td>
<td>.01</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Young adult age</td>
<td>-.21**</td>
<td>-.08</td>
<td>-.02</td>
<td>-.04</td>
<td>-.00</td>
<td>.12†</td>
<td>.04</td>
<td>.04</td>
<td>.10</td>
<td>-.01</td>
<td>.02</td>
<td>.07</td>
<td>.65**</td>
</tr>
<tr>
<td>15</td>
<td>Adolescent/young adult gender</td>
<td>-.12†</td>
<td>-.13*</td>
<td>-.05</td>
<td>.04</td>
<td>-.01</td>
<td>.03</td>
<td>.17**</td>
<td>-.10</td>
<td>.04</td>
<td>.02</td>
<td>.03</td>
<td>.00</td>
<td>.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.08</td>
<td>3.55</td>
<td>3.75</td>
<td>3.31</td>
<td>3.30</td>
<td>35.32</td>
<td>3.05</td>
<td>51.16</td>
<td>14.83</td>
<td>14.83</td>
<td>42.67</td>
<td>45.06</td>
<td>16.05</td>
<td>26.19</td>
</tr>
<tr>
<td>2</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.46</td>
<td>.68</td>
<td>.47</td>
<td>.68</td>
<td>.67</td>
<td>3.71</td>
<td>.70</td>
<td>13.18</td>
<td>2.05</td>
<td>2.29</td>
<td>3.90</td>
<td>5.37</td>
<td>1.63</td>
<td>.96</td>
</tr>
</tbody>
</table>

Note. *Adolescent/Young adult gender: 0 = female, 1 = male. †p < .10; *p < .05; **p < .01
Figure 2.1. The mediation model linking mother- and father-adolescent relationship quality to young adult occupational prestige through adolescent career adaptivity. Covariates tested include mothers’ and fathers’ education and age at Time 1, adolescents’ age at Time 2 and young adults’ age.
Figure 2.2. The career adaptation process as proposed in the career construction theory (Savickas, 2002; 2013)
Figure 2.3. Estimated model with standardized path coefficients and standard errors (in parenthesis). Nonsignificant covariates, including mothers’ education and ages of mothers, fathers, adolescents and young adults, were omitted from the final model. Significance level: † p <.10, ***p<001, **p<.01, *p<.05.
CHAPTER 3

IMPLICATIONS OF MEXICAN-ORIGIN YOUTH'S WORK EXPERIENCES FOR RELATIONSHIPS WITH FATHERS

Introduction

Work is a key context for youth development. Currently, youth employment is widespread in the U.S.: Data from the U.S. Bureau of Labor Statistics (2018), for example, indicate that 50.7% of youth aged 16 to 24 were employed in 2017. And, a body of research has documented effects of youth employment and work experiences on a range of outcomes, including social and coping skills (Mortimer & Finch, 1996), behavioral adjustment (Mortimer, Finch, Ryu, Shanahan, & Call, 1996; McMorris & Uggen, 2000), concurrent academic achievement and future educational attainment (Hwang & Domina, 2017; Mortimer et al., 1996; Staff & Mortimer, 2008), and adult occupational outcomes (Carr, Wright, & Brody, 1996).

We know very little, however, about the work experiences of racial/ethnic minority youth (Hwang & Domina, 2017). Available data document that employment among Latino youth approximates the national average (48.9% in 2017; U.S. Bureau of Labor Statistics, 2018) and that, on average, these youth are employed in what is considered middle to high intensity work (i.e., over 5 hours per week). As such, the workplace is a salient context in the lives of many Latino youth (Hwang & Domina, 2017). From an ecological perspective (García Coll et al., 1996), however, the role of work in their development may differ from what has been learned from studies of majority Anglo youth, including because of Latino youth’s experiences of racial/ethnic segregation, wage inequality, and stereotyping and discrimination in the labor market and at work (Browne & Misra, 2003; Hwang & Domina, 2017). Study of the work experiences of Latino youth and implications of work for their functioning in other domains is
therefore imperative.

Accordingly, in this study I adopted an ethnic homogenous approach to examine the work experiences of Mexican-origin youth and their implications for youth’s family relationships. Although a body of literature has documented effects of work on family experiences—a process termed, work-family spillover (Perry-Jenkins & Wadsworth, 2017)—the focus has been predominantly on adults, and we know much less about spillover among employed youth. Theories, however, direct attention to the interface between youth work and family as a process warranting investigation. According to the ecological theory, for example, the exchanges between youth work and family constitute a key mesosystem for youth development (Bronfenbrenner, 1986; Zimmer-Gembeek & Mortimer, 2006), and a cultural ecological perspective highlights the role of sociocultural stressors and supports for minority youth (García Coll, et al., 1996; McLoyd, 1998). In addition, family systems theory highlights family as an open system shaped by family members’ experiences outside the home (Minuchin, 1985)—such as youth’s work.

Mexican-origin youth are an important focus of research on work-family spillover because of their engagement in the work, but also because their culture emphasizes interdependence among family members, making family relationships especially influential in their lives (García Coll & Vázquez García, 1995; Updegraaff, McHale, Whiteman, Thayer, & Delgado, 2005). In this study, I used longitudinal data to examine whether and how Mexican-origin youth’s work experiences, including work hours and workplace discrimination, were linked two years later to their relationship quality with fathers—the primary breadwinners in many Mexican-origin families—controlling for concurrent relationship quality and fathers’ corresponding work experiences. I also tested moderation of these linkages by youth gender and mothers’
employment status as informed by theories and previous findings about different effects of work on men’s versus women’s family experiences (Roehling, Jarvis, & Swope, 2005; Wheeler, Updegraff, & Crouter, 2015) and mothers’ central role in regulating father-youth dynamics (Puhlman & Pasley, 2013).

**Positive and Negative Sides of Youth Work Effects on Family Relationships**

The ecological model highlights the interface between work and family as a mesosystem for individual development, including for youth who work (Bronfenbrenner, 1986; Zimmer-Gembeck & Mortimer, 2006). This tenet is echoed by the family systems theory that family is an open system subject to influences from the external contexts, among which the work context plays a critical part (Minuchin, 1985). Reflecting the theories, literature on adult work-family interface has accumulated in recent decades (Perry-Jenkins & Wadsworth, 2017). Collectively this literature has directed attention to both positive and negative effects of work on family relationships, termed work-family enrichment (Greenhaus & Powell, 2006) and conflict (Frone, Russell, & Cooper, 1992), respectively.

Although enrichment and conflict have not been systematically examined among youth, literature has identified both positive and negative elements of youth work that can directly or indirectly affect their family lives. On the positive side, youth’s paid work can provide a financial contribution to the family’s welfare, adding to the household budget and reducing youth’s financial dependence on parents (Shanahan, Elder, Burchinal, & Conger, 1996a; 1996b). Work can also promote youth’s social skills, such as through their communication with supervisors and cooperation with co-workers, and coping skills from experiencing challenges in both job tasks and interpersonal relationships at work. These positive effects may benefit family interactions, including communication with family members and coping with family conflict and
challenges (Mortimer & Finch, 1996; Mortimer & Staff, 2004). Further, a positive work environment may serve as an “arena of comfort” for youth that alleviates the impact of stressors, including those from the family life and its surrounding contexts (Call, 1996).

In contrast, consistent with the literature on adult work-family conflict, youth’s work can take their time away from home and family; moreover, negative experiences at work, such as pressure and discrimination, may result in fatigue and psychological distress that interfere with effective coping and family obligations (Perry-Jenkins & MacDermid, 2013). In addition, via their employment, youth may assume a more adult-like, independent role in the family (Zimmer-Gembeck & Mortimer, 2006), which may challenge parents’ power and authority and thus create tensions (Fuligni, 1998). These negative implications can be especially salient for Mexican-origin families given the cultural emphases on youth’s family obligations (Fuligni, Tseng, & Lam, 1999), interdependence among family members (García Coll & Vázquez García, 1995), and parental (especially paternal) authority (Fuligni, 1998).

In this study, I focused on examining youth work effects on their relationship quality with fathers. Fathers in Latino or Mexican-origin families are often the sole or primary breadwinners in line with the cultural stereotype of *machismo* (Arciniega, Anderson, Tovar-Blank, & Tracey, 2008; Casas, Wagenheim, Banchero, & Mendoza-Romero, 1994; Pinto & Coltrane, 2009; Wildsmith, Ramos-Olazagasti, & Alvira-Hammond, 2018). As such, they may be especially susceptible to influences of youth paid work. A working youth can share the father’s financial responsibility, which can alleviate his stress but also threaten his culturally stereotypical authority and power at home (Casas et al., 1994). The culturally grounded concept of *machismo* highlights fathers’ role in protecting the family from the outside world, including from youth’s workplace stressors, possibly making fathers more reactive than other family members to youth’s
experiences (Arciniega et al., 2008). Thus, tests of the links between these youth’s work experiences and their relationships with fathers can illuminate their work-family spillover, an understudied process.

**Youth Work and Father-Youth Relationships: Potential Curvilinear Linkages**

The coexistence of both positive and negative implications of youth work may be one basis for mixed findings among prior studies—which have led to debate about how youth employment affects their life and development. These discussions have directed attention to effects of specific dimensions of youth’s experiences at work on their outcomes, such as work hours, autonomy and control, and task complexity, and schedule conflict between school and work (Zimmer-Gembeck & Mortimer, 2006). Among these work experience dimensions, work intensity, indexed by work hours, has been studied the most. Research so far has suggested that low to moderate intensity (e.g., less than 20 hours) may be optimal across a range of youth outcomes, compared to both non-work and high intensity. For example, longitudinal analyses using a nationally representative sample have revealed that, compared to non-working youth, youth who were working less than 20 hours per week had higher grades and fewer school problem behaviors and were less likely to smoke, whereas youth who were working more than 20 hours had lower grades and were more likely to drink, controlling for gender and other background characteristics (Mortimer et al., 1996). Using the same sample, another study found that working less than 20 hours was associated with youth’s more positive communication with fathers, compared to the non-workers, but this association was nonsignificant for communication with mothers (Shanahan et al., 1996b). Further, a recent study indicated that among Latino youth, high-intensity work in adolescence (>20 hours) predicted lower likelihood of obtaining a Bachelor’s degree whereas neither low (<5 hours) nor moderate (5 to 20 hours) intensity had any
effect (Hwang & Domina, 2017). Overall, this pattern of previous findings, in combination with the complexities in youth work effects, suggests a potential curvilinear association between youth work hours and their outcomes, including those in their family life—such as relationships with fathers. Nevertheless, most previous studies used dichotomous variables to capture work hours and compared different intensity levels of work to the non-work status, but seldom examined work hours as a continuous variable to characterize working youth. Thus, these prior studies failed to identify the optimal number of work hours for particular outcomes in any group of youth.

Accordingly, in this study I tested a curvilinear association, potentially in an inverted U-shape pattern, between youth work hours and father-youth relationship quality among employed Mexican-origin youth. In a low to moderate range, the positive effects of work hours on father-youth relationship quality may manifest as youth contribute to the family financially (Shanahan et al., 1996a; 1996b) and gain coping skills and psychological resources that can benefit family interactions (Call, 1996; Mortimer & Finch, 1996). Past a turning point, however, in the moderate to high range, negative effects may outweigh positive effects and manifest as a negative association between work hours and relationship quality. As mentioned above, by working longer hours, youth may experience more time conflicts with their family obligations and experience more psychological distress that spills over to negative family interactions (Perry-Jenkins & MacDermid, 2013). Further, they can also display more independence and power assertion (Zimmer-Gembeck & Mortimer, 2006), which conflict with the cultural emphasis on family member interdependence and challenges fathers’ authority and sole/primary breadwinner role in Mexican-origin families (Casas et al., 1994; Fuligni, 1998; García Coll & Vázquez García, 1995; Pinto & Coltrane, 2009).
Although work experiences are multi-dimensional and youth work effects are not exclusive to work hours, other dimensions have received limited attention in prior research (Zimmer-Gembeck & Mortimer, 2006). Accordingly, in this study in addition to work hours, I tested the effect of youth workplace discrimination on their relationship quality with fathers. Racial/ethnic discrimination at work is a dimension unique to minority employees, including Latino youth. From a social justice perspective, such discrimination should be eliminated; in reality it is pervasive in workplaces (Deitch et al., 2003). On the condition of its existence, research has documented its negative effects, as well as detected some potentially positive effects of workplace discrimination on family life. On one hand, negative effects of minority youth’s experiences of discrimination are evident in their psychological well-being, including negative affect, depression, anxiety (Schmitt, Postmes, Branscombe, & Garcia, 2014), and substance use (Okamoto, Ritt-Olson, Soto, Baezconde-Garbanati, & Unger, 2009), all of which may have negative implications for their family relationships. On the other hand, research on youth resilience suggests that negative experiences at work can foster youth’s psychological and coping resources (Mortimer & Staff, 2004), which may also benefit family interactions, including coping with family conflicts and challenges. Moreover, family members can play a protective role in youth’s discrimination experiences, with those experiences inducing family support to the benefit of family relationships (Noh & Kaspar, 2003). In Mexican-origin families, this mechanism may be especially likely for youth and their fathers due to fathers’ role as protector and their own experiences of workplace discrimination (Arciniega et al., 2008). Therefore, although I found no prior study that explicitly examined the effect of youth’s workplace discrimination on their family experiences, parallel to the hypothesis about the effects of work hours, I also tested whether a curvilinear association (specifically, an inverted U-shape pattern)
described the association between youth workplace discrimination and youth-father relationship quality, such that in a low to moderate range, discrimination was positively related, but above this range, discrimination would be more harmful than beneficial and thus characterized by a negative association with father-youth relationship quality. Also, as in the examination of the implications of work hours, in testing these curvilinear links, the goal was to not only illuminate the complexities of youth work effects, but also to reveal the optimal or turning points in work-family spillover effects.

**The Roles of Youth Gender and Mother’s Employment Status**

According to the ecological model, mesosystem processes interact with individuals’ personal characteristics (Bronfenbrenner, 1986). In this study, I examined the role of youth gender in work-family interface effects, a personal characteristic that is intertwined with youth’s family and work experiences (Zimmer-Gembeck & Mortimer, 2006). In addition, although the focus of this study was on the implications of youth work for father-youth relationship quality, I tested the potential moderating effects of mothers’ employment status based on the family systems tenet that family members’ experiences are interdependent (Minuchin, 1985). I expected that these potential moderation effects would exhibit different patterns, depending on whether the links between youth work experiences and father-youth relationship quality were curvilinear. If curvilinear, the moderators may shift the optimal or turning points of the hypothesized inverted U-shape linkage: Given that the turning point is the level of work hours or workplace discrimination (i.e., the independent variables) at which negative outweigh positive effects, the turning point should be at higher levels under conditions wherein positive effects are stronger and/or negative effects are weaker. If instead, the observed linkages failed to conform to a curvilinear pattern, moderation effects would condition the direction and strength of the linear
associations between youth work experiences and their relationships with their fathers.

I first tested youth gender as a moderator. Prior studies of adults highlight the role of gender in work-family spillover noting that in general, the work role is more central to men’s identities whereas the family role is more central to women’s (Cinamon & Rich, 2002). This gender difference is magnified in Latino families given that gender-stereotypical roles are highlighted in their culture and family socialization (Roehling et al., 2005; Pinto & Coltrane, 2009). In line with this cultural gender expectation, compared to daughters, scripts for sons’ roles may involve longer work hours but fewer household responsibilities. Although no studies have yet examined work and family processes among Latino youth, research on Latino parents has found higher levels of work-family conflict among women than men (Roehling et al., 2005; Wheeler et al., 2015). Correspondingly, compared to sons, by working longer hours, daughters from Mexican-origin families may be more likely to violate cultural expectations and negative work-family spillover may manifest as less positive relationships with fathers. Thus, in the curvilinear association linking work hours and father-youth relationship quality, the turning points where negative effects outweigh positive effects may appear at lower levels for daughters than for sons; instead, if linear, the association should be more negative for daughters than for sons. In contrast, the positive, resilience mechanism linking workplace discrimination to family relationships may be stronger for daughters than for sons given that among Latino youth, young women are more likely to seek support from parents to alleviate culture-related stressors, including discrimination, than are young men (Crockett et al., 2007). In this case, the turning point of the curvilinear effect of workplace discrimination would be at a higher level for daughters than for sons, or the linear effect should be more positive for daughters than for sons.

I also examined mothers’ employment status as a potential moderator of the links between
youth work experiences and their relationship quality with fathers. Prior research characterizes mothers as youth’s primary caregivers, who play a key role in regulating father-youth dynamics (Puhlman & Pasley, 2013). Given that family processes are interdependent (Minuchin, 1985), youth’s work-family spillover and father-youth dynamics may depend in part on mothers’ experiences. In particular, maternal employment has been an important context for father-youth relationship, such that fathers are expected to take on more caregiving responsibilities, including Mexican-origin families, when mothers are also working (Formoso, Gonzales, Barrera, & Dumka, 2007). Especially relevant to this study, Formoso and colleagues (2007) found that maternal employment mitigated the negative effects of family stressors on father-youth relationship quality. Accordingly, maternal employment may also reduce negative spillover from youth work experiences to father-youth relationship quality. Further, maternal employment may be linked to less stereotypical family roles, including for fathers in Latino families (Pinto & Coltrane, 2009), thereby desensitizing fathers to the potentially threatening effects of youth’s work. Therefore, the negative effects of youth work hours and workplace discrimination on their relationship quality with fathers may be less prominent in families with employed mothers, such that the turning points of the curvilinear associations emerge at higher levels or the linear effects prove more positive.

**The Present Study**

In sum, grounded in the ecological model that highlights mesosystems for youth development and literature on the work-family interface among adults, this study used longitudinal data from Mexican-origin families to investigate the effects of work experiences, including work hours and workplace discrimination among youth in later adolescence and young adulthood, on their relationship quality with fathers two years later. The tests controlled for the
father-youth relationship quality concurrent with youth work experiences and fathers’ corresponding work experiences, in order to account for their potential confounding effects on relationship quality. The cultural emphasis on fathers’ sole or primary breadwinning role and authority in the family served as basis for this study’s focus on father-youth relationship quality. Informed by the potential co-existence of both positive and negative aspects of youth work effects as well as previous findings about curvilinear associations between youth work intensity and educational and familial outcomes, I hypothesized main effects of these work experiences to be curvilinear, in inverted U-shape patterns, such that each linkage would be positive in a low to moderate range until an optimal or turning point, after which the linkage would become negative.

Further, based on the person-context interaction tenet of the ecological model and the family systems tenet that family processes are interdependent, I tested youth gender and mothers’ employment status as potential moderators of the linkages between youth work experiences and father-youth relationship quality. On the condition of these linkages being curvilinear, moderations could shift turning points at which negative effects outweigh positive effects by changing the strengths of these effects; whereas in the case of linkages not being curvilinear, moderation would emerge for the linear effects of work hours and workplace discrimination. In particular, I expected the turning point for work hours to be at lower levels, and/or the linear effect of work hours to be less positive or more negative, for daughters with non-employed mothers, and the turning point for workplace discrimination to be at lower levels, and/or the linear effect of discrimination to be less positive or more negative, for sons whose mothers were non-employed. In addition to Time 1 father-youth relationship quality and fathers’ work experiences, controls included youth age, whether they lived at home, whether they were attending school, and parents’ and youth’s immigration status (i.e., nativity) given that these have
been associated with likelihood and intensity of paid work (Wiemers, 2014; Zimmer-Gembeck & Mortimer, 2006), parent expectations for youth achievement (Glick & White, 2004), and relationship quality with parents (Seiff-Krenke, 2006).

Method

Participants

I used data from a larger longitudinal study of 246 Mexican-origin families (i.e., mothers, fathers) raising adolescent/young adult offspring (i.e., a target 7th grader and an older sibling; author citation). The larger study was aimed at illuminating family socialization in adolescence through young adulthood (Phase 1, 2002/2003; Phase 2; 2007/2008; Phase 3, 2009/2010), and the role of sociocultural factors in these processes (Updegraff et al., 2005). The larger study started by recruiting Mexican-origin families through schools in a southwestern metropolitan area. Eligibility criteria were: (1) families had to have mother, father, younger sibling in the 7th grade and older sibling under the age of 21 who lived at home; (2) mothers self-identified as Mexican or Mexican American and were the biological mother of both participating siblings; and (3) fathers were the biological or long-term adoptive father (a minimum of ten years) who reported working at least 20 hours/week. Although not a criterion, 93% of fathers also were of Mexican descent. At Phase 1, 18.3% of participating families met federal poverty guidelines, similar to the 18.6% of two-parent Mexican-origin families living in poverty in the county from which the sample was drawn (U.S. Census Bureau, 2000).

The present study used data at Phases 2 and 3 (termed Times 1 and 2 hereafter) and includes families (N = 127) wherein fathers and at least one participating youth were employed at Time 1 for the study purpose of examining youth work experiences and accounting for fathers’ work experiences. T-tests and chi-squared analyses revealed that, under this inclusion criterion,
other than mother, father, and youth age and older sibling gender and nativity, families examined in this study differed from those excluded on Phase 1 on demographic characteristics. At Phase 1, compared to excluded families, in included families: family income was higher, $M = $62,499 ($SD = $50,754) v. $43,241 ($SD = $36,500), $t (229) = 3.43, p < .001; mothers and fathers were more educated, $M = 11.02 ($SD = 3.72) v. 9.61 ($SD = 3.58), t (244) = 3.02, p = .003 and $M = 10.43 ($SD = 4.45) v. 9.28 ($SD = 4.23), t (243) = 2.07, p = .04, for mothers’ and fathers’ years of education, respectively; more mothers and fathers were born in the U.S., 36.2% v. 21.8%, $χ^2 (1) = 5.45, p = .02, and 35.4% v. 22.0%, $χ^2 (1) = 4.70, p = .03, for mothers’ and fathers’ U.S. nativity, respectively; target younger siblings were often born in U.S., 68.5% v. 55.9%, $χ^2 (1) = 3.90, p = .04 and male, 57.5% v. 43.7%, $χ^2 (1) = 4.13, p = .04. Thus, I included these variables as controls in the following analyses.

A total of 187 youth (52.4% female, 64.7% born in U.S., 50.8% older siblings) who were employed at Time 1, were included in the analyses. At Time 1, youth’s average age was 19.33 ($SD = 1.78$, range = 17-25) and 64.2% were attending school (55.5% in high school, 3.4% in technical school, 21.3% in community college and 15.4% in four-year universities) and 77.5% were living at home. For youth who were not attending school, the average highest level of education was 12.18 ($SD = 1.16$, range = 12-16; on a scale where 12 = high school graduate and 16 = college degree). Average age was 46.92 ($SD = 5.37$) for fathers and 44.44 ($SD = 4.58$) for mothers, respectively. The median family income was $60,000 ($SD = $55,214, range = $7,000 - $315,000) at Time 1, and 68.5% mothers were employed. Most youth (89%), 40.1% of fathers, and 57.9% of mothers were interviewed in English and the rest in Spanish.

**Procedure**

Families participated in structured in-home interviews lasting two to three hours at each
time point. Bilingual interviewers conducted the interviews separately with each family member, entering data into laptop computers. Families received a $125 honorarium for participation at Time 1. At Time 2, each family member received a $75 honorarium for his or her participation. The University’s Institutional Review Board approved all procedures.

Measures

All measures were forward translated into Spanish and back translated into English by two separate individuals. Final translations were reviewed by a third, native Mexican-origin translator, and discrepancies were resolved by the research team (Knight, Roosa, & Umaña-Taylor, 2009).

Work experiences. Youth’s and fathers’ work hours and workplace discrimination were measured at the Time 1 home interviews. Work hours were measured by asking participants how many hours they spent on work across all jobs each week. The measure of workplace discrimination was created by combining Hughes and Dodge’s (1997) measures of Institutional Discrimination and Interpersonal Prejudice in the Workplace, which were correlated ($r = .46$, $p < .001$; author citation), to form a single scale. Using a 4-point scale ($1 = strongly disagree$, $4 = strongly agree$) youth and fathers rated their agreement with 12 items such as, “Mexicans or Mexican Americans get the least desirable assignments,” And, “People you work with assume that Mexicans or Mexican Americans are not as competent as others.” Item ratings were averaged to create a mean score, with higher scores reflecting more discrimination at work. Cronbach’s alphas were .88 and .91 for youth and fathers, respectively. In the analyses, youth’s work experiences were treated as independent variables and fathers’ experiences as controls.

Father-youth relationship quality. At Times 1 and 2, both youth and fathers separately reported on their relationship quality. Youth’s reports were assessed using the Acceptance Scale
of the Child’s Report of Parental Behavior Inventory (CRPBI, Schludermann & Schludermann, 1970), and fathers’ reports were assessed using the Parent Version of this scale (Schwarz, Barton-Henry, & Pruzinsky, 1985). This scale has been cross-validated for ethnic and language equivalence on a Hispanic sample (Knight, Tein, Shell, & Roosa, 1992). Youth and fathers responded to 8 items such as, “My father/I understand(s) my/my child’s problems and worries,” using a 5-point scale (1 = almost never, 5 = almost always). Item ratings were averaged to create a mean score, with higher scores reflecting better relationship quality, and Cronbach’s alphas range from .84 to .94 across times and reporters. In following analyses, youth’s and fathers’ reports at Time 2 were treated as dependent variables and their reports at Time 1 were included as controls.

**Moderators.** Youth gender was coded as 0 = female and 1 = male. Maternal employment was determined by mothers’ response to the question, “Are you currently employed?” at Time 1 and was coded as 0 = non-employed and 1 = employed.

**Control variables.** In addition to fathers’ work experiences and father-youth relationship quality at Time 1, these included youth age and whether youth were living at home (0 = no, 1 = yes) and attending school (0 = no, 1 = yes), as well as family demographic characteristics. As noted above, these included family income, mothers’ and fathers’ education levels and parents’ and youth’s nativity (0 = born outside U.S., 1 = born in U.S.).

**Analytic Strategy**

I applied multivariate multilevel modeling to include multiple dependent variables (i.e., youth’s and fathers’ reports of relationship quality) in each model and to use clustered data (i.e., siblings within families (Snijders & Bosker, 2012). To apply this modeling approach, I created a stacked dataset as illustrated in Table 3.1: Two reports of each relationship quality were stacked
into one outcome variable for analysis, with the dichotomous “reporter” variable created to distinguish fathers’ and youth’s reports. Further, experiences of older and younger siblings were also stacked, marked by the dichotomous “birth order” variable. As a result, for each family there were four lines of data for analysis if both siblings were included in this study. I applied two-intercept, two-level models where intercepts and coefficients were estimated for youth’s and fathers’ relationship reports separately while father-youth dyadic interdependencies were taken into account by using heterogeneous compound symmetry in residual estimation. For each intercept, random effects were estimated to allow for within-family variances between siblings. All continuous independent and control variables were centered around their sample means.

Two models were tested for the main effects of work experiences and moderation effects, respectively. All continuous independent variables were centered before analyses. Model 1 tested curvilinear associations between work hours and workplace discrimination at Time 1 and father-youth relationship quality at Time 2 by estimating a linear and a quadratic term for each work experience variable with significant quadratic terms indicating curvilinear associations. Controls included fathers’ work experiences and relationship quality reports at Time 1 in order to account for youth work-family contemporaneous associations and father work-family spillover, as well as demographic characteristics at Time 1 (i.e., youth age, living at home, attending school, family income, mothers’ and fathers’ education levels, and parents’ and youth’s nativity) as mentioned above. On the condition of significant quadratic terms, the turning point of each curvilinear association was obtained at the level where the first derivative of the regression equation equals zero, and the precision of each turning point, that is, the 95% confidence interval (CI) estimated using the delta method (i.e., approximating the distribution of the turning point based on a Taylor expansion; Plassmann & Khanna, 2007). Further, simple slopes of each curvilinear association
were tested with the Johnson-Neyman (JN) technique (Miller, Stromeyer, & Schwieterman, 2013). In Model 2, moderation by youth gender and mothers’ employment status at Time 1 were tested by entering interaction terms between these moderators and the linear term of each youth work experience variable. Nonsignificant quadratic terms were eliminated in this model to prevent confounding with the tests and interpretations of the moderations. In Model 2, interactions with the linear terms, in combination with simple slope follow-up tests, revealed whether the moderators shifted the turning points of the curvilinear associations (on the condition of the existence of curvilinear associations, that is, significant quadratic terms), or whether the moderators conditioned the directions and/or strengths of the linear effects of youth work experiences (on the condition of nonsignificant quadratic terms). Nonsignificant interactions and control variables were omitted from the final models (Aiken & West, 1991). Analyses were conducted using PROC MIXED in SAS Version 9.4 (SAS Institute Inc., Cary, NC).

Results

Preliminary Analyses

Descriptive data for study variables (before centering) are shown in Table 3.2. On average, youth were employed part-time (worked hours less than 35 hours/week), though there was considerable variation, with a range from 1 to 100 hours per week. I recoded one outlier for the work hours variable (100 hours per week), into the second highest score reported (75 hours per week), in order to prevent this case from creating a spurious association. Their average workplace discrimination was below the midpoint of the 4-point scale. Fathers were, on average employed full-time (range = 8-80 hours/week) and had a mean workplace discrimination level around the midpoint of the 4-point scale. Finally, relationship quality as reported by both youth
and fathers was high, well above the midpoint of the 5-point scale. The bivariate correlation between youth work hours and workplace discrimination was low ($r = .23$), indicating that these are different dimensions of work experiences. For relationship quality, although within-person cross-time correlations were high ($r = .76$ for both fathers’ and youth’s reports), there were still changes in rank order as indicated by moderate between-person correlations ($r = .44$ to .55), indicating that youth and fathers have somewhat different perspectives on their relationship quality. Correlations between youth work experiences and relationship qualities were all nonsignificant except for that between youth workplace discrimination and youth-reported relationship quality at Time 1 ($r = -.15$), indicating that those associations may be nonlinear or subject to moderation and merit examination via tests of both linear and quadratic terms as well as interactions with the potential moderators.

**Associations Between Youth Work Hours and Workplace Discrimination and Father-Youth Relationship Quality**

The results of the multivariate multilevel analyses testing the curvilinear associations between youth work hours and workplace discrimination and father-youth relationship quality are shown in Table 3.3. Among control variables, mother’s nativity was significantly linked to fathers’ reports of relationship quality, and Time 1 relationship quality reports were significantly linked to Time 2 reports, but other effects were all nonsignificant. Thus, the final models only included mother’s nativity (as linked to fathers’ reports of relationship quality) and Time 1 fathers’ and youth’s reports of relationship quality as the control variables. In addition, moderators were all nonsignificant in Model 1 and thus eliminated. The test of Model 1 revealed that, with Time 1 relationship quality controlled and accounting for interdependencies in fathers’ and youth’s reports, youth workplace discrimination at Time 1 had a significant curvilinear effect
on fathers’ reports of relationship quality at Time 2 (i.e., significant quadratic term). As shown in Figure 3.1, this curvilinear link conformed to an inverted U-shape pattern, consistent with the study hypothesis regarding the main effect of youth workplace discrimination on father-youth relationship quality. The equation for the link between youth workplace discrimination (x) and fathers’ reports of relationship quality (y) in this model is: \( y^2 = -0.213x^2 - 0.006x + 4.050 \). Setting the first derivative of this equation equal to 0, in combination with the delta method, the result revealed the turning point of this link to be at 1.85, with 95% CI = [1.53, 2.17], indicating that relationship quality was at the highest level when workplace discrimination was at this point.

Follow-up with the JN technique showed that within this curvilinear link, the simple slope was significant and positive between scores from 1 to 1.15 (i.e., 95% CIs all above 0), nonsignificant between scores of 1.15 and 2.16 (i.e., 95% CIs across 0), and significant and negative between scores of 2.16 and 4 (i.e., 95% CIs all below 0). This pattern indicated that, for youth workplace discrimination at Time 1: At a very low level (1—1.15), discrimination was positively linked to Time 2 fathers’ reports of father-youth relationship quality, at a moderate level (1.15—2.16), the positive and negative effects of discrimination balanced out, and at a high level (2.16—4), the negative effects of workplace discrimination outweighed potential positive effects. Inconsistent with the hypothesis for youth work hours, however, neither the quadratic nor the linear term were significantly associated with relationship quality reports.

**Moderation by Youth Gender and Mothers’ Employment Status**

As shown in Table 3.3, Model 2 is the final model that emerged in tests of moderation of the links between youth work experiences and father-youth relationship quality by youth gender and mothers’ employment status. The nonsignificant quadratic terms revealed in Model 1, including that of work hours for both youth’s and fathers’ reports of relationship quality, and that
of workplace discrimination for youth’s reports of relationship quality, were eliminated from Model 2. This means that moderation tests were focused on whether and how gender and mothers’ employment status conditioned the linear effects of work hours on youth and father relationship reports, and the linear effect of discrimination of youth relationship reports. Given the curvilinear effect in Model 1, for the link between workplace discrimination and fathers’ relationship reports, the moderation tests were focused on whether the moderators could shift the turning point of this linkage. Nonsignificant interactions were omitted from the final model.

Results of Model 2 revealed that both gender and mother employment moderated the linear effect of work hours on youth reports of relationship quality. Follow-up tests of gender moderation showed that for sons, work hours had a trend level negative effect on their perceptions of relationship quality with fathers, simple slope: $\gamma = -.008, SE = .003, p = .052$, whereas this effect was nonsignificant for daughters, simple slope: $\gamma = .007, SE = .005, p = .140$. Further, follow-up tests of the mother employment moderation showed that when mothers were not employed, work hours had a trend level positive effect, simple slope: $\gamma = .011, SE = .006, p = .079$; in contrast, when mothers were employed, the linear effect was negative at a trend level, simple slope: $\gamma = -.007, SE = .004, p = .066$. Both of these results were inconsistent with the hypotheses in suggesting that the effects of work hours on father-youth relationship quality were more negative for sons and for youth whose mothers were employed.

Turning to workplace discrimination, a significant interaction emerged between the workplace discrimination linear term and youth gender in relation to youth reports of father-youth relationship quality. Follow-up tests showed that this linear effect was positive for daughters, simple slopes: $\gamma = .213, SE = .106, p = .047$, but was nonsignificant for sons, simple slopes: $\gamma = -.191, SE = .124, p = .124$, which was consistent with the hypothesis that the effect of
workplace discrimination could be more positive for daughters than sons. Finally, neither of the
moderations were significant for workplace discrimination in relation to fathers’ reports,
meaning that the moderators did not shift the turning point in this curvilinear association.

Robustness Check: Results for Mother-Youth Relationship Quality and Power Analysis

Given that the study focus on father-youth relationship quality was based on literature
regarding the cultural emphasis on fathers’ sole or primary breadwinning role and *machismo*
values, the effects of youth work experiences revealed above should be unique to father-youth
relationship quality. To explore this premise, the study models were re-estimated with a focus on
youth’s and mothers’ reports of mother-youth relationship quality at Time 2 as the dependent
variables, which were measured using the same approach as father-youth relationship quality
reports. Estimation of the models revealed that all the main effects tested and the interaction
effects found for father-youth relationship quality were nonsignificant in the models for mother-
youth relationship quality, suggesting that the study findings were unique to fathers in these
Mexican-origin families.

Given the study’s relatively small sample size and the high correlations between the Time 1
and Time 2 relationship reports, some nonsignificant interaction patterns may have been due to
limited power. Therefore, in an additional follow-up step I conducted a power analysis focused
on the multivariate multilevel model where all the interactions between the two work experience
linear terms and the two moderators in relation to fathers’ and youth’s reports (in total, eight
interactions) were estimated. I used the R ‘simr’ package (Green & MacLeod, 2016) to obtain the
statistical power of the interaction terms based on Monte Carlo simulation with 1,000 repetitions.
Results showed that power estimates for the significant interaction terms in Model 2
were .38, .07, and .82, that is, only the interaction between workplace discrimination and youth
gender in relation to youth’s reports of relationship quality had high power (i.e., over .80). Further, for nonsignificant interactions, power estimates ranged from .05 (interaction between work hours and mother employment in relation to fathers’ reports) to .48 (interaction between workplace discrimination and mother employment in relation to youth’s reports), all in low to medium levels. Overall, the power analysis revealed that, except for one interaction, none of the other interactions tested in Model 2 had high power. Thus, caution is needed to interpret some of the nonsignificant interactions, and these results also highlight the need for future research to test these interaction patterns in a larger sample for replication.

**Discussion**

Although work is a key context for youth development, it is understudied among racial/ethnic minority youth, including Latino youth (aged 16 to 24) for whom employment is prevalent in the U.S. (U.S. Bureau of Labor Statistics, 2018). Grounded in the ecological theory highlighting the work-family mesosystem for youth development (Bronfenbrenner, 1986), the established literature on work-family interface among adults (Perry-Jenkins & Wadsworth, 2017), and the cultural ecological perspective emphasizing environmental influences on minority youth (García Coll et al., 1996), this study examined implications of Mexican-origin youth’s work experiences for their family relationships in later adolescence and young adulthood. I adopted an ethnic homogeneous approach to use longitudinal data from Mexican-origin families for this examination. In particular, based on the cultural emphasis on fathers’ sole or primary breadwinning role and paternal authority in the family, this study tested whether and how youth work hours and workplace discrimination (Time 1) were linked to father-youth relationship quality reported by youth and fathers two years later (Time 2). These links were expected to be curvilinear given the co-existence of positive and negative sides of youth work effects (Mortimer
& Staff, 2004; Zimmer-Gembeck & Mortimer, 2006), including those pertaining to Mexican-origin families (Arciniega et al., 2008; Fuligni, 1998; García Coll & Vázquez García, 1995) and previous findings suggesting a curvilinear pattern of these effects (Hwang & Domina, 2017; Shanahan et al., 1996b). I also tested youth gender and mothers’ employment status as moderators for these links, built on theory and prior literature highlighting the role of these factors in work-family interface and father-youth dynamics (Bronfenbrenner, 1986; Formoso et al., 2007; Minuchin, 1985; Roehling et al., 2005; Zimmer-Gembeck & Mortimer, 2006).

Controlling for Time 1 father-youth relationship quality reports, results of the multivariate multilevel models revealed an inverted U-shape curvilinear association between workplace discrimination and father-youth relationship quality (reported by fathers), and linear effects of work hours and workplace discrimination on youth reports of relationship quality qualified by youth gender and mothers’ employment status. This study contributes to a better understanding of the work-family mesosystem in youth development, especially among Latino youth whose work experiences have been rarely examined. Findings also extended the work-family interface literature, which has been almost exclusively focused on adults, by revealing potential effects of youth work experiences to their family relationships and the complexities in these effects involving curvilinear and interaction patterns.

**Links Between Youth’s Work Experiences and Relationships with Fathers**

By examining links between youth work experiences and father-youth relationship quality, this study contributes to a better understanding about implications of Latino youth’s work, a salient yet understudied context (Hwang & Domina, 2017) for this racial/ethnic group, and more broadly, youth work effects on their family relationship, another area that has received limited attention in recent decades. In addition, this study extended previous research on youth work
effects, which has focused almost exclusively on work hours in face of work’s multi-
dimensionality (Zimmer-Gembeck & Mortimer, 2006), by examining racial/ethnic
discrimination at work, a pervasive phenomenon that is unique to racial/ethnic minority groups.
Using data from two time points across two years, I was able to control for the concurrent
linkages between youth work experiences and father-youth relationship qualities and test the
longitudinal implications of youth work hours and workplace discrimination, both in the form of
main effects with potential curvilinear patterns and interactions with potential moderators—
youth gender and mothers’ employment status. And, the multivariate multilevel models allowed
examination of both youth’s and fathers’ reports of their relationship quality as dependent
variables in a same model while accounting for dyadic interdependencies in residuals, enabling
study of two relationship perspectives and adding statistical power for the analysis.

Results from the main effect models estimating both linear and quadratic terms of work
hours and workplace discrimination at Time 1 in relation to youth’s and fathers’ reports of
relationship quality at Time 2, controlling relationship quality reports at Time 1, revealed the
hypothesized inverted-U shape, curvilinear link between workplace discrimination and fathers’
reports of relationship quality. In this link, the significant positive slope at the very low levels of
discrimination indicated that mild encounters with this negative work experience can potentially
benefit Latino youth’s relationships with fathers. This positive effect is consistent with previous
findings that small to moderate negative work experiences may promote youth’s coping abilities
(Mortimer & Staff, 2004), and that discrimination experiences may induce family support (Noh
& Kaspar, 2003), especially from fathers who assume the protector role according to their
cultural values (Arciniega et al., 2008). However, at a higher level (from slightly below the
median to the highest scores), the link between discrimination and father-youth relationship
quality was negative. Importantly, the turning point of this curvilinear link (corresponding to the highest predicted level of relationship quality) emerged at a low level of workplace discrimination. Thus, this pattern is consistent with previous findings from research focused on linear effects of youth discrimination in documenting predominantly negative effects of discrimination on youth’s well-being (Okamoto et al., 2009; Schmitt et al., 2014), but extended prior research by showing the point at which a negative effect may begin to manifest. Future research should examine the mechanisms underlying this curvilinear link by testing mediators such as youth’s and families’ coping strategies in response to youth’s discrimination experiences at work as well as youth’s adjustment outcomes.

Beyond the curvilinear link found between workplace discrimination and fathers’ reports of relationship quality, three other hypothesized quadratic effects (work hours in relation to both youth’s and fathers’ relationship reports and workplace discrimination in relation to youth’s reports), were nonsignificant. Although previous research has suggested potential curvilinear effects of work hours including on academic achievement, behavioral adjustment and positive communication with fathers in predominantly White samples (Mortimer et al., 1996; Shanahan et al., 1996b) and on educational attainment among Latino youth (Hwang & Domina, 2017), such a pattern failed to emerge in this study. These different results may be due to differences in the ranges for low, moderate and high work hours across samples, a lack of statistical power to detect such effects, or that this pattern does not hold. Rather, the effect of work hours was qualified by youth gender and mothers’ employment status.

**Moderating Effects of Youth Gender and Maternal Employment**

This study also advances understanding of factors that may condition the effects of youth work experiences on father-youth relationship quality. From an ecological perspective, personal
characteristics such as youth gender, interact with contextual factors, including the work-family mesosystem (Bronfenbrenner, 1986; Zimmer-Gembeck & Mortimer, 2006). Relatedly, a family systems perspective highlights interdependencies among family subsystems (Minuchin, 1985), suggesting that mothers’ employment experiences may play a role in youth’s work-family interface and father-youth dynamics. Findings from this study documented the moderation effects of both of these factors on the linear effects of youth work experiences on father-youth relationship quality. In relation to youth’s relationship quality reports, youth gender moderated the linear effect of both work hours and workplace discrimination, and mother employment moderated the linear effect of workplace discrimination. In contrast, findings revealed no moderation effects on the links between youth work experiences and fathers’ reports of relationship quality.

Findings from tests of youth gender moderation indicated that the effects of work hours and discrimination were more positive for daughters than sons. In the case of work hours, the linear effects on father-youth relationships were positive among daughters but negative among sons. This finding was inconsistent with the study hypothesis, which was based on literature regarding the gender-stereotypical roles in Latino families that emphasize the male role as provider and breadwinner (Pinto & Coltrane, 2009). This pattern is also inconsistent with the previous finding that in Latino families, women experience more work-family conflict than men (Roehling et al., 2005; Wheeler et al., 2015). Instead, this pattern is consistent with prior findings on gender differences in work-family spillover effects among predominantly White adults, showing that work-family conflict is stronger among men whereas work-family enrichment is stronger among women (Grzywacz & Butler, 2005). It may be that this latter pattern applies to Latino youth: Whereas sons’ work stressors spillover to family such as because work is a core component of
their identities, daughters’ work is empowering including for their family roles—as well as potentially less threatening to their fathers. Also relevant to the differing meanings and motivations for work for Mexican-origin men and women, whereas sons may be more likely to work longer hours in response to family needs and obligations, daughters may work longer hours more out of self-determination rather than pressure from the family. Thus, possibly different motivations may make long hours at work a stressor for young men that can have negative spillover to their family relationships, but more of a personal growth opportunity for young women that can provide psychological resources and benefit their relationships. Accordingly, a future direction for a better understanding of this pattern is to use qualitative and quantitative data from Mexican-origin youth to investigate mechanisms, including their motivations, underlying their work experiences.

Congruent with the effects of work hours, findings also revealed a positive effect of workplace discrimination for daughters but not for sons. This pattern was consistent with the hypothesis about youth gender moderation of workplace discrimination effects and previous findings indicating stronger resilience mechanisms linking contextual cultural-related stressors to family relationships for young women than for young men in Latino families (Crockett et al., 2007). As noted, this gender moderation pattern of workplace discrimination echoed the pattern found for work hours in suggesting a stronger work-family enrichment process for daughters than for sons.

Turning to the moderation effect of mothers’ employment status, findings revealed that the effect of youth work hours on father-youth relationship quality was negative when mothers were employed, but positive when mothers were not employed—in contrast to the study hypothesis. This pattern was also inconsistent with the previous findings that maternal employment buffers
the negative effects of family stressors on father-youth relationship quality in Mexican-origin families (Formoso et al., 2007). Given that mothers play a key role in regulating father-youth dynamics (Puhlman & Pasley, 2013), it is possible that mothers’ time at work may make them less available for this regulating role and thus, the relationship between fathers and youth more vulnerable to stressors, including youth’s long work hours. In addition, though maternal non-employment may be linked to more stereotypical gender roles in Latino families (Pinto & Coltrane, 2009) and exacerbate the threats to fathers’ authority from youth’s long work hours, mothers’ availability at home may mitigate tensions between fathers and youth, leaving the positive aspects of youth work effects more salient in leading to better father-youth relationships.

An important next step is to directly measure the regulating role of mothers in father-youth dynamics in Mexican-origin families as a function of maternal employment. More generally, it also will be important to collect data on parents’ and youth’s perceptions of the meanings and motivations as well as the benefits and stressors of sons’ and daughters’ employment.

Notably, all the moderation patterns found in this study involved youth’s —but not fathers’— reports of father-youth relationship quality. Although the null results could be due to limited power as shown by the post-hoc power analysis results, these results also may mean that youth gender and maternal employment status may play less salient roles in fathers’ perceptions of relationship quality. Instead, as suggested, fathers’ perceptions were directly subject to youth’s experiences of workplace discrimination, a salient stressor outside the family that may trigger fathers’ protector role.

Limitations and Future Directions

In the face of its strengths, limitations of this study guide further directions for research. First, to increase the homogeneity of the sample and rule out some third variable explanations,
the sample was exclusively two-parent, Mexican-origin families with at least two offspring youth in adolescence and young adulthood, and fathers were employed at least part time. This limits the generalizability of study findings to other families, such as single-father families or families in which fathers were unemployed and where youth may feel more obligated to contribute financially to the family. Future research needs to examine youth work effects on father-youth relationship quality on more diverse samples of families. Second, limits to statistical power may have concealed some effects, especially interaction effects that may otherwise emerge as significant. Thus, a future direction is to test the study hypotheses in larger samples. Third, this study used a longitudinal design spanning across two years to examine the links between youth work experiences and father-youth relationship quality on a relatively large time scale. Work experiences and family dynamics, however, may vary on a daily basis and thus daily diary studies on these links would be a promising future direction towards a better understanding of the work-family mesosystem for youth’s family experiences and their adjustment, more generally. Finally, despite the longitudinal design, the study design was correlational, and thus intervention studies targeted at promoting youth work experiences are needed to examine causal links between youth work and their family relationships.

In sum, this study used an ethnic homogeneous design and longitudinal data to test links between youth work experiences and father-youth relationship quality in Mexican-origin families, and examined the role of youth gender and mothers’ employment status in moderating these links. By examining two dimensions of youth work—work hours and workplace discrimination—the study findings illuminate youth work effects on their family relationships, an understudied area, especially among Latino youth and with regard to relationships with fathers. The findings also contribute to the theory by providing empirical evidence about the youth work-
family mesosystem as highlighted in the ecological model, and expand on the work-family literature, which has focused on adults’, especially parents’ and spouses’ work, by targeting the workplaces of youth as a context with implications for the family system. Both the curvilinear association and the linear links that interacted with youth gender and mother employment illuminate complexities in youth work effects on their family relationships and suggest future directions to test underlying mechanisms. Finally, the findings direct attention to intervention programs targeted at work experiences among Latino youth for whom employment is prevalent.
References


different track from the very beginning. In H. E. Fitzgerald, B. M. Lerner, & B. Zuckerman (Eds.), *Children of poverty: Research, health, and policy issues* (pp. 57–83). New York: Garland.


U.S. Census Bureau (2000). Projections of the total resident population by 5-year age groups, race, and Hispanic origin with special age categories. *Middle series, 2001 to 2005.*


Table 3.1.
An Illustration of Data Stacked Within Each Family for Analysis With Multivariate Multilevel Models

<table>
<thead>
<tr>
<th>Family</th>
<th>Birth Order</th>
<th>Youth Work Experiences</th>
<th>Father Work Experiences</th>
<th>Relationship Reporter</th>
<th>Outcome Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>O</td>
<td>Work_O</td>
<td>Work_F</td>
<td>O</td>
<td>O-F Relationship Quality</td>
</tr>
<tr>
<td>1</td>
<td>O</td>
<td>Work_O</td>
<td>Work_F</td>
<td>F</td>
<td>F-O Relationship Quality</td>
</tr>
<tr>
<td>1</td>
<td>Y</td>
<td>Work_Y</td>
<td>Work_F</td>
<td>Y</td>
<td>Y-F Relationship Quality</td>
</tr>
<tr>
<td>1</td>
<td>Y</td>
<td>Work_Y</td>
<td>Work_F</td>
<td>F</td>
<td>F-Y Relationship Quality</td>
</tr>
</tbody>
</table>

*Note. O = Older sibling; Y = Younger sibling; F = Father.*
Table 3.2.
Means (M), Standard Deviations (SD), and Correlations Between Study Variables Stacked Across Siblings (N = 187)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Youth work hours (T1)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Youth workplace discrimination (T1)</td>
<td>.23**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Father work hours (T1)</td>
<td>-.09</td>
<td>-.12</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Father workplace discrimination (T1)</td>
<td>.01</td>
<td>.14†</td>
<td>-.02</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Relationship quality (T1): Youth reports</td>
<td>-.04</td>
<td>-.15*</td>
<td>.04</td>
<td>-.09</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Relationship quality (T1): Father reports</td>
<td>.01</td>
<td>-.03</td>
<td>-.18*</td>
<td>-.08</td>
<td>.41***</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Relationship quality (T2): Youth reports</td>
<td>-.06</td>
<td>-.09</td>
<td>-.06</td>
<td>-.07</td>
<td>.76***</td>
<td>.55***</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Relationship quality (T2): Father reports</td>
<td>-.00</td>
<td>-.09</td>
<td>-.25**</td>
<td>-.06</td>
<td>.44***</td>
<td>.76***</td>
<td>.51***</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Youth gender</td>
<td>.12</td>
<td>.13†</td>
<td>-.05</td>
<td>.18*</td>
<td>-.06</td>
<td>-.04</td>
<td>-.13</td>
<td>.02</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>10. Mother employment (T1)</td>
<td>-.12†</td>
<td>.06</td>
<td>.05</td>
<td>.16*</td>
<td>-.06</td>
<td>-.03</td>
<td>-.07</td>
<td>-.03</td>
<td>.03</td>
<td>-</td>
</tr>
<tr>
<td>11. Youth age (T1)</td>
<td>.35***</td>
<td>.01</td>
<td>-.04</td>
<td>-.07</td>
<td>.07</td>
<td>.00</td>
<td>.03</td>
<td>-.05</td>
<td>.11</td>
<td>-.00</td>
</tr>
<tr>
<td>12. Youth living at home (T1)</td>
<td>-.22**</td>
<td>-.11</td>
<td>.00</td>
<td>.01</td>
<td>-.01</td>
<td>.05</td>
<td>.10</td>
<td>-.00</td>
<td>.30***</td>
<td>-.41***</td>
</tr>
<tr>
<td>13. Youth attending school (T1)</td>
<td>-.45***</td>
<td>-.13†</td>
<td>.06</td>
<td>.05</td>
<td>.03</td>
<td>-.06</td>
<td>.06</td>
<td>.02</td>
<td>-.11</td>
<td>.05</td>
</tr>
<tr>
<td>14. Family income (T1)</td>
<td>-.11</td>
<td>-.22**</td>
<td>.16*</td>
<td>-.19*</td>
<td>.17*</td>
<td>-.14†</td>
<td>.08</td>
<td>-.12</td>
<td>-.07</td>
<td>.10</td>
</tr>
<tr>
<td>15. Mother education (T1)</td>
<td>-.15*</td>
<td>-.14†</td>
<td>.02</td>
<td>.01</td>
<td>.07</td>
<td>-.05</td>
<td>.05</td>
<td>-.04</td>
<td>.01</td>
<td>.13†</td>
</tr>
<tr>
<td>16. Father education (T1)</td>
<td>-.20**</td>
<td>-.08</td>
<td>.04</td>
<td>-.05</td>
<td>.08</td>
<td>-.11</td>
<td>.03</td>
<td>-.12</td>
<td>.07</td>
<td>.15*</td>
</tr>
<tr>
<td>17. Youth nativity</td>
<td>-.08</td>
<td>-.13†</td>
<td>.05</td>
<td>.03</td>
<td>.05</td>
<td>-.10</td>
<td>.02</td>
<td>-.10</td>
<td>-.06</td>
<td>.03</td>
</tr>
<tr>
<td>18. Mother nativity</td>
<td>-.15*</td>
<td>-.10</td>
<td>.06</td>
<td>-.02</td>
<td>-.04</td>
<td>-.14†</td>
<td>-.07</td>
<td>-.23**</td>
<td>-.07</td>
<td>.05</td>
</tr>
<tr>
<td>19. Father nativity</td>
<td>-.12†</td>
<td>-.09</td>
<td>.08</td>
<td>-.01</td>
<td>.05</td>
<td>-.16*</td>
<td>-.04</td>
<td>-.18*</td>
<td>-.01</td>
<td>.05</td>
</tr>
</tbody>
</table>

Note. Descriptives based on raw data (prior to centering and recoding). T1 = Time 1, T2 = Time 2. Youth gender: 0 = female, 1 = male. Mother employment: 0 = not employed, 1 = employed. Youth living at home: 0 = youth did not, 1 = youth did live at home. Youth attending school: 0 = not attending school, 1 = attending school. Nativity: 0 = born outside U.S., 1 = born in U.S. †p<.10, *p<.05, **p<.01, ***p<.001.
Table 3.3.

Unstandardized Coefficients (γ) and Standard Errors (SE) From Two-Intercept Multivariate Multilevel Models Testing Curvilinear Associations Between Youth’s Work Experiences (Work Hours and Workplace Discrimination; Time 1) and Fathers’ and Youth’s Reports of Relationship Quality (RQ_{Father} and RQ_{Youth}; Time 2), and Their Moderation by Youth Gender and Mothers’ Employment Status

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>γ</td>
<td>(SE)</td>
</tr>
<tr>
<td>Intercepts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ_{Father}</td>
<td>4.050*** (.066)</td>
<td>4.006*** (.111)</td>
</tr>
<tr>
<td>RQ_{Youth}</td>
<td>3.798*** (.071)</td>
<td>3.807*** (.119)</td>
</tr>
<tr>
<td>Covariates and moderators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother nativity→RQ_{Father}</td>
<td>-1.95* (.091)</td>
<td>-2.00 (.093)</td>
</tr>
<tr>
<td>Time 1 RQ→RQ_{Father}</td>
<td>.613*** (.057)</td>
<td>.602*** (.058)</td>
</tr>
<tr>
<td>Time 1 RQ→RQ_{Youth}</td>
<td>.633*** (.051)</td>
<td>.599*** (.052)</td>
</tr>
<tr>
<td>Gender→RQ_{Father}</td>
<td>- -</td>
<td>.041 (.072)</td>
</tr>
<tr>
<td>Gender→RQ_{Youth}</td>
<td>- -</td>
<td>-1.87* (.091)</td>
</tr>
<tr>
<td>Mother employment→RQ_{Father}</td>
<td>- -</td>
<td>.061 (.111)</td>
</tr>
<tr>
<td>Mother employment→RQ_{Youth}</td>
<td>- -</td>
<td>.055 (.125)</td>
</tr>
<tr>
<td>Youth work experiences main effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours→RQ_{Father}</td>
<td>-0.001 (.002)</td>
<td>-0.001 (.002)</td>
</tr>
<tr>
<td>Hours × Hours→RQ_{Father}</td>
<td>.000 (.000)</td>
<td>- -</td>
</tr>
<tr>
<td>Hours→RQ_{Youth}</td>
<td>-0.001 (.003)</td>
<td>0.018* (.007)</td>
</tr>
<tr>
<td>Hours × Hours→RQ_{Youth}</td>
<td>-0.002 (.001)</td>
<td>- -</td>
</tr>
<tr>
<td>Discrimination→RQ_{Father}</td>
<td>-0.006 (.067)</td>
<td>-0.014 (.067)</td>
</tr>
<tr>
<td>Discrimination × Discrimination→RQ_{Father}</td>
<td>-2.13** (.078)</td>
<td>-2.10** (.079)</td>
</tr>
<tr>
<td>Discrimination→RQ_{Youth}</td>
<td>0.013 (.092)</td>
<td>0.196† (.105)</td>
</tr>
<tr>
<td>Discrimination × Discrimination→RQ_{Youth}</td>
<td>0.009 (.112)</td>
<td>- -</td>
</tr>
<tr>
<td>Moderations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours × Gender→RQ_{Youth}</td>
<td>-0.015* (.006)</td>
<td>-0.015* (.007)</td>
</tr>
<tr>
<td>Hours × Mother employment→RQ_{Youth}</td>
<td>-0.015* (.007)</td>
<td>-0.337* (.164)</td>
</tr>
<tr>
<td>Discrimination × Gender→RQ_{Youth}</td>
<td>- -</td>
<td></td>
</tr>
<tr>
<td>Random effects—Intercepts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (RQ_{Father})</td>
<td>.098** (.035)</td>
<td>.102** (.035)</td>
</tr>
<tr>
<td>Variance (RQ_{Youth})</td>
<td>.090* (.046)</td>
<td>.082* (.041)</td>
</tr>
<tr>
<td>Covariance (RQ_{Father},RQ_{Youth})</td>
<td>.048 (.030)</td>
<td>.046 (.028)</td>
</tr>
<tr>
<td>Random effects—Residuals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (RQ_{Father})</td>
<td>.111*** (.025)</td>
<td>.109*** (.024)</td>
</tr>
<tr>
<td>Variance (RQ_{Youth})</td>
<td>.249*** (.047)</td>
<td>.215*** (.042)</td>
</tr>
<tr>
<td>Covariance (RQ_{Father},RQ_{Youth})</td>
<td>-0.002 (.023)</td>
<td>-0.011 (.022)</td>
</tr>
</tbody>
</table>

Note. Nonsignificant covariates and interaction terms were omitted from final models.

†p<.10, *p<.05, **p<.01, ***p<.001.
Figure 3.1. Curvilinear association between youth workplace discrimination and father-youth relationship quality reported by father; the turning point is indicated by the dotted line and the 95% CI, by dashed lines.
CHAPTER 4

ADOLESCENT FAMILY EXPERIENCES PREDICT YOUNG ADULT EDUCATIONAL ATTAINMENT:

A DATA-BASED CROSS-STUDY SYNTHESIS WITH MACHINE LEARNING

Introduction

Educational attainment in young adulthood is critical to individuals’ well-being throughout adulthood (IOM & NRC, 2015). From a developmental perspective, it is important to determine whether and how experiences in childhood and adolescence are related to educational attainment (Pettit, Davis-Kean, & Magnuson, 2009). Built on a range of theoretical frameworks, including the life course perspective (Elder, 1998), social capital theory (Coleman, 1988), the ecological model (Bronfenbrenner, 1986), and the expectancy-value model of achievement (Eccles, 2011), hypothesis-testing studies have identified adolescent family experiences that predict young adults’ educational attainment (Benner, Boyle, & Sadler, 2016; Gordon & Cui, 2012; Parcel, Dufur, & Zito, 2010; Monserud & Elder, 2011; Sun, McHale, & Updegraff, 2017). These studies have identified correlates of achievement, including multiple components of family structure, family socioeconomic characteristics, family relationships and parenting practices, and parent characteristics, though to date, individual studies have been limited to examination of a relatively small subset of potentially important family experiences.

Recognition of the multi-dimensional, multi-component nature of family experiences is evident in the design of large-scale, longitudinal projects. For example, the National Longitudinal Study of Adolescent Health (Add Health; Harris & Udry, 1994-2008), a longitudinal study of a U.S. nationally representative sample that spans across adolescence and into young adulthood, includes numerous measures of adolescent family experiences, and a body
of research has linked these to young adults’ educational attainment. Indeed, a literature search revealed 101 studies (published between August 2016 and March 2019) that used Add Health data to examine adolescent family experiences (Wave I; 1994-1995, Grades 7 – 12) as predictors of educational attainment in young adulthood (Wave III (2001-2002); Wave IV (2008). The goal of the current study was to synthesize findings from these previous studies by analyzing Add Health data using an innovative approach with machine learning (ML) to address three research questions. (1) How accurately does this broad range of adolescent family factors predict young adult educational attainment? (2) When examined concurrently, which family experience factors are the best predictors of young adult educational attainment? And (3) What complex patterns, including nonlinearities and interactions involving this range of family factors, merit further examination?

Cross-Study Synthesis With A Machine Learning-Based Approach

For this cross-study synthesis, I identified prior studies using Add Health data to examine adolescent family predictors of young adult educational attainment ($N = 101$ studies) by: (1) searching for journal articles within the Add Health publications database (https://www.cpc.unc.edu/projects/addhealth/publications) with keywords including “education(al)” “academic” “college” “postsecondary” “high school” “attainment” “achievement” “attendance” “enrollment” “completion” “graduation”, and (2) selecting studies that tested models with at least one family experience variable measured at Wave I, including independent variables or covariates, and young adult educational attainment outcomes at Wave III or IV as dependent variables. Table 4.1 summarizes the educational attainment outcomes, the Wave I family experience independent variables/covariates, and the statistical methods used in these studies.
To analytically synthesize these studies, I applied a ML-based approach because ML allows for simultaneous consideration of how a large number of predictor variables are related to an outcome. Using this approach I included nearly all of the adolescent family experience variables that emerged in the literature search to predict college enrollment and graduation in young adulthood. Compared to the regression based, hypothesis-testing approach used in prior analyses of Add Health data, ML is a data-driven, exploratory approach focused on building models with existing data on predictors and outcomes in an effort to maximize predictive accuracy (McArdle, 2013; Yarkoni & Westfall, 2017). In contrast to an hypothesis-driven approach, a basic protocol within ML, termed supervised learning, involves first training the model with a pre-determined algorithm and known data on predictors and their outcome, then predicting outcome values with data on the predictors from new cases, and finally testing the model prediction performance by comparing the predicted values to the true outcome values in these cases. With its capacity to simultaneously include and test a large number of predictors, ML is especially suitable for the current cross-study synthesis of a range of family experience variables: In traditional statistical approaches, due to potential problems of overfitting and inflation of Type I error from multiple testing, a limited number of predictors can be examined in any given model or study (Burke, Ammerman, & Jacobucci, 2019; Whelan et al., 2014; Yarkoni & Westfall, 2017). ML also is well-suited to address the three major study questions raised above due to its capacity to maximize predictive power, its functions of feature importance estimation and feature selection, which can identify the strongest predictors of an outcome, and using some algorithms (e.g., random forests), its capacity to freely estimate nonlinear and interactive effects among predictors. These unique features of ML are detailed below.

Predicting Young Adult Educational Attainment From Adolescent Family Experiences
How accurately can a comprehensive set of family experiences in adolescence predict educational attainment in young adulthood? Making predictions of future outcomes is a key goal of developmental science (Baltes, Reese, & Nesselroade, 1977; Yarkoni & Westfall, 2017). Accordingly, the studies listed in Table 4.1 have collectively estimated numerous hypothesis-testing models (mostly regression-based) that reveal significant associations between family experiences and educational outcomes. Results of each study target a set of family experiences that predict educational attainment. Families, however, are complex systems, embedded in larger contexts, that include multiple, interacting subsystems which together influence family members’ outcomes (Minuchin, 1985). Thus, separate studies, with each focused on a subset of potential family influences, do not provide a comprehensive picture of the wide range of family systems components and processes that operate together to predict young adult educational attainment.

As described by Yarkoni and Westfall (2017), extant studies test “simple models that appear theoretically elegant but have very limited capacity to predict actual human behavior” (p.1101). In addition, with the primary goal of testing statistical significance of associations between predictors and outcomes, these studies have identified significant predictors in separate models, but the predictive power of the overall models have not been a focus. In particular, reporting prediction accuracy is a rare practice in hypothesis-testing studies, including among the 101 studies reviewed here. Thus, the current study aimed to contribute to the literature on the role of family experiences in educational attainment by testing the prediction accuracy of a comprehensive set of these experiences (Yarkoni & Westfall, 2017).

The ML-based approach provides new opportunities for predicting human behavior and development due to its capacity to train comprehensive prediction models that allow for complex patterns involving a large number of predictors and test those models using data independent of
data of the training set (Rosenberg, Casey, & Holmes, 2018; Yarkoni & Westfall, 2017). For example, recent research has used ML to train and test predictive models for suicidal thoughts and behaviors (Burke et al., 2019), child maltreatment (Chouldechova, Putnam-Hornstein, Benavides-Prado, Fialko, & Vaithianathan, 2018), adolescent alcohol misuse (Whelan et al., 2014), and failing or repeating grades in high school (Lakkaraju et al., 2015). Although exploratory, these studies advance understanding about developmental change and, because of their high predictive power revealed in the testing procedure, can identify at-risk individuals as targets for preventive interventions. Following these examples, the current study was designed to build comprehensive predictive models of young adult educational attainment from adolescent family experiences by analytically synthesizing findings from prior studies with ML. Based on the longitudinal, national representative and high-dimensional Add Health data, training and testing the predictive models with independent subsets of the data can reveal how accurately these family experiences, when incorporated simultaneously in the same model, predict educational attainment. Also, this novel integrative modeling approach can provide new insights about the role of families as multi-dimensional and multi-component systems in later young adult educational attainment.

**Identifying Key Family Experiences for Educational Attainment**

*Which family experiences are the key predictors of educational attainment?* Theoretical frameworks emphasize the role of family experiences in later educational attainment. First, the life course perspective (Elder, 1998) highlights “linked lives” between generations and thus focuses on parents’ experiences, attainments, expectations and behaviors in explaining youth development and achievement. Coleman’s (1988) capital theory illuminates three dimensions of family resources (i.e., capital) that are important to offspring’s achievement, including
financial/material capital (e.g., family income), human capital (e.g., parents’ educational and occupational attainments), and social capital (e.g., family structure such as mother/father presence and family emotional support). Family socialization perspectives, such as Eccles’ (2011) model of achievement choices, target parents’ expectations and practices, in addition to family resources, in youth’s achievement outcomes. From among these theoretically important factors, in translating research findings for practice, a key step is to identify modifiable, cost-effective factors of practical (not just statistical) significance to serve as the targets for prevention and intervention (Holder, 2010). This means that family and parenting processes may be of more practical significance than family structural and status characteristics.

Drawing from these theories, studies using Add Health data have collectively examined relations between a wide variety of adolescent family experiences and young adult educational attainment. Starting with family structure, presence of two biological married parents, smaller household size and fewer siblings, and earlier birth order are related to higher achievement (e.g., Feliciano & Lanuza, 2017; Fletcher & Lehrer, 2009; Mangino, 2014; Monserud & Elder, 2011). Implications of family socioeconomic characteristics have been widely studied, with results indicating that socioeconomic resources, including higher family income and parents’ education attainment and occupational prestige are related to higher educational attainment (e.g., Faas, Benson, & Kaestle, 2013; Fasang, Mangino, & Brückner, 2014), whereas family receipt of public assistance and welfare and parents’ economic hardship are related to lower attainment (e.g., Gillette, & Gudmunson, 2014; Humberstone, 2018). Potentially modifiable family processes that have been the focus of intervention research, also have been linked to young adult educational attainment, including parents’ educational expectations (e.g., Gordon & Cui, 2012) and positive parenting practices, including school involvement (e.g., Ashtiani & Feliciano, 2018; Gordon &
Cui, 2012), shared activities, particularly family meals (e.g., Boardman, Alexander, Miech, MacMillan, & Shanahan, 2012; Monserud & Elder, 2011), less control/ more autonomy granting (Monserud & Elder, 2011) but more supervision/ monitoring (e.g., Ryabov; 2013) and intergenerational closure (i.e., knowing adolescents’ friends’ parents; Glanville, Sikkink, & Hernández, 2008) and better family relationship quality, particularly parent-adolescent relationship quality (e.g., Erickson, McDonald, & Elder, 2009). Some studies, however, have documented negative effects of high levels of parent-adolescent relationship quality and family social support on educational attainment (Ryabov, 2016; Turley, Desmond, & Bruch, 2010).

Finally, parent and family sociocultural characteristics, such as parents’ nativity (i.e., born outside of U.S.; Benner & Wang, 2014), religiosity (Stokes, 2008), and English language spoken at home (Mears & Siennick, 2016) have been positively linked to attainment, whereas parent health conditions and health risk behaviors such as smoking (Boardman et al., 2012) have been negatively related to young adults’ educational attainment. Notably, adding to the number of variables that has been examined, many of these factors were measured and tested separately for mothers and fathers.

Building on prior studies’ efforts to document significant effects of each of these many family variables, I used the ML approach to establish the relative importance of each. That is, with the prediction models trained, ML can reveal the relative contributions of the numerous predictors, termed feature importance (Brick, Koffer, Gerstorf, & Ram, 2017), and identify the most important predictors of a given outcome from among numerous predictors. For example, one study used over 40 predictors across domains including genetics, demographics, personal history, cognition, personality, and brain activity, to predict adolescents’ current and future binge drinking status; findings revealed that the personal history domain (e.g., family history of drug
use, romantic history) most strongly predicted both outcomes (Whelan et al., 2014). In addition to feature importance, ML also enables recursive feature elimination (RFE). In this procedure, the least important predictor(s) is eliminated from the predictive model in successive runs, each time followed by estimation of the model’s overall performance (i.e., prediction accuracy); elimination can be ended at the point where prediction performance drops substantially. With RFE, researchers can isolate a small set of predictors that together have predictive power that is equivalent to that of the original, larger set of predictors. As an example, Brick and colleagues (2017) used RFE to pare down 598 predictors of life satisfaction and subjective health status to 20 predictors that had equivalent, strong predictive power.

Following these examples, this study used feature importance estimation in ML to rank the contributions of the many adolescent family experiences that have been examined in prior studies using the Add Health dataset to predict young adult educational attainment, and used RFE to identify the subset of adolescent family experience factors that are highly predictive of individual differences in young adult educational attainment. Identifying the most important family experience factors from the very large array of possibilities can guide future research and practice to selectively measure, analyze, and devote intervention resources in the most cost-effective targets (Brick et al., 2017).

**Discovering Nonlinear and Interactive Patterns in Family Experience Factors That Predict Educational Attainment**

*What complex patterns, including nonlinearities and interactions, merit further examination?* Beyond the main linear effects most often examined in hypothesis-testing studies, given the complexities of family systems dynamics, other patterns may exist. According to family systems theory, family processes are circular and interdependent (Minuchin, 1985), a
tenet that is echoed in the ecological model’s focus on nonlinearities and interactive effects of person, process and context in development (Bronfenbrenner & Morris, 2006). Grounded in such ideas, some research has examined nonlinear and interactive patterns in family experience factors and their links to young adult education attainment. Using Add Health data to examine nonlinearity, for example, Mangino (2014) investigated the effect of family income in the top 1% of the population. Other studies using Add Health data have tested two-way interactions such as between parental educational expectations and both parent education and family income (Mahatmya & Smith, 2017). Beyond Add Health studies, some research has documented interaction (i.e., moderation) effects suggestive of both amplifying and compensatory risk and protective effects involving family socioeconomic characteristics, parent expectations and practices, and family relationship quality in relation to educational achievement (Benner et al., 2016; Sun et al., 2017). Such findings provided the basis for the current study’s examination of nonlinear and interactive effects of adolescent family experiences in young adults’ educational attainment. Traditional statistical methods, however, have faced challenges in exhaustively exploring and testing all possible nonlinear and interaction patterns among the large number of family experience factors identified in prior research toward identifying those that are important in predicting outcomes (McArdle, 2013). Thus, these methods have been considered as “often ill-suited as operational models for developmental investigations in the discovery mode” (Bronfenbrenner & Morris, 2006, p. 802).

Accordingly, with the ML approach implemented in this study, I was able to speed up the discovery of complex patterns of adolescent family experience effects on young adult educational attainment: Beyond including a large number of predictors within a same model, ML has the capacity to freely estimate nonlinear and interactive effects (including high-order
interactions beyond two- and three-way) among these predictors in building predictive models using algorithms such as random forests (Breiman, 2001). Moreover, comparing models built with ML algorithms that impose linearity and exclusivity, such as (regularized) logistic regression, to those built with algorithms that allow nonlinearities and interactions, can potentially reveal complex patterns for future examination. These complex patterns can be depicted using partial dependence plots, a follow-up visualization method for interpreting complex ML models with nonlinear and interaction effects (Friedman, 2001; Molnar, 2019).

Thus, in this study I compared the feature importance and RFE results of models that only accommodate linear effects to those of models that additionally accommodate non-linearity and higher-order interactions. Further, by depicting important predictors unique to the latter models with partial dependence plots, I aimed to discover possible nonlinearities and interactions among the family experience factors that suggested examples worth further examining in hypothesis-testing studies.

**The Current Study**

In sum, built on an expansive and growing body of knowledge about how various aspects family experiences influence educational attainment, this study took an important novel step to analytically synthesize prior studies and address three main research questions (as introduced above) using a data-driven synthesis approach with ML. In particular, this cross-study synthesis used Add Health data, a large-scale longitudinal study of a U.S. nationally representative sample of youth, to build comprehensive models with adolescents’ family experiences as predictors of their young adult educational attainment using an ML approach. Based on 101 published studies that used Add Health data (Table 4.1), I trained and tested ML models focusing on a range of family experiences measured in adolescence (Add Health Wave I), including family structure,
family socioeconomic characteristics, family relationships and parenting, and parent and other family characteristics—in total 55 variables as summarized in Table 4.2, to predict educational attainment in young adulthood (Wave IV). Educational attainment was indexed by two dichotomous variables, college enrollment and graduation with a Bachelor’s degree (termed college graduation hereafter), outcomes that are important education milestones with significant implications for adulthood employment, family formation, and health (IOM & NRC, 2015) and have been the primary focus of previous publications using Add Health data to study young adult education attainment. In particular, I applied two ML algorithms in this study—regularized logistic regression (an algorithm imposing linearity and exclusivity among predictors) and random forests (an algorithm allowing complexities such as interactions and nonlinearities among predictors; Breiman, 2001), studying the same set of family experience variables to predict young adult educational attainment. Both algorithms have been implemented in previous studies predicting individual development in domains such as substance use (Whelan et al., 2014), romantic relationship (Joel, Eastwick, & Finkel, 2017), and suicidal thoughts and behaviors (Burke et al., 2019).

With respect to the first research question, the predictive performance of these ML models, such as prediction accuracy, determined how well family experiences in adolescence, taken together, predicted young adult educational attainment. To address the second question, feature importance estimation and RFE results established the rank order and the relative contributions of the predictors and identified the most important adolescent family experience factors for predicting college enrollment and graduation. Finally, to address the third question, I compared the feature importance and RFE results between two algorithms—regularized logistic regression and random forests, in combination with follow-up visualization with partial dependence plots,
to illuminate nonlinearities and interactions among the family experience factors. More
generally, this study was designed to provide for family and developmental research with a
paradigm of how ML can be applied to large publicly available data resources (e.g. Add Health)
toward synthesizing and building a next layer of knowledge.

**Method**

**Data**

**Sample.** To conduct the cross-study synthesis, I used the public-use sample from Add
Health, about 31% of the entire restricted dataset. I chose the public-use data due to its wide
accessibility, which makes the current study and paradigm readily replicable and key exploratory
results, such as nonlinear and interactive patterns among predictors, readily available for future
hypothesis-testing studies. Further, using these data I was able to draw generalizable conclusions
by adjusting results using sample weights (Wave IV) provided by Add Health. Given that I used
family experience variables collected from adolescents and parents in home interviews at Wave I
to predict educational attainment at Wave IV, analyses focused on adolescents in the public
dataset who participated in both Wave I and Wave IV (N = 5,114). At Wave I (1994-1995),
adolescents were in grades 7-12 (age weighted $M = 15.94$, $SD = 1.77$; 50.53% female). Young
adults’ mean age (weighted) at Wave IV (2008) was 28.88 ($SD = 1.76$, range = 24.42-33.92).
Ethnic composition of the sample (weighted proportions) was 67.33% White, 16.49% African
American, and 11.09% Hispanic/Latino origin, 3.56% Native American, 2.98% Asian and 6.52%
other. I also used data collected from parents at Wave I ($N = 4,548$; age weighted $M = 41.51$, $SD$
$= 6.35$; 92.71% mothers).

**Variables.** A collection of family experience variables and two educational attainment
variables were identified for synthesis based on review of the publications listed in Table 4.1.
Family experiences (predictors). Among the family experience variables that were measured at Wave I and examined in at least one of the studies listed in Table 4.1, I selected variables as predictors for the ML models based on these inclusion criteria: (1) the most commonly used measures of composite scores for a family experience construct that has been indexed by different (but usually overlapping) survey items across studies (e.g., mother/father-adolescent relationship quality, family support); (2) separate measures for two parents, if available (instead of a generic parent score, e.g., mother/father education instead of parent education) to allow comparison of that mothers’ and fathers’ roles in young adult education attainment; (3) continuous variables rather than categorical variables when available and used in prior studies to maximize variability (e.g., mother/father education levels rather than dummy codes for higher versus lower education levels). As shown in Table 4.2, the final list of variables included measures of family structure, family socioeconomic characteristics, family relationships and parenting, and parent and other family characteristics; I also included adolescent demographic characteristics (biological sex, age, race/ethnicity, nativity) which have often been treated as covariates in prior studies. Although 55 variables were initially identified, two were excluded given that over 25% of cases had missing data (Brick et al., 2017), which left 53 variables, shown in Table 4.2, to serve as predictors in the ML models. Bivariate correlations between these variables were all lower than $r = .80$ and thus were all retained for model training (Brick et al., 2017).

Educational attainment (outcomes). Educational attainment was operationalized using two dichotomous variables, college enrollment and graduation, derived from young adults’ response to the question “What is the highest level of education that you have achieved to date?” at Wave IV (1 = 8th grade or less, 2 = some high school, 3 = high school graduate, 4 = some
vocational/technical training, 5 = completed vocational/technical training, 6 = some college, 7 = completed college (bachelor’s degree), 8 and above = post baccalaureate education). College enrollment was coded as 0 for young adults who scored from 1 to 4, and 1 for those who scored 5 and above to this question. College graduation (with bachelor’s degree) was coded as 0 for young adults who scored from 1 to 5, and 1 for those who scored 6 and above. Youth who reported ‘some college’ and were still attending college (N = 418) were eliminated from the analyses for college graduation.

Data Preparation

Prior to analysis with ML, I addressed missing data and sample weighting.

Missing data. For adolescents’ reports at Wave I, missing data originated from legitimate skips (e.g., adolescents without resident fathers skipped questions about their fathers’ parenting) or responses such as “don’t know” or “refused” to the questions. In addition, some parent data were missing due to non-participation. For missing data due to legitimate skips, I followed the dummy variable adjustment approach (Allison, 2001; Cohen & Cohen, 1985): First, a dummy variable was created for the skip pattern (e.g., “Father” variable where 0 = no resident father and 1 = has resident father); then, missing data due to legitimate skips were imputed with the variable means. With the legitimate skips imputed, before further imputation I deleted cases/participants with over 25% missingness, similarly to the approach to variable selection noted above. This step resulted in 4,598 cases for the college enrollment outcome variable (weighted proportion = 62.94% for value 1) and 4,180 cases for the outcome variable, college graduation (weighted proportion = 33.23% cases with value 1). The remaining data only had 1.88% missingness. I handled these missing data with augmented multiple multivariate imputation using chained equations (MICE) with 5 imputations and 500 iterations by the R ‘mice’ package (van Buuren &
Groothuis-Oudshoorn, 2011), for which all the family experience variables were included in the estimation equations. For parsimony in the following analysis, I pooled the 5 imputed datasets by taking the mode for categorical variables and mean for continuous variables.

**Sample weighting.** In following analyses, I accounted for sample weights at Wave IV in the Add Health dataset for youth who participated both Waves I and IV such that cases with higher weight values were weighted more in model training and testing.

**Data Analysis**

Using an ML approach, three steps of data analysis mapped onto the three research questions addressed in this study.

**Question 1: Do family experiences in adolescence predict college enrollment and graduation in young adulthood?** I used two ML algorithms, regularized logistic regression and random forests, to quantify the extent to which adolescents’ family experiences, when examined together, were predictive of college enrollment and graduation in young adulthood. Within each algorithm and for each outcome, I conducted model training and model testing with stratified nested cross-validation (Krstajic, Buturovic, Leahy, & Thomas, 2014), and evaluated model prediction performances as answers to this research question. The algorithms and procedure are detailed below.

**Algorithm 1: Regularized logistic regression.** This algorithm is based on logistic regression wherein the logit function of the binary outcome is modeled using a linear combination of the predictors. Without specification, all predictors are assumed to have linear and exclusive associations with the outcome—consistent with traditional statistical methods based on linear (logistic) regression. Within the ML approach that aims at maximizing the prediction accuracy of the outcome given existing data, coefficients in the logistic regression
model are estimated with the maximum likelihood function. In addition, to prevent overfitting, the estimation also simultaneously accounts for a penalty function—the L2 penalty function in this study—for regularization of the coefficients. Thus, given data for training, estimation of the regularized logistic regression is to minimize:

$$\lambda \frac{1}{2} w^T w + \sum_{i=0}^{n} \log(\exp(-y_i(X_i^T w + c)))$$,

where the first component is the penalty function with the parameter $\lambda$ determining the weight of the cost function and subject to tuning (i.e., in model training, selecting the best value of the parameter to optimize model performance), and the second component is a negative log likelihood function, with $w$ and $c$ indicating the vector of the coefficients and the intercept.

**Algorithm 2: Random forests.** Random forests are an ensemble method of decision tree models such as Classification and Regression Trees (CARTs; Breiman, 1984, 2001). A decision tree classifies the outcome with recursive partitioning based on the predictor values, and returns combinations of predictors with their thresholds for determining the level of the outcome, for example, “if family income > 30,000 and 3< family social support <4, then college enrollment = 1, otherwise college enrollment = 0”. Thus, decision trees allow for nonlinearities and interactions among the predictors. A salient drawback of a single decision tree is that it is prone to overfitting. Accordingly, random forests regularize overfitting by constructing predictions from a collection of decision trees (Breiman, 2001). Further regularization and optimization of random forests involves tuning the parameters (termed, hyperparameters in the ML literature), such as the number of trees, the maximum depth (i.e., number of partitions) of each tree, and the minimum sample size required for further partitioning. In this study, I focused on tuning the combination of: (1) the number of trees, a parameter that is important to determine for random forests (Couronné, Probst, & Boulesteix, 2018) and that impacts model convergence (Probst,
Wright, & Boulesteix, 2018); and (2) the maximum depth of each tree, a parameter controlling the complexity of the prediction mechanisms given that a higher-depth tree allows more space for nonlinearities and interactions (Brandmaier, von Oertzen, McArdle, & Lindenberger, 2012).

**Model training and testing with stratified nested cross-validation.** Given a dataset, an ML approach usually splits it into two independent subsets, the training set and the testing set (termed, ‘known data’ and ‘unknown data’), trains the models with the predictor and outcome data from the training set, and tests model performance on the testing set so that the output indicates how well the model replicates in unknown cases. To fully utilize the dataset and obtain a more stable estimation of the model prediction performances and feature importance indices, *k*-fold cross-validation is usually used in this training-testing procedure, a process in which the whole dataset is split into *k* folds and the model is trained *k* times, each time with the *k*-1 folds and then tested with the remaining one fold (Geisser, 1975). Moreover, within each model training run, cross-validation can be further conducted within the training set for model tuning in order to determine the optimal value(s) of the model parameter(s)—a nested cross-validation technique (Varma & Simon, 2006). As depicted in Figure 4.1, in this study I applied a nested five-fold cross-validation, with the inner loop for model tuning (i.e., for each parameter value, training with four folds and validation with the remaining one) and the outer loop for model training with the optimal parameter value(s) and testing. *Stratified* nested cross-validation was used so that the class proportions (which were unbalanced) were equivalent across folds (Krstajic et al., 2014). The tuning procedure was performed on the cost function parameter, $\lambda$, for regularized logistic regression models, and on the combination of number of trees and maximum depth, for random forests models. I used the classification accuracy index (i.e., the percentage of correct predictions) in model tuning to select the best parameter value(s) for which the averaged
accuracy scores across the five inner-loop validations was/were highest.

**Model evaluation.** The main analytical task here was to quantify the prediction relation between adolescents’ family experience and educational attainment in young adulthood. To do so, I indexed the model performance with two indices mostly often used in ML-based studies, classification accuracy and the area-under-the-curve (AUC) receiver-operator characteristic (ROC) value, abbreviated as AUC and representing capability of the model to distinguish the two outcome classes (Fawcett, 2006). The chance levels of classification accuracy (i.e., sum of squared weighted true positive rate and squared weighted true negative rate) were 53.35% and 55.62% for college enrollment and college graduation, respectively, and the chance level of AUC was .50. Averaged across the 5 folds in the cross-validation outer loop, the prediction accuracy levels indicate how well adolescents’ family experiences predict their college enrollment and graduation in young adulthood, and the AUC levels indicate how well these predictive models distinguish whether or not they were enrolled in and graduated from college.

**Question 2: Which family experiences are the key predictors of college entrance and graduation?** To identify which family experience were the strongest predictors of educational attainment, I conducted feature importance estimation and recursive feature elimination (RFE). For each trained model, I obtained the feature importance estimations indicating the relative contribution of each predictor to the model prediction. For models trained with regularized logistic regression, feature importance is usually indicated by the standardized regression coefficient for each predictor, ranging from -1 to 1, with 0 as the least important; in addition, the sign of each coefficient indicates the direction of the variable’s effect. For random forests, however, there have been a variety of approaches to estimate feature importance, and in this study, I used the mean decrease in (Gini) impurity, a widely-used index for random forests
feature importance that is indexed as the relative contribution of each predictor to the homogeneity (i.e., ‘purity’) of partitioning/classification summed across trees (Breiman, 1984). In this study, the possible range of random forests feature importance was 0 to 0.47 (i.e., 1 minus the sum of squared weighted proportion of value = 1 and squared weighted proportion of value = 0) for predictors in the model for college enrollment, and 0 to 0.44 in the model for college graduation, with higher scores indicating higher relative contributions and importance. I obtained the predictors’ feature importance indices for each predicted outcome with each algorithm by averaging across the feature importance estimations from the five trained models in the outer loop of the nested cross-validation. Further, I conducted RFE, a procedure that begins from the original model with the whole set of predictors and estimation of prediction accuracy and feature importance, followed by elimination of the predictor with the lowest importance, leading to a subset that includes the strongest predictors. Then, with the subset of predictors, the model is trained again, with the estimation of prediction accuracy and feature importance and further elimination of the least important predictor. In this study, the recursive procedure stops when only one predictor remains, and I identified the most important predictors from the model that retained the equivalent prediction accuracy as the original model and for which an additional elimination led to an evident drop of the accuracy.

**Question 3: What complex patterns merit further examination?** The third step of the analyses was to identify complex associations between family experience and educational attainment that might be examined more specifically in future work. For each outcome, I compared models trained by the two algorithms the sets of the most important predictors identified by RFE. I expected that predictors identified as important in the random forests model, but not as important in the regularized logistic regression model, would potentially be involved
in nonlinear and interactive effects. Given the complexity of random forests, an approach to interpret the model is to visualize it with partial dependence plots, which can depict the effect of each predictor as well as combination of predictors on the outcome probability, holding the rest of the predictors constant (Friedman, 2001; Molnar, 2019). Here, I examined and interpreted relations between single important predictors and the outcome probability using 2D partial dependence plots. Likewise I examined relations between pairs of important predictors and educational attainment using 3D plots; from among the numerous possibilities I chose findings that may have practical implications for intervention/prevention programs. Due to limitations in human perceptions, plots higher than three dimensions, that is, for interactions at three-way and higher, have been rarely used in ML (Molnar, 2019) and thus were not considered in this study.

**Software and Packages.**

All data analyses with the ML approach were implemented in Python3.6.0. Using the ‘Scikit-learn’ package (v0.20.3; Pedregosa et al., 2011), I implemented regularized logistic regression models within the ‘linear_model.LogisticRegression’ class, and random forests models within the ‘ensemble.RandomForestClassifier’ class. Further, I obtained partial dependence plots from the random forests models using the ‘Skater’ package (Choudhary, Kramer, & datascience.com team, 2018). Programming scripts are documented in the publicly available repository at https://github.com/xiaoransun/ML.family.edu.AddHealth for future research to replicate and extend the study approach and findings.

**Results**

**Do Family Experiences in Adolescence Predict Educational Attainment in Young Adulthood?**

The first question pertained to model performance, specifically, testing comprehensive models to determine how well the range of family experiences that have been studied in prior
research using Add Health data, can predict college enrollment and graduation. Table 4.3 shows performances of models predicting college enrollment and graduation trained by regularized logistic regression and random forests with the collection of 53 variables capturing adolescent family experiences that were examined in prior studies using Add Health data. In predicting college enrollment, these two algorithms had classification accuracies higher than chance level by 20.08% and 18.98% and AUC by 0.29 and 0.28, respectively. For college graduation, classification accuracies were also higher than chance level by over 23.48% and 23.45%, and AUC by 0.33 and 0.32, respectively. The accuracy results indicated that adolescents’ family experiences examined in previous studies using Add Health data, when taken together, can predict their educational attainment across about 13 years and among ‘unknown’ cases (those not involved in model training). In addition, the AUCs indicated that the comprehensive models could also distinguish whether or not young adults in ‘unknown’ cases enrolled in and graduated from college.

**Which Family Experiences Are Key Predictors of Educational Attainment?**

The second research question was aimed at identifying key predictors through feature importance estimation and RFE. Figures 4.2 and 4.3 show these results for predicting college enrollment. RFE reduced the original predictor collection to 12 predictors and 18 predictors, for regularized logistic regression and random forests, respectively; these models performed as well as the original models with the 53 predictors examined in the 101 relevant Add Health studies. Specifically, predictors identified by RFE as important in both algorithms (with directions of effects as estimated by the regularized logistic regression model above and beyond one another) were mother education (positive effect), family income (positive effect), father education (positive effect), adolescent biological sex (female youth higher), mother occupational prestige
(positive effect), and mother educational expectations (positive effect), all of which also ranked high (top 15) in feature importance across the two algorithms, indicating that these were key family experiences for college enrollment no matter whether the model included nonlinear and interactive effects. Although family socioeconomic characteristics dominated these important predictors, mothers’ educational expectations was a positive parenting factor that is potentially modifiable. Further, important predictors that were unique to the model trained with regularized logistic regression included smoker(s) in household (negative effect), being resident with two biological parents (positive effect), parent participating in parent-teacher association (positive effect), father nativity (father born outside US, higher), family receiving welfare (negative effect), and birth order (negative effect), indicating that these were key family experiences for college enrollment when the model imposed linearity and exclusivity among predictors while controlling for one another. Notably, these predictors were all dichotomous variables. In contrast, predictors including adolescent age, parent age, mother-adolescent shared activities, family social support, intergenerational closure, mother-adolescent relationship quality, father educational expectations, father-adolescent shared activities, household size, parent religiosity, father occupational prestige, and adolescent shared dinner with parents were important in the random forests model but not in the regularized logistic regression model, suggesting that these predictors (all continuous variables with more than 2 levels) could have nonlinear effects and/or be involved in interactions in relation to college enrollment probability.

Turning to college graduation, Figures 4.4 and 4.5 show the feature importance estimation and RFE results. Again, RFE reduced the original predictor collection, in this case to 17 predictors and 19 predictors for regularized logistic regression and random forests, respectively, to retain the prediction performance for the original models with 53 predictors. Predictors
identified as important in both algorithms were mother education (positive effect), family income (positive effect), father education (positive effect), parent in parent-teacher association (positive effect), mother educational expectations (positive effect), intergenerational closure (positive effect), parent age (positive effect), mother occupational prestige (positive effect), parental control (negative effect), and shared dinner with parents (positive effect). Further, important predictors that were unique to the model trained with regularized logistic regression included smoker(s) in household (negative effect), being resident with two biological parents (positive effect), biological sex, family receiving welfare (negative effect), father nativity (father born outside US, higher), mother obesity (negative effect), and birth order (negative effect), indicating that these were important family experiences for college graduation when the model imposed linearity and exclusivity among predictors while controlling for one another. Notably, these predictors were all dichotomous variables. In contrast, predictors including adolescent age, family social support, mother-adolescent shared activities, father occupational prestige, father-adolescent relationship quality, mother-adolescent relationship quality, father-adolescent shared activities, parent religiosity, and father educational expectations, which were all continuous variables with more than 2 levels, were important in the random forests model but not in the regularized logistic regression model, suggesting that these predictors could have nonlinear effects and/or be involved in interactions in relation to college graduation probability.

**What Complex Patterns of Predictors Merit Further Examination?**

The third research question was directed at identifying nonlinearities and interactions between the family experience predictors. To do so, following the RFE results I used partial dependent plots to illustrate links to college enrollment and graduation probability from predictors identified as important in the random forests model but not as important in the
regularized logistic regression model. Importantly, given ML’s exploratory nature, these effects cannot be considered statistically significant as in hypothesis-testing studies that use traditional statistical methods. Rather, the nonlinearities and interactions detected here direct attention to examining these effects in future hypothesis-testing studies in order to better understand families’ role in young adults’ educational attainment. From among the numerous possibilities for partial dependence plots, I chose example plots involving potentially modifiable family experience predictors that could serve as a focus for prevention and intervention studies after being further examined in hypothesis-testing studies. Plots for all patterns, however, can be obtained using the programming scripts and detailed instructions documented in the publicly accessible repository for this study at https://github.com/xiaoransun/ML.family.edu.AddHealth.

**Nonlinear patterns depicted by 2D partial dependence plots.** Illustrating patterns of relations between single predictors and the outcomes, Figure 4.6 shows 2D partial dependence plots for predictors of *college enrollment* probability (upper panel) and *college graduation* probability (lower panel). As elaborated below, these visualizations suggest nonlinear relations between family experiences and educational attainment that may worth examining in future hypothesis-testing studies.

**Mother-adolescent shared activities and college enrollment.** For this association, a positive linear-like, monotonically increasing pattern was evident at lower levels; however surpassing the level of four (out of ten shared activities with mothers in the past four weeks), the association turned relatively flat, suggesting a possible ‘good enough’ threshold for this factor with regard to college enrollment and directing future research to test log-transformed or curvilinear term of this variable or estimate spline models of its effects.

**Family social support and college enrollment.** This association was relatively flat at low
and high levels, but linear-like, monotonically increasing between the levels three and four (on a five-point scale), indicating a possible ‘activation’ point for this factor to begin to matter for college enrollment, as well as a ‘good enough’ threshold. Further, a negative dip in the association emerged at the highest level, indicating a potential negative effect of overinvolved parenting or enmeshment given that adolescence is an important period for autonomy development (Collins & Laursen, 2004). This result pattern suggests future research to examine the curvilinear effect or a spline model for how family social support for adolescents may influence their future educational attainment.

**Shared dinner with parents and college enrollment.** For this association, a curvilinear, monotonically increasing pattern emerged. Overall, more frequent shared dinner (in the past seven days) was associated with higher enrollment probability, but the slope was steeper at lower levels than at higher levels, suggesting a stronger positive effect of additional family meals in the low frequency range. Further, surpassing the level of four, the association was almost flat, again suggesting a possible ‘good enough’ threshold for this family experience factor.

**Family social support and college graduation.** Different from the pattern found in predicting college enrollment, the association here conformed to a curvilinear pattern whereby both low and high levels were associated with higher graduation probability, but probability was lower in the middle range. This pattern was different from both theory on the positive effects of family emotional resources and a prior result using Add Health data to show that family social support was negatively associated with college enrollment (Turley et al., 2010). Further hypothesis-testing studies are needed to examine and explain this pattern.

**Father educational expectations and college graduation.** This association was almost flat below the level of four (out of a five-point scale), but a strong positive effect emerged after four,
suggesting an ‘activation’ point. This result directs future research to test spline models or dichotomize this factor to examine at which level fathers’ expectations for adolescents begin to matter for their future college graduation outcome.

*Father-adolescent shared activities and college graduation*. This association was positive overall, but a negative dip emerged at high levels surpassing eight activities with fathers (out of ten activities, in the past four weeks), suggesting a ‘good enough’ point at the level of eight, and even a slight detrimental effect of ‘too close’ relationships with fathers.

**Two-way interaction patterns depicted by 3D partial dependence plots.** Visualizing patterns of relations between pairs of predictors and the outcomes, Figure 4.7 shows 3D partial dependence plots for two-way interactions predicting *college enrollment* probability (upper panel) and *college graduation probability* (lower panel), suggesting interaction effects between aspects of family experience on educational attainment that may worth examining in future hypothesis-testing studies. For each education outcome, I selected two plots to display: (1) for the interaction between a potentially modifiable family relationship/parenting practice factor and a family socioeconomic characteristic that may condition the effect of family processes and suggest a targeted intervention, and (2) the interaction between two potentially modifiable family processes, which would suggest how these may combine to influence educational attainment.

*Shared dinner with parents and father education predicting college enrollment*. The plot for this interaction pattern revealed that, although the positive effect of increasing the frequency of shared dinner with parents from zero to two occasions in past seven days was evident across all levels of father education, the benefit of increasing dinner frequency from two to four appeared stronger for adolescents with fathers whose education levels were greater than six (i.e., education after high school). By allowing nonlinearities in this interaction pattern, in contrast to
prior analyses that focused on interactions between linear effects, this result not only suggested an amplifying effect of high father education on the relation between shared dinner and college enrollment, but also indicated at what levels this amplification may be most likely to emerge.

**Intergenerational closure and father educational expectations predicting college enrollment.** The plot for this interaction pattern suggested a stronger effect of intergenerational closure in the range from one to three (i.e., parent talked to one to three parents of the adolescent’s friends in the last four weeks) at lower levels of father expectations. In other words, in this range, the positive effect of intergenerational closure may compensate for lower paternal educational expectations, both potentially modifiable family process factors.

**Father-adolescent shared activities and father occupational prestige predicting college graduation.** Here, a compensation pattern emerged: the positive effect of father-adolescent shared activities, especially at low levels (i.e., zero to two father-adolescent shared activities in the past four weeks), was more evident when fathers’ occupational prestige was lower. This pattern suggests the potential efficacy of parenting interventions targeting fathers in relatively low prestige jobs.

**Mother-adolescent relationship quality and father-adolescent shared activities predicting college graduation.** A compensation pattern also emerged in this interaction: Higher maternal relationship quality (above 4.25 on this 5-point scale) had a stronger positive effect on college graduation probability when adolescents reported fewer shared activities with fathers.

Taken together, these interaction patterns revealed the presence of both amplification and compensation effects involving different family experiences in adolescence, including interactions indicating how family socioeconomic characteristics may condition the effects of modifiable, family process factors, as well as interactions between modifiable family process
factors. Moreover, these interaction patterns, coupled with nonlinear patterns, also indicated effects were evident at some levels more than others.

Discussion

Educational attainment outcomes in young adulthood, especially college enrollment and graduation, are critical to individual well-being throughout adulthood (IOM & NRC, 2015). The ability to predict these outcomes from earlier experiences can advance both developmental theory and evidence-informed practice (Baltes et al., 1977; Pettit et al., 2009). Theoretical frameworks have highlighted important roles of adolescents’ family experiences in their future educational attainment, and numerous studies, mostly using an hypothesis-testing approach with regression-based methods, have built on these theories to test associations between a wide range of adolescent family experiences and young adult educational attainment. These include at least 101 studies which, to date, have used the Add Health longitudinal data to examine one or more adolescent family experience variables to predict future educational attainment. The current study took a further step to synthesize findings from across these 101 studies towards a more comprehensive understanding of the role of adolescent family experiences in young adult educational attainment. Specifically, analyses addressed three questions of both theoretical and practical importance: (1) *How accurately do the adolescent family experience factors examined in prior studies using Add Health data collectively predict young adult educational attainment?* (2) *When all are considered together, which family experience factors are the key predictors of young adult educational attainment?* (3) *What complex patterns, including nonlinearities and interactions, merit further examination?* To address these questions I applied an innovative, ML-based approach to synthesize adolescent family experiences examined in these studies, in total 53 variables, to train and test models predicting young adult college enrollment and graduation.
about 13 years later. With the models I also conducted feature importance estimation and RFE to identify key predictors, and illustrated complex patterns among the predictors using partial dependence plots. In addition to addressing questions about the role of family experiences in young adult education attainment, the ML-based approach used here provides a model for developmental and family scientists to synthesize existing literature and discover new findings using secondary, large-scale datasets.

**Adolescent Family Experiences Predict Young Adult College Enrollment and Graduation**

This study was designed first, to determine how accurately adolescent family experiences, examined in prior Add Health research, can together predict education attainment. I applied two ML algorithms, regularized logistic regression and random forests to train models predicting college enrollment and graduation, with the 53 family experience variables as predictors, and test model performances indexed by classification accuracy and AUC. Although the ML-based approach in this study was exploratory, the nested cross-validation process I employed was able to prevent model overfitting and revealed strong model prediction performances when applied to ‘unknown’ cases that were not involved in the training process (Yarkoni & Westfall, 2017).

Specifically, the results indicated that the prediction models for both algorithms and both outcomes had accuracies and AUCs that were well-above chance levels: When taken together, the adolescent family experience predictors examined in prior studies, predicted educational attainment outcomes 13 years later. In addition to relatively accurate prediction, the high AUCs indicated that the models were also able to distinguish the classes (i.e., 0 or 1) of the dichotomous outcomes, that is, whether or not the young adults were enrolled in college and whether or not they graduated from college. This finding supports theoretical tenets in family and developmental science highlighting the important role of family life in adolescence in future
educational attainment. And, forecasting adolescents’ future educational outcomes with information about their families also has practical implications: These built prediction models can be used to develop targeted prevention/intervention programs that promote youth’s education attainment—with long term implications for important adult life domains such as employment, family formation, and health (IOM & NRC, 2015).

**Key Family Predictors of Educational Attainment**

Given well-performing prediction models I next conducted feature importance estimation and RFE to identify the most important predictors, among the 53 family experience variables included in the models, for college enrollment and graduation. In previous studies these family experience variables were only examined in separate models. In this cross-study synthesis, I extended prior research by simultaneously including these predictors in the ML models through which I could compare their relative importance and identify the most important ones. First, the feature importance estimation provided rankings of the relative contribution of each family experience variable to the prediction of the outcomes with each algorithm; the feature importance indices provided by the regularized logistic regression indicated the directions of the predictions as well. Then, built on the feature importance estimation, RFE further identified collections of predictors—proportional to the entire 53-variable collection—that retained equivalent prediction accuracies for each outcome with each algorithm. For predicting college enrollment, 12 and 18 predictors were identified, respectively, for the two algorithms, and for predicting college graduation, 17 and 19 predictors were identified. This finding, that subsets of the predictors retained equivalent levels of prediction accuracy as the 53 predictors, has practical value in identifying family experiences that can be the focus of future data collection as well as prevention and intervention programs.
Findings revealed that family socioeconomic characteristics, especially family income and mothers’ and fathers’ education were consistent, key predictors, ranking high across the two algorithms and the two outcomes. These results support theories highlighting the role of family socioeconomic resources in youth achievement, including the life course perspective that illuminates transmission of attainments between generations (Elder, 1998), the capital theory that emphasizes influences of family financial and human capital on offspring’s achievement (Coleman, 1988), and family socialization perspectives that regard parent educational attainments as important models for youth attainment (Eccles, 2011). Given that theory and prior research rarely consider the relative importance of the range of family experiences they target, the current findings extend prior literature in documenting the predictive role of family socioeconomic characteristics compared to other dimensions of family experiences.

Although family socioeconomic characteristics were strong and consistent predictors, modifiable family relationships and parenting factors, of potential significance for prevention and intervention, also emerged as key predictors (Holder, 2010). For the college enrollment outcome, mother educational expectations emerged as a key predictor identified by RFEs across the two algorithms. For the college graduation outcome, predictors included parents’ involvement in parent-teacher associations, mother educational expectations, intergenerational closure, parental control, and shared dinner with parents. Such results are consistent with major theoretical perspectives in documenting that beyond family structure/context characteristics, (potentially modifiable) family processes also matter for young adult educational attainment.

Other key predictors emerged as important in one algorithm but not the other. Such results could be due to the algorithm differences, that is, regularized logistic regression constrains the associations between each predictor and the outcome to be linear and exclusive, whereas random
forests allow nonlinearities and interactions among predictors. In particular, the key predictors unique to the regularized logistic regression models were all dichotomous variables for which complex patterns, especially nonlinearity, were less likely to emerge, whereas the key predictors unique to the random forests models were all continuous variables with more than two levels that could more readily be involved in complex patterns. That is, the measurement and quantification methods for predictors could potentially influence their feature importance estimation depending on the nature of the ML algorithm implemented. Future studies should test different approaches to measuring the same factors and implement other algorithms to determine whether particular constructs emerge as key predictors across models.

In addition to predictors identified as important in the models, those not identified as important (i.e., ranked low in feature importance and eliminated by RFE) also provide useful information that adds to current literature. In particular, the results revealed that in both algorithms and for both outcomes, adolescents’ race/ethnicity never emerged as a key predictor.

**Directions for Future Research: Nonlinear and Interaction Patterns**

To discover complex patterns among predictors that may merit future examination, I applied partial dependence plots to interpret associations between predictors and college enrollment and graduation probability in the random forest models where nonlinearities and interactions were allowed. In particular, for each outcome I focused on interpreting potentially modifiable, family process predictors identified as important in the random forests model but not in the regularized logistic regression model given that these predictors were likely to be involved in nonlinear and/or interactive effects and be of practical significance.

The results with the 2D plots showed nonlinear patterns in associations between individual family experience variables and the outcomes, revealing ‘activation’ points surpassing which
predictors became more closely linked to the outcome, ‘good enough’ points at which the association between predictor and outcome turned from steep to flat, and negative dips at the highest levels of an otherwise positive predictor. With further validation by hypothesis-testing studies, these particular points within the nonlinear patterns may provide guidance for intervention programs to modify aspects of family dynamics in precise thus potentially cost-effective ways. For example, interventions can enhance factors with positive effects to levels higher than the ‘activation’ points, but stop by the ‘good enough’ points or the points turning into negative dips to save resources or prevent possible iatrogenic effects of family over-involvement.

The two-way interactions illustrated by the 3D plots showed how nonlinear effects of the modifiable family process predictors (i.e., in the dimension of family relationships and parenting) were qualified by family structure/context factors, in particular, family socioeconomic characteristics. These also showed how pairs of family process factors combined to predict college enrollment and graduation probability. Results indicated both amplifying and compensatory patterns between predictors, consistent with research findings using other datasets (Benner et al., 2016; Parcel et al., 2010; Sun et al., 2017). Different from previous studies, however, because these interactions also involved nonlinearities, the results also revealed at what levels of the predictors amplifying and/or compensatory effects may emerge—potentially important information to intervention and prevention. For example, the results pertaining to shared dinner with parents and father education in predicting college enrollment suggested that, although increasing the frequency of shared dinner from zero to twice a week could potentially benefit adolescents’ future educational outcomes, increasing the frequency from two to four appeared to benefit adolescents whose fathers had more than high school education.

As noted above, given the exploratory nature of the ML-based approach, this study used the
plots to *discover*, rather than *test*, complex patterns. These discovered patterns can provide directions for future hypothesis-testing studies to examine their statistical significance, and the significant patterns can then guide practice. This study’s use of publicly available, secondary data and documentation of programming scripts in a publicly accessible repository will facilitate such next steps. Moreover, if future investigators are interested in examining patterns of other important predictors identified in this study (especially in the random forests models), an exploratory first step may be to use programming scripts and models developed for this study and available online, to obtain partial dependence plots that can guide the selection of nonlinear and interaction models for estimation in the research.

**Limitations and Future Directions**

In face of its strengths, this study has limitations that may guide future research directions. First, this study, as a data-based cross-study synthesis, focused on Add Health data. Adolescents were in Grades 7-12 in 1994-1995 at Wave I and followed through their mid-20s, which may limit the generalizability of results to other cohorts of adolescents and young adults. The focus on using family experience variables examined in Add Health-based studies also prevented examination of other family experiences not examined in these studies or measured in the entire sample of Add Health such as sibling and interparental relationships. Thus, a future direction is to apply the study approach to other large-scale, longitudinal datasets to learn, for other population cohorts, whether family experiences in adolescence predict young adult educational attainment and whether the key predictors and complex patterns differ from those found in the current study. Second, although the model prediction performances were above chance levels, the accuracies or AUCs were less than perfect (i.e., 100%). Reasons could include: (1) Given the goal of examining *family experiences in adolescence* as predictors of young adult educational
attainment, adolescents’ personal characteristics (e.g., school achievement, career goals) and experiences in other contexts (e.g., in school, with peers) were not included as predictors, although such factors could potentially add to the ML model performances in predicting young adult educational attainment; (2) Following previous developmental studies, regularized logistic regression and random forests were used as algorithms, but other algorithms (e.g., support vector machine, neural networks) may be able to build models that perform better. Thus, if aiming for better-performing models predictive of young adult educational attainment, future directions include incorporating predictors beyond family experience variables to build prediction models and implementing other algorithms to find the optimal algorithm with best prediction performances. Third, despite the capacity of ML models to simultaneously include a large number of predictors and that of random forests model to allow nonlinearities and interactions, the models were unable to model mediation processes among predictors, which could otherwise help to better understand how family dynamics influence one another to predict educational attainment. A future direction toward revealing mediation mechanisms among these predictors is to implement exploratory mediation analysis via regularization (XMed), a recently developed ML-based approach (Serang & Jacobucci, 2019). Moreover, although random forests allow estimations of high-order interactions, due to limitations in human perception and constraints in interpretation techniques for models trained with random forests, it was difficult to illustrate interaction patterns beyond the two-way (Molnar, 2019). This limitation calls for future advancement of methods for interpreting and visualizing those complex patterns.

Overall, this study has important theoretical, methodological, and practical contributions. Theoretically, this study advanced understanding of how earlier family experiences may have implications for young adult college enrollment and graduation—outcomes that are important to
adult well-being across multiple domains. In particular, by addressing the three main research questions, this study (1) documented the significance of families’ role in adolescence for young adults’ educational attainment around 13 years later, (2) identified important factors in family experiences from among multiple dimensions of predictors of educational attainment, and (3) illuminated complex patterns in how these factors predicted educational attainment. As a methodological contribution, the ML-based approach for the synthesis based on a large-scale, longitudinal dataset provides a paradigm for synthesizing other studies in the family and developmental science fields. Practically, findings have implications for prevention and intervention programs to use key predictors to forecast adolescents’ future educational outcomes, identify at-risk youth, and with further validation in hypothesis-testing studies, target modifiable factors for change in precise and thus potentially cost-effective ways.
References

†Publications using Add Health data that examined associations between adolescent family experiences at Wave I and educational attainment outcomes at Wave III and/or IV.


1752-1758.


†Fomby, P. (2013). Family instability and college enrollment and completion. *Population*


†Needham, B. L. (2009). Adolescent depressive symptomatology and young adult educational


†Rosenbaum, J. E. (2018). Disabilities and degrees: Identifying health impairments that predict
lower chances of college enrollment and graduation in a nationally representative sample.

*Community College Review, 46, 145-175.*


Table 4.1.
Overview of Publications (N = 101) Using Add Health Data That Examined Effects of Family Experience Variables at Wave I on Educational Attainment Outcomes at Waves III or IV

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Educational Attainment Outcome</th>
<th>Family Experience Variables (Wave I)</th>
<th>Statistical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albrecht &amp; Albrecht, 2011</td>
<td>Number of years of education completed (Wave III; continuous)</td>
<td>Married v. single-parent family; Biological parents; Family income (in quantiles); Parent education</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Amis, Hussey, &amp; Okunade, 2014</td>
<td>High school graduation; College enrollment; College graduation with Bachelor’s degree; Graduate school enrollment (Wave IV; dichotomous)</td>
<td>Number of siblings; Mother education; Mother occupation; Father occupation; Mother/father-adolescent relationship quality</td>
<td>Probit regression</td>
</tr>
<tr>
<td>Ashtiani &amp; Feliciano, 2018</td>
<td>College enrollment, Bachelor’s degree attainment (Wave IV; dichotomous)</td>
<td>Two-parent v. single-parent family; Parent education; Mother/father educational expectations; Parent in parent-teacher association</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Bauldry, Shanahan, Russo, Roberts, &amp; Damian, 2016</td>
<td>Educational attainment (Wave IV; continuous—recoded: 1 = less than high school education, 5 = more than a four-year college degree)</td>
<td>Parent education; Family income (log)</td>
<td>Structural equation modeling</td>
</tr>
<tr>
<td>Benner &amp; Wang, 2014</td>
<td>Educational attainment (Wave III; continuous—recoded: 1 = high school dropout, 7 = 4-year college degree or higher)</td>
<td>Two biological parents v. other family structure; Family income; Parent nativity</td>
<td>Structural equation modeling</td>
</tr>
<tr>
<td>Benson, Johnson, &amp; Elder, 2012</td>
<td>Educational attainment (Wave IV; original coding)</td>
<td>Two biological parents v. other family structure; Parent education; Family income (log)</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
<td>--------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Bifulco, Fletcher, Oh, &amp; Ross, 2014</td>
<td>High school dropout, Attend some college, Complete one/two/three/four years of college or more (Wave III &amp; IV; dichotomous)</td>
<td>Mother education; Parent nativity and years in U.S.; Parent age</td>
<td>Multilevel modeling</td>
</tr>
<tr>
<td>Bissell-Havran, Loken, &amp; McHale, 2012</td>
<td>Sibling pairs’ college attendance for at least one year (Wave III; categorical: neither sibling, only sister, only brother, both siblings attended)</td>
<td>Two-parent biological v. other family structure; Parent education; Family income (log); Mother educational expectations (average and difference between siblings); Mother involvement in education (2-item; average and difference between siblings)</td>
<td>Multinomial logistic regression</td>
</tr>
<tr>
<td>Boardman, Alexander, Miech, MacMillan, &amp; Shanahan, 2012</td>
<td>High school graduation, college enrollment, college graduation (Wave IV; dichotomous)</td>
<td>Married v. single-parent family; Number of children in household; Parent education; Family income (log); Parent-adolescent shared activities (9-item); Parent age; Parent smoke; Parent obese</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Brinbaum &amp; Lutz, 2017</td>
<td>High school graduation (Wave III; dichotomous)</td>
<td>Parent education</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Brody, Yu, Miller, &amp; Chen, 2016</td>
<td>College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Single-parent status; Parent employment; Parent education; Family income (above v. below poverty); Parent receive public assistance; Parent economic hardship</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------</td>
<td>-------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Burge &amp; Beutel, 2018</td>
<td>College enrollment (Wave III; Categorical—recoded: did not enroll, enrolled in two-year college, enrolled in four-year college)</td>
<td>Two biological/adoptive parent v. other family structure; Parent education; Family receive public assistance; Mother/father educational expectations</td>
<td>Multinomial logistic regression</td>
</tr>
<tr>
<td>Campbell, 2009</td>
<td>College enrollment (Wave III; categorical: attending 2-year college, attending 4-year college, v. neither)</td>
<td>Two biological parents at home; Parent education; Family income (log)</td>
<td>Multinomial logistic regression</td>
</tr>
<tr>
<td>Carbonaro &amp; Workman, 2016</td>
<td>High school dropout (Wave III; dichotomous)</td>
<td>Parent education</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Cavanagh, Riegle-Crumb, &amp; Crosnoe, 2007</td>
<td>High school dropout (Wave III; dichotomous)</td>
<td>Two biological parent, single parent, stepparent, v. other family structure; Number of older sisters; Parent education</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Cherng, Calarco, &amp; Kao, 2013</td>
<td>College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Mother education; Family income</td>
<td>Logistic regression with fixed effects</td>
</tr>
<tr>
<td>Choi, Raley, Muller, &amp; Riegle-Crumb, 2008</td>
<td>Four-year college enrollment (Wave III; dichotomous)</td>
<td>Two-parent, single-parent, stepparent, v. other family structure; Parent education; Family income (standardized)</td>
<td>Multilevel logistic regression</td>
</tr>
<tr>
<td>Crosnoe, 2007</td>
<td>College enrollment (Wave III; dichotomous)</td>
<td>Two biological parents v. other family structure; Parent education</td>
<td>Logistic regression</td>
</tr>
</tbody>
</table>
Table 4.1. (Continued)

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Educational Attainment Outcome</th>
<th>Family Experience Variables (Wave I)</th>
<th>Statistical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doherty, Willoughby, &amp; Wilde, 2016</td>
<td>College enrollment (Wave IV; dichotomous)</td>
<td>Resident father presence</td>
<td>Relative risk ratios</td>
</tr>
<tr>
<td>Donlan, Prescott, &amp; Zaff, 2016</td>
<td>High school graduation, College enrollment (Wave III; dichotomous)</td>
<td>Mother involvement with schoolwork (2-item)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Drevon, Almazan, Jacob, &amp; Rhymer, 2016</td>
<td>High school graduation (Wave III; dichotomous); Number of years of education completed (Wave III; continuous)</td>
<td>Parent-adolescent relationship quality (3-item)</td>
<td>Logistic regression, Linear regression</td>
</tr>
<tr>
<td>Enayati &amp; Karpur, 2019</td>
<td>College enrollment (Wave IV; dichotomous)</td>
<td>Biological mother; Mother marital status; Mother education; Family receive welfare (3-item); Mother disability status</td>
<td>Linear probability model with fixed effects</td>
</tr>
<tr>
<td>Erickson, McDonald, &amp; Elder, 2009</td>
<td>Educational attainment (Wave III; highest degree achieved—ordinal)</td>
<td>Two biological parent v. other family structure; Parent education; Family income; Parent in parent-teacher association; Adolescent-parent relationship quality (4-item)</td>
<td>Ordinal logit model</td>
</tr>
<tr>
<td>Faas, Benson, &amp; Kaestle, 2013</td>
<td>Educational attainment (Wave IV; original coding)</td>
<td>Parent education; Parent occupational prestige; Family income (recoded into 1—10); Mother/father educational expectations</td>
<td>Structural equation modeling</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
<td>--------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Fasang, Mangino, &amp; Brückner, 2014</td>
<td>High school graduation (Wave III; dichotomous)</td>
<td>Two biological parent, step-parent, single-parent, v. other family structure; Parent education; Parent occupation; Family receive welfare; English as home language; Parent in parent-teacher association; Parent in school fund-raising; Parent-adolescent relationship quality (3-item); Parent intergenerational closure;</td>
<td>Multilevel modeling</td>
</tr>
<tr>
<td>Feliciano &amp; Ashtiani, 2012</td>
<td>College enrollment, College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Two-parent v. other family structure; Parent education; Family income (in poverty); Parent educational expectations</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Feliciano &amp; Lanuza, 2017</td>
<td>Number of years of education completed (Wave IV; continuous—recoded)</td>
<td>Two-parent v. other family structure; Household size; Parent education; Parent occupational prestige; Family income</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Fine, 2015</td>
<td>College graduation with a Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Family socioeconomic status (composite of parent education and occupational prestige)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Fishman &amp; Min, 2018</td>
<td>Number of years of education completed (Wave IV; continuous—recoded)</td>
<td>Household size; Mother relationship status; Birth order; Parent education; Parent occupation; Family income (log); Mother age at time of birth (linear, curvilinear, cubic; categorical with 5-year interval)</td>
<td>Linear regression; Multilevel modeling</td>
</tr>
</tbody>
</table>
Table 4.1.  
*(Continued)*

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Educational Attainment Outcome</th>
<th>Family Experience Variables (Wave I)</th>
<th>Statistical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fletcher, 2008</td>
<td>High school dropout, college enrollment (Wave III; dichotomous)</td>
<td>Parents married; Mother education; Family income</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Fletcher, 2010</td>
<td>Number of years of education completed (Wave III; continuous), High school dropout, college enrollment (Wave III; dichotomous)</td>
<td>Parents married; Mother education; Family income</td>
<td>Linear/logistic regression with fixed effects</td>
</tr>
<tr>
<td>Fletcher, 2015</td>
<td>College enrollment (Wave III; dichotomous)</td>
<td>Parents married; Number of siblings; Has older siblings; Mother education; Family income; Parent educational expectations; Mother-adolescent relationship quality</td>
<td>Logistic regression with fixed effects</td>
</tr>
<tr>
<td>Fletcher &amp; Lehrer, 2009</td>
<td>Number of years of education completed (Wave III; continuous)</td>
<td>Parents married; Birth order; Mother education; Family income; Parent age</td>
<td>Linear regression with fixed effects</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------</td>
<td>--------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Fomby, 2013</td>
<td>College enrollment, College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Two parents, single mother, v. mother and stepfather family; Mother education; Mother occupation; Family income (4 levels); Mother educational expectations; Mother-adolescent relationship quality; Mother supervision (3-item); Frequency of family dinners; Mother in parent-teaching association; Mother intergenerational closure; Mother health; Mother met the adolescent’s teachers; Mother in school fundraising; Mother age at birth; Mother cohabitating at birth; Mother single at birth;</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>French, Homer, Popovici, &amp; Robins, 2015</td>
<td>Educational attainment (Wave IV; categorical—recoded, 7 levels)</td>
<td>Intact v. non-intact family; Single-parent family; Residential mother/father presence; Birth order Mother/father education; Mother/father occupation; Parent receive welfare</td>
<td>Ordered probit model</td>
</tr>
<tr>
<td>Fruith &amp; Chan, 2018</td>
<td>Educational attainment (Wave IV; continuous—recoded: 1 = did not finish high school, 10 = doctoral or professional degree)</td>
<td>Parent education</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
<td>-------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Gibbs, Erickson, Dufur, &amp; Miles, 2015</td>
<td>College enrollment (Wave IV; categorical: never enrolled in post-secondary education, enrolled in a vocational or technical college, enrolled in a 4-year college)</td>
<td>Two biological parents v. other family structure; Mother/father work status; Family income (log) Parent receive public assistance</td>
<td>Multinomial logistic regression</td>
</tr>
<tr>
<td>Gillette, &amp; Gudmunson, 2014</td>
<td>Educational attainment (Wave IV; original coding)</td>
<td>Family income; Family receive welfare (3-item); Duration of father absence</td>
<td>Structural equation modeling with Bayesian estimation</td>
</tr>
<tr>
<td>Glanville, Sikkink, &amp; Hernández, 2008</td>
<td>High school dropout (Wave III; dichotomous)</td>
<td>Two biological parent, single-parent, step family, v. other family structure; Parent education; Family income Intergenerational closure; Family social support (3-item)</td>
<td>Structural equation modeling</td>
</tr>
<tr>
<td>Glass, Sutton, &amp; Fitzgerald, 2015</td>
<td>Education attainment (Wave IV; categorical—recoded; no college, some college, college degree or higher)</td>
<td>Two biological parents v. other family structure; Number of siblings; Parent education; Family income (log)</td>
<td>Multinomial logistic regression</td>
</tr>
<tr>
<td>Gordon &amp; Cui, 2012</td>
<td>Number of years of education completed (Wave IV; continuous—recoded)</td>
<td>Two biological parents, stepfamilies, single-mother, single-father, v. other family structure; Parent education; Mother/father involvement with school work (3-item); Mother/father educational expectations (2-item); Mother/father-adolescent relationship quality (2-item);</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
<td>--------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Gorry, 2016</td>
<td>High school graduation (Wave III &amp; IV; dichotomous)</td>
<td>Parent education</td>
<td>Logistic regression with fixed effects</td>
</tr>
<tr>
<td>Gorry, 2017</td>
<td>High school graduation (Wave III; dichotomous), Number of years of education completed (Wave IV; continuous—recoded)</td>
<td>Parent education</td>
<td>Logistic/linear regression with fixed effects</td>
</tr>
<tr>
<td>Greenman, 2013</td>
<td>High school graduation, College enrollment (Wave III; dichotomous)</td>
<td>Two biological parents, stepparent, single parent, v. no parents present; Parent education; Family income</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Holmes, 2018</td>
<td>Educational attainment (Wave IV; categorical—recoded: less than high school, high school graduate, vocational or technical school, some college, completed bachelor's degree, and completed schooling beyond a bachelor's degree)</td>
<td>Married v. not married parents; Parent education; Parent nativity; Smoker(s) in the household</td>
<td>Multinomial logistic regression</td>
</tr>
<tr>
<td>Humberstone, 2018</td>
<td>High school graduation (Wave III; dichotomous)</td>
<td>Presence of a parent figure; Parent education; Parent receive public assistance</td>
<td>Multilevel logit modeling</td>
</tr>
<tr>
<td>Huynh-Hohnbaum, Bussell, &amp; Lee, 2015</td>
<td>High school graduation (Wave IV; dichotomous)</td>
<td>Parent education</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
<td>-------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Johnston, 2017</td>
<td>Number of years of education completed (Wave III; continuous), High school graduation (Wave III; dichotomous)</td>
<td>Mother marital status; Mother education; Parent receive public assistance</td>
<td>Coarsened exact matching; Propensity score</td>
</tr>
<tr>
<td>Kane, Morgan, Harris, &amp; Guilkey, 2013</td>
<td>Number of years of education completed (Wave IV; continuous—recoded)</td>
<td>Two-parent family v. other family structure; Parent education; Family income-to-needs ratio</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Kapinos &amp; Yakusheva, 2016</td>
<td>Educational attainment (Wave IV; categorical—recoded: less than high school, high school, any college, 4-year college, and higher)</td>
<td>Two-parent family v. other family structure; Number of siblings in household; Birth order; Mother age; Mother education; Family income; Mother nativity; Mother smoke</td>
<td>Ordinal logistic regression</td>
</tr>
<tr>
<td>Karpur, Nazarov, Brewer, &amp; Bruyère, 2014</td>
<td>College enrollment (Wave IV; dichotomous)</td>
<td>Mother marital status; Biological mother; Mother education; Parent receive welfare (3-item); Parent in parent-teacher association; Parent health; Mother disability status</td>
<td>Linear probability model</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
<td>--------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Keller &amp; Tillman, 2008</td>
<td>College enrollment (Wave III; dichotomous)</td>
<td>Two-biological parents, stepparents, single-parent, v. other family structure</td>
<td>Logistic regression; Multinomial logistic regression</td>
</tr>
<tr>
<td></td>
<td>College attendance type (Wave III; categorical—recoded: not attending, attending two-year college, attending four-year college)</td>
<td>Parent education; Family income (recoded into categorical—5 levels); Mother work status; Language spoken at home; Parental control (6-item); Parent involvement (4-item); Parent educational expectations</td>
<td></td>
</tr>
<tr>
<td>Kim, 2018</td>
<td>Number of years of education completed (Wave IV; continuous—recoded), College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Mother education; Family income</td>
<td>Linear/logistic regression with fixed effects</td>
</tr>
<tr>
<td>Lundberg, 2013</td>
<td>College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Mother education</td>
<td>Linear probability model</td>
</tr>
<tr>
<td>Mahatmya, &amp; Smith, 2017</td>
<td>Matching between college expectations (Wave I) and college attendance (Wave III)</td>
<td>Two-parent v. other family structure; Parent education; Family income; Parent educational expectations; Parent-adolescent relationship quality (16-item: 7-item for father, 9-item for mother); Parent-adolescent shared activities (22-item: 11-item for mother and father, respectively)</td>
<td>Multilevel multinomial logistic regression</td>
</tr>
<tr>
<td>Mangino, 2012</td>
<td>High school graduation, College enrollment (Wave III; dichotomous)</td>
<td>Two biological parents v. other family structure; Family income (linear, quadratic)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------</td>
<td>------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Mangino, 2014</td>
<td>College graduation with Bachelor’s degree given college enrollment (Wave IV; dichotomous)</td>
<td>Two biological parents v. other family structure; Number of siblings; Parent education; Family income (linear, quadratic, top 1%); Intergenerational closure</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Maralani, 2014</td>
<td>College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Female-headed v. other household; Parent education; Parent educational expectations; Parent smoke</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Maslow, Haydon, McRee, &amp; Halpern, 2012</td>
<td>College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Parent education; Mother/father-adolescent relationship quality (4-item)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>McLeod, Uemura, &amp; Rohrman, 2012</td>
<td>Educational attainment (Wave IV; categorical—recoded: received no degree, received GED or high school equivalency, received high school diploma, completed technical training, completed some college classes, received bachelor’s degree, received higher degree).</td>
<td>Two biological parents, single parent, v. other family structure; Parent education; Family income; Family receive public assistance</td>
<td>Ordinal logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Mears &amp; Siennick, 2016</td>
<td>Educational attainment (Wave III &amp; Wave IV; continuous—recoded: 1 = 8th grade or less, 5 = one or more years of graduate school)</td>
<td>Intact v. nonintact family; Number of siblings; Family socioeconomic status (composite: parent education, parent occupation); Language spoken at home; Alcoholic biological parent; Illegal drugs in household; Mother/father-adolescent relationship quality (2-item)</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Migali &amp; Zucchelli, 2017</td>
<td>High school graduation (Wave III &amp; IV; dichotomous)</td>
<td>Mother/father education; Mother/father occupation; Family access to medical care; Family unhealthy behaviors (composite: parent smoke or drink, family member smoke); Parent health; Parent report of partner’s health;</td>
<td>Linear probability model</td>
</tr>
<tr>
<td>Miller, &amp; Barnes, 2015</td>
<td>College graduation (Wave IV; dichotomous)</td>
<td>Mother education; Mother occupation; Mother receive public assistance; Mother/father-adolescent relationship quality (2-item) Mother smoke</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Mollborn, &amp; Everett, 2015</td>
<td>High school graduation, College enrollment, College graduation with a Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Parent education; Parent educational expectations</td>
<td>Logistic regression</td>
</tr>
</tbody>
</table>
Table 4.1.  
(Continued)

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Educational Attainment Outcome</th>
<th>Family Experience Variables (Wave I)</th>
<th>Statistical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monserud &amp; Elder, 2011</td>
<td>High school completion, College enrollment, College graduation with a Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Households with two biological parents, two biological parents with a grandparent, never married mother with a grandparent, divorced mother with a grandparent, grandparents with no parents present, never-married mother, divorced mother, and stepparent (All dichotomous); Number of siblings; Duration of current household structure; Mother/grandmother number of marital transitions; Mother/grandmother education; Family income (log); Mother/grandmother receiving public assistance; Mother involvement with schoolwork (2-item); Mother educational expectations; Mother-adolescent relationship quality (4-item); Parental control (7-item); Shared dinner with parent; Mother intergenerational closure</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Muller, Riegle-Crumb, Schiller, Wilkinson, &amp; Frank, 2010</td>
<td>Current college attendance or college graduation (Wave III; dichotomous)</td>
<td>Parent education</td>
<td>Multilevel modeling</td>
</tr>
</tbody>
</table>
Table 4.1.  
(Continued)

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Educational Attainment Outcome</th>
<th>Family Experience Variables (Wave I)</th>
<th>Statistical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needham, 2009</td>
<td>High school dropout, College enrollment (Wave III; dichotomous)</td>
<td>Intact v. nonintact family structure; Mother/father education; Family income</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Nguyen et al., 2012</td>
<td>Educational attainment (Wave IV; categorical—recoded: &lt; high school, high school/GED, some college/AA degree, &gt;= 4-year college)</td>
<td>Two biological parents, two parents, single parent/other family structure; Parent education; Parent receive public assistance; Mother/father-adolescent relationship quality (5-item)</td>
<td>Multinomial logistic regression</td>
</tr>
<tr>
<td>Nichols, Loper, &amp; Meyer, 2016</td>
<td>Educational attainment (Wave IV; continuous—recoded: 1 = 8th grade or less, 10 = professional training beyond 4-year college)</td>
<td>Parent education; Parent receive public assistance; Adolescent connectedness to family and parents (13-item)</td>
<td>Multilevel modeling</td>
</tr>
<tr>
<td>Okunade, Hussey, &amp; Karakus, 2009</td>
<td>On-time high school graduation (Wave III; dichotomous)</td>
<td>Live with both parents; Mother education; Mother occupation</td>
<td>Probit regression</td>
</tr>
<tr>
<td>Owens, 2010</td>
<td>High school graduation, College graduation with Bachelor’s degree (Wave III; dichotomous)</td>
<td>Mother/father education; Family income; Family receive public assistance; Mother/father educational expectations; Duration of living in current residence;</td>
<td>Multilevel logit modeling</td>
</tr>
<tr>
<td>Park, 2018</td>
<td>College enrollment (Wave IV; dichotomous)</td>
<td>Parent marital status; Parent education; Family income (log); Family support (4-item)</td>
<td>Karlson-Holm-Breen (KHB) analyses</td>
</tr>
<tr>
<td>Paschall, Bersamin, Finan, &amp; Zhang, 2019</td>
<td>Number of years of education completed (Wave III &amp; IV; continuous)</td>
<td>Parent education</td>
<td>Multilevel modeling</td>
</tr>
</tbody>
</table>
Table 4.1. (Continued)

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Educational Attainment Outcome</th>
<th>Family Experience Variables (Wave I)</th>
<th>Statistical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patacchini, Rainone, &amp; Zenou, 2017</td>
<td>Number of years of education completed (Wave IV; continuous—recoded)</td>
<td>Household size; Parent education; Parent occupation</td>
<td>Linear regression; 2SLS estimation with network model</td>
</tr>
<tr>
<td>Pate, Maras, Whitney, &amp; Bradshaw, 2017</td>
<td>Number of years of education completed (Wave III; continuous)</td>
<td>Family socioeconomic status (composite: family income, mother education, father education)</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Pearson &amp; Wilkinson, 2017</td>
<td>High school graduation &amp; college enrollment (Wave III; dichotomous) College graduation with Bachelor’s degree (Wave IV, dichotomous)</td>
<td>Two biological parents, stepparent family, single parent, v. other family type; Parent education</td>
<td>Logistic regression with fixed effects</td>
</tr>
<tr>
<td>Perreira, Harris, &amp; Lee 2006</td>
<td>High school dropout (Wave III; dichotomous)</td>
<td>Two-parent, stepparent, single-mother, single-father, v. other family type; Number of siblings at home; Parent education; Mother work status</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Rees &amp; Sabia, 2009</td>
<td>High school graduation, College enrollment (Wave III; dichotomous)</td>
<td>Parent separated, widowed, divorced, v. other status; Single parent; Number of siblings; Birth order; Parent education; Family income (log); Mother occupation; Mother-adolescent relationship quality (4-item); Parent involvement with schoolwork (4-item)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Rees &amp; Sabia, 2011</td>
<td>High school graduation, College enrollment (Wave III; dichotomous)</td>
<td>Parent never married, divorced, separated, widowed, v. married status; Has an older sibling; Parent education;</td>
<td>Propensity score</td>
</tr>
<tr>
<td>Rosenbaum, 2018</td>
<td>College enrollment, Community college graduation, 4-year college graduation (Wave IV; dichotomous)</td>
<td>Mother education; Family income (log)</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Ryabov, 2009</td>
<td>College enrollment (Wave III; dichotomous)</td>
<td>Resident with both parents, single parent, v. other; Number of siblings in household; Mother/father educational expectations (2-item); Parent-adolescent shared activities (10-item); Family social support (4-item)</td>
<td>Multilevel logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
<td>--------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Ryabov, 2011</td>
<td>High school graduation (Wave III; dichotomous)</td>
<td>Two-parent, single-parent, non-parent family; Parent education; Family income (Transformed with Box-Cox method); Mother/father educational expectations (2-item); Parent-adolescent shared activities (10-item); Parent supervision (3-item)</td>
<td>Multilevel modeling</td>
</tr>
<tr>
<td>Ryabov, 2013</td>
<td>College enrollment (Wave III; dichotomous)</td>
<td>Two-parent, single-parent, non-parent family; Parent education; Family income (Transformed with Box-Cox method); Mother/father educational expectations (2-item); Parent-adolescent shared activities (10-item); Parent supervision (3-item)</td>
<td>Multilevel modeling</td>
</tr>
<tr>
<td>Ryabov, 2016</td>
<td>High school graduate, College enrollment, College graduation with a Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Two-parent v. other family structure; Parent education; Family income (log); Parent-adolescent shared activities (10-item); Family social support (4-item)</td>
<td>Multilevel logistic regression</td>
</tr>
<tr>
<td>Ryabov, 2018</td>
<td>High school graduation, college enrollment, College graduation with a Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Two-parent household v. other; Parent education; Family income</td>
<td>Multilevel logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
<td>-------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Ryabov, 2019</td>
<td>Educational attainment (Wave IV; continuous—recoded: 1 = <em>less than high school</em>, 5 = <em>more than a 4-year degree</em>)</td>
<td>Two-parent, single-parent, v. non-parent family structure; Family socioeconomic status (composite of parent education, parent occupational prestige, and family income); Parent-adolescent shared activities (10-item); Parent-adolescent relationship quality: Parental attachment (4-item) &amp; Parental engagement (5-item)</td>
<td>Structural equation modeling</td>
</tr>
<tr>
<td>Sabia &amp; Rees, 2009</td>
<td>High school graduation, College enrollment (Wave III; dichotomous)</td>
<td>Parent single, divorced, separated, widowed, v. other; Number of biological siblings; Has an older sibling; Family income (log); Parent education</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Sabia &amp; Rees, 2011</td>
<td>Number of years of education completed (Wave III, continuous); High school graduation, college enrollment (Wave III, dichotomous)</td>
<td>Parent marital status; Parent education; Family income</td>
<td>Linear/logistic regression with fixed effects</td>
</tr>
<tr>
<td>Sabia &amp; Rees, 2015</td>
<td>High school graduation, College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Parent marital status; Number of biological siblings; Has an older sibling; Parent education; Family income</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Sabia, Wang, &amp; Cesur, 2017</td>
<td>High school graduation, College enrollment (Wave IV; dichotomous)</td>
<td>Parent marital status; Parent education; Family income (log)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------</td>
<td>--------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Shanahan, Bauldry, Roberts, Macmillan, &amp; Russo, 2014</td>
<td>Educational attainment (Wave IV; continuous—recoded: 1 = less than high school, 5 = more than a 4-year degree)</td>
<td>Two biological parents, two parent (one not biological), single mother, single father, v. other family structure; Parent education</td>
<td>Seemingly unrelated regression</td>
</tr>
<tr>
<td>Smith, Wiersma-Mosley, Ham, &amp; Moon, 2016</td>
<td>Educational attainment (Wave III; continuous—recoded: 0 = no high school diploma, 3 = completing one to five or more years of graduate school)</td>
<td>Parent education; Adolescent connectedness to family and parent (6-item; composite of parent-adolescent relationship quality and family social support)</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Spriggs &amp; Halpern, 2008</td>
<td>College enrollment (Wave III; dichotomous)</td>
<td>Two biological parent v. other family structure; Parent education</td>
<td>Poisson regression</td>
</tr>
<tr>
<td>Staff &amp; Kreager, 2008</td>
<td>High school dropout, College graduation with Bachelor’s degree (Wave III; dichotomous), years of education (Wave III; continuous)</td>
<td>Two biological parent v. other family structure; Parent education</td>
<td>Linear/logistic regression</td>
</tr>
<tr>
<td>Stokes, 2008</td>
<td>High school graduation (Wave III; dichotomous)</td>
<td>Two biological parent v. other family structure; Parent education; Family income; Parent involvement with schoolwork (3-item) Parent educational expectations; Mother/father relationship quality (5-item); Intergenerational closure; Attend church (a shared activity) with mother/father; Parent religiosity (2-item)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
<td>--------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Tabler &amp; Utz, 2015</td>
<td>Number of years of education completed (Wave IV; continuous—recoded)</td>
<td>Mother education</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Turley, Desmond, &amp; Bruch, 2010</td>
<td>College enrollment (Wave III; dichotomous)</td>
<td>Single biological parent, biological parent with partner, v. no biological parent in the household; Parent education; Mother/father-adolescent relationship quality (5-item); Family social support (3-item); Mother/father-adolescent shared activities: involvement (5-item) &amp; communication (4-item)</td>
<td>Structural equation modeling</td>
</tr>
<tr>
<td>Ueno, Roach, &amp; Peña-Talamantes, 2013</td>
<td>Educational attainment (Wave IV; categorical—recoded: 1= less than high school diploma or GED, 2 = high school diploma or GED, 3 = post-secondary vocational or technical certificate or license, 4 = associate degree, 5 = Bachelor’s degree, 6 = graduate or professional degree)</td>
<td>Resident with two biological parents; Parent education</td>
<td>Ordered logistic regression</td>
</tr>
<tr>
<td>Walsemann, Lindley, Gentile, &amp; Welihinda, 2014</td>
<td>Educational attainment (Wave IV; categorical—recoded: 1= high school diploma or less, 2 = some college or Associate’s degree, 3 = Bachelor’s degree or higher)</td>
<td>Nuclear, step-family, female-headed, extended family, v. other household; Family socioeconomic status (composite: parent education, parent occupation, family poverty)</td>
<td>Ordered logit regression</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Educational Attainment Outcome</td>
<td>Family Experience Variables (Wave I)</td>
<td>Statistical Method</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Wilczak, 2014</td>
<td>Educational attainment (Wave III; categorical—recoded: high school dropout, high school graduation, college enrollment)</td>
<td>Two biological parents v. other family structure; Family socioeconomic status (composite of parent education and occupation status)</td>
<td>Multinomial logistic regression</td>
</tr>
<tr>
<td>Wilkinson, 2010</td>
<td>High school graduation, College current attendance or graduation (Wave III; dichotomous)</td>
<td>Two biological parents v. other family structure; Parent education; One, two, or no Latino parent; Language spoken at home; Parent involvement with schoolwork (6-item; 3 items for mother and father, respectively)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Wilkinson &amp; Pearson, 2015</td>
<td>College enrollment (Wave III; dichotomous); College graduation with Bachelor’s degree (Wave IV; dichotomous)</td>
<td>Two biological parents v. other family structure; Parent education</td>
<td>Logistic regression</td>
</tr>
</tbody>
</table>

*Note. Most family experience variables are single items in measures; composite variables with multiple items are noted with number of items. For similar constructs using different terminologies in different studies (e.g., mother-adolescent relationship quality, closeness, warmth) that shared raw survey items, I chose the commonly used terminologies to list in this table.*
<table>
<thead>
<tr>
<th>Category</th>
<th>Constructs</th>
<th>Variables</th>
<th>Example Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family structure</td>
<td>Resident mother/father presence</td>
<td>‘Mother’, ‘Father’</td>
<td>French et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>Resident with two biological parents</td>
<td>‘TwoBioParent’</td>
<td>Benson et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>Household size</td>
<td>‘HHsize’</td>
<td>Patacchini et al. (2017)</td>
</tr>
<tr>
<td></td>
<td>Number of siblings</td>
<td>‘SibNum’</td>
<td>Fletcher (2015)</td>
</tr>
<tr>
<td></td>
<td>Birth order</td>
<td>‘BirthOrder’</td>
<td>Fletcher &amp; Lehrer (2009)</td>
</tr>
<tr>
<td>Family socioeconomic characteristics</td>
<td>Mother/Father education levels</td>
<td>‘momeduc’, ‘dadeduc’</td>
<td>Faas et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>Mother/Father occupational prestige</td>
<td>‘momjob’, ‘dadjob’</td>
<td>Feliciano &amp; Lanuza (2017)</td>
</tr>
<tr>
<td></td>
<td>Family income (log)</td>
<td>‘famic’</td>
<td>Bauldry et al. (2016)</td>
</tr>
<tr>
<td></td>
<td>Parent receiving public assistance</td>
<td>‘PAassistance’</td>
<td>Gibbs et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>Family receive welfare (3-item)</td>
<td>‘welfare’</td>
<td>Enayati &amp; Karpur (2019)</td>
</tr>
<tr>
<td></td>
<td>Parent economic hardship</td>
<td>‘PAeohard’</td>
<td>Brody et al. (2016)</td>
</tr>
<tr>
<td>Family relationships and parenting</td>
<td>Mother/Father involvement with schoolwork (3-item)</td>
<td>‘minvolve’, ‘dinvolve’</td>
<td>Gordon &amp; Cui (2012)</td>
</tr>
<tr>
<td></td>
<td>Parent in school fund-raising</td>
<td>‘PAfund’*</td>
<td>Fasang et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>Parent met teachers</td>
<td>‘PAteacher’*</td>
<td>Fomby (2013)</td>
</tr>
<tr>
<td></td>
<td>Mother/Father educational expectations (2-item)</td>
<td>‘mexp’, ‘dexp’</td>
<td>Gordon &amp; Cui (2012)</td>
</tr>
<tr>
<td></td>
<td>Parental control (7-item)</td>
<td>‘control’</td>
<td>Monserud &amp; Elder (2011)</td>
</tr>
<tr>
<td></td>
<td>Mother/father supervision (3-item)</td>
<td>‘mspV’, ‘dspv’</td>
<td>Ryabov (2013)</td>
</tr>
<tr>
<td></td>
<td>Family social support (4-item)</td>
<td>‘famsup’</td>
<td>Ryabov (2016)</td>
</tr>
<tr>
<td></td>
<td>Shared dinner with parents</td>
<td>‘dinner’</td>
<td>Monserud &amp; Elder (2011)</td>
</tr>
<tr>
<td></td>
<td>Intergenerational closure</td>
<td>‘PAclosure’</td>
<td>Glanville et al. (2008)</td>
</tr>
</tbody>
</table>
### Table 4.2. (Continued)

<table>
<thead>
<tr>
<th>Parent characteristics</th>
<th>Mother/Father nativity</th>
<th>‘mnativity, ‘dnativity’</th>
<th>Holmes (2018)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent age</td>
<td>‘PAge’</td>
<td></td>
<td>Fletcher &amp; Lehrer (2009)</td>
</tr>
<tr>
<td>Parent health</td>
<td>‘PAhealth’</td>
<td></td>
<td>Fomby (2013)</td>
</tr>
<tr>
<td>Parent smoking</td>
<td>‘PAsmoke’</td>
<td></td>
<td>Boardman et al. (2012)</td>
</tr>
<tr>
<td>Mother/Father alcoholic</td>
<td>‘malcohol’, ‘dalcohol’</td>
<td></td>
<td>Mears &amp; Siennick (2016)</td>
</tr>
<tr>
<td>Mother/Father obese</td>
<td>‘mobese’, ‘dobese’</td>
<td></td>
<td>Boardman et al. (2012)</td>
</tr>
<tr>
<td>Mother/Father disabled</td>
<td>‘mdisable’, ‘ddisable’</td>
<td></td>
<td>Enayati &amp; Karpur (2019)</td>
</tr>
<tr>
<td>Parent religiosity (2-item)</td>
<td>‘PArelig’</td>
<td></td>
<td>Stokes (2008)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other family characteristics</th>
<th>Smoker(s) in household</th>
<th>‘HHsmoke’</th>
<th>Holmes (2018)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Illegal drugs in household</td>
<td>‘HHdrug’</td>
<td>Mears &amp; Siennick (2016)</td>
</tr>
<tr>
<td></td>
<td>Family access to medical care</td>
<td>‘fammed’</td>
<td>Migali &amp; Zucchelli (2017)</td>
</tr>
<tr>
<td></td>
<td>English as home language</td>
<td>‘EnglishHome’</td>
<td>Fasang et al. (2014)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adolescent demographic characteristics</th>
<th>Biological sex</th>
<th>‘biosex’</th>
<th>Holmes (2018)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adolescent age</td>
<td>‘YAge’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nativity</td>
<td>‘nativity’</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Total = 55 variables. All variables were measured at Wave I.

*Variables that were missing over 25% and thus were removed for imputation and ML models.
Table 4.3.
Results of Tuning and Prediction Performances of Models Predicting College Enrollment and Graduation Trained by Regularized Logistic Regression and Random Forests

<table>
<thead>
<tr>
<th></th>
<th>Regularized logistic regression</th>
<th>Random forests</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\lambda)</td>
<td>Accuracy</td>
<td>AUC</td>
</tr>
<tr>
<td>Predicting college enrollment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CV1</td>
<td>0.1</td>
<td>72.09%</td>
<td>0.7834</td>
</tr>
<tr>
<td>CV2</td>
<td>0.001</td>
<td>73.64%</td>
<td>0.7945</td>
</tr>
<tr>
<td>CV3</td>
<td>1</td>
<td>73.49%</td>
<td>0.7758</td>
</tr>
<tr>
<td>CV4</td>
<td>500</td>
<td>75.23%</td>
<td>0.8142</td>
</tr>
<tr>
<td>CV5</td>
<td>100</td>
<td>72.71%</td>
<td>0.7962</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>73.43%</td>
<td>0.7928</td>
</tr>
<tr>
<td>Predicting college graduation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CV1</td>
<td>50</td>
<td>79.58%</td>
<td>0.8421</td>
</tr>
<tr>
<td>CV2</td>
<td>500</td>
<td>78.51%</td>
<td>0.8271</td>
</tr>
<tr>
<td>CV3</td>
<td>10</td>
<td>81.46%</td>
<td>0.8627</td>
</tr>
<tr>
<td>CV4</td>
<td>10</td>
<td>77.98%</td>
<td>0.8386</td>
</tr>
<tr>
<td>CV5</td>
<td>10</td>
<td>77.99%</td>
<td>0.8193</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>79.10%</td>
<td>0.8379</td>
</tr>
</tbody>
</table>

Note. CV = Cross-validation (outer loop). Chance level of accuracy = 53.35% and 55.62% for college enrollment and graduation, respectively. Chance level of AUC = 0.50.
Figure 4.1. An illustration of nested 5-fold cross-validation implemented in this study. (Figure crafting in reference to Raschka, 2013-2019)
Figure 4.2. Feature importance rankings for predicting college enrollment with regularized logistic regression (left) and random forests (right). In the left panel, blue bars indicate negative coefficients and violet bars indicate positive coefficients.
Figure 4.3. Recursive feature elimination results for predicting college enrollment with regularized logistic regression (upper) and random forests (lower) and the models selected in order to identify important features (i.e., predictors). Display order of the features is in the reversed elimination order. Features identified as important in both algorithms are bolded.
Figure 4.4. Feature importance rankings for predicting college graduation with regularized logistic regression (left) and random forests (right). In the left panel, blue bars indicate negative coefficients and violet bars indicate positive coefficients.
Figure 4.5. Recursive feature elimination results for predicting college graduation with regularized logistic regression (upper) and random forests (lower) and the models selected in order to identify important features (i.e., predictors). Display order of the features were in the reversed elimination order. Features identified as important in both algorithms are bolded.
Figure 4.6. Partial dependence plots (2D) of selected important family experience predictors in the random forests model predicting the probability of college enrollment (upper panel) and college graduation (lower panel).
Figure 4.7. Partial dependence plots (3D) of the random forests model predicting the probability of college enrollment (upper panel) from the interactions between shared dinner with parents and father education (left) and between intergenerational closure and father educational expectations (right), and model predicting the probability of college graduation (lower panel) from the interactions between father-adolescent shared activities and father occupational prestige (left) and between mother-adolescent relationship quality and father-adolescent shared activities (right).
CHAPTER 5

GENERAL DISCUSSION

This dissertation contributes to understanding of the mechanisms underlying families’ role in youth career development. Built on existing theoretical frameworks on life-span career development, including the life-career theory (Super, 1980), the social cognitive career theory (Lent, Brown, & Hackett, 1994), the expectancy-value model of achievement (Eccles, 2011), the developmental-contextual model of career development (Porfeli & Vondracek, 2009), and the career construction theory (Savickas, 2002), I proposed an integrative model shown in Figure 1.1 to illustrate paths connecting the family system, youth career development processes, and career attainment in young adulthood. The dissertation’s three studies addressed paths in this model that were highlighted in theory but relatively less studied in empirical research, that is, paths c, d, and f. Study 1 provided empirical evidence for path c, in combination with path b, by showing that adolescents’ career adaptivity (a career development process) mediated the longitudinal link between parent-adolescent relationship quality (a family process) and young adult occupational prestige (an aspect of career attainment). Study 2 supported path d by documenting implications of youth work experiences (a component of youth career development) for their relationship quality with fathers (a family process). Study 3 took an innovative approach to model and reveal complexities in the multi-dimensional, multi-component family system processes and resources surrounding adolescents that together predict their educational attainment in young adulthood, an important step towards career attainment, thereby findings contributed to a better understanding of paths f and b. Together, these three studies contribute to the theoretical understanding of family and youth career development, advance methodologies applied in this research field, and provide important practical implications. Findings from these three studies also suggest future directions toward a better understanding of the interplay between family and youth career
Theoretical Contributions

By examining paths in the theoretical model (Figure 1.1) that have been understudied in empirical research but highlighted across theories, findings of this dissertation provide novel empirical evidence and advance understanding of the interplay between family systems and youth career development processes and outcomes. In particular, Study 1 illuminated the combined implications of paths b and c to show that, career adaptivity in adolescence serves as a mechanism linking adolescents’ family relationships, over time, to career attainment in young adulthood. The study’s focus on career adaptivity provided support for the career construction theory, which highlights career adaptivity as a fundamental component in youth career development that is subject to family influences and important to future career attainment (Savickas, 2002; 2013). Beyond providing empirical support, this study also informed revisions and advancements that can be made to the career construction theory and research testing this theory. First, the term ‘career adaptivity’ that has been widely used in the literature was originally proposed as ‘career adaptive readiness’, and the inconsistency in the terminology usage may hinder understanding and examinations of the theoretical model for career adaptation. Second, the lack of clear operationalization of career adaptivity in the theory may have constrained research developing measurement model for this construct and led previous research to use it as a generic umbrella concept, as compared to the other three career adaptation components for which valid and reliable measures have been developed (Savickas & Porfeli, 2012; Savickas, Porfeli, Hilton, & Savickas, 2018). This calls for an expansion of the theory that clearly defines and operationalizes career adaptivity, the fundamental component in the career adaptation process. Third, to better inform the role of family in adolescent career adaptation and development.
guide further research, an important next step improving the theory is to develop tenets pertaining to gendered patterns in family influences, especially what roles that mothers and fathers would play in boys’ and girls’ career adaptation. More broadly, testing the hypothesized mediation model (Figure 2.1) also provided empirical evidence for a variety of theoretical frameworks on career development that highlight family influences on future career attainment through youth career development processes, such as the social cognitive career theory (Lent et al., 1994) and the expectancy-value model of achievement (Eccles, 2011).

Study 2’s findings about longitudinal implications of youth work experiences for father-youth relationship quality provided empirical support for bi-directionality in person-context interactions that has been emphasized across theories, such as the developmental-contextual model of career development (Porfeli & Vondracek, 2009). This study extended prior research that has focused on family influences on youth career development but neglected to examine the reverse direction of effect, that is, youth career development processes in relation to family dynamics (path d). Moreover, results of this study illuminated complexities in this youth work-to-family linkage, including curvilinear patterns and interactions with youth gender and mothers’ employment status.

Turning to Study 3, by training and testing comprehensive prediction models with 53 adolescent family experience variables as predictors of college enrollment and graduation about 13 years later, identified key family experiences in adolescence to young adult educational attainment—an important step in career development towards career attainment. That is, the findings provided strong support for path b and theoretical tenets highlighting the role of family in youth career development, especially its long-term implications (Eccles, 2011; Savickas, 2002; Super, 1980). Further, by identifying key predictors out of the wide array of family
experience variables, this study highlighted processes and resources in the family system that may be especially important in shaping youth career development and achievement. As such, this study provided novel insights for career development theoretical frameworks wherein the relative contributions of different family experiences have not been a focus of study. Moreover, by illuminating complex patterns, including nonlinearities and interactions among predictors (i.e., path f), this study provided evidence for theoretical frameworks, especially the developmental-contextual model of career development (Porfeli & Vondracek, 2009), which highlight systemic underpinnings of contexts for career development, including complex dynamics among multiple components and sub-systems of families.

Beyond contributions to the integrative model in Figure 1.1 and career development theories, findings from these studies also inform general theoretical frameworks of human development and family dynamics, including the ecological model (Bronfenbrenner & Morris, 2006) and the family systems perspective (Minuchin, 1985). With regard to the ecological model, Study 2 illuminated the work-family mesosystem, with findings indicating exchanges between these two contexts, by means of a novel focus on youth work experiences, and Study 3 revealed the nonlinearities and interactions of processes and resources within the family contexts of individual development as highlighted in this model (Bronfenbrenner & Morris, 2006). With respect to the family systems perspective (Minuchin, 1985), Study 3 revealed nonlinearities and interactions within the family system—which are often discussed but less frequently examined in empirical research. Further, Study 2 targeted youth work experiences as an important context for family system functioning and highlighted interdependencies among family members and processes by examining the role of maternal employment in youth work effects on father-youth relationship quality.
Methodological Strengths and Advancements

Strengths and contributions of this dissertation also lie in the study methodologies, including sampling and data usage, measurement, and statistical analyses. With regard to data, all three studies used longitudinal data, and in particular, Studies 1 and 3 used data spanning across adolescence and young adulthood to study long-term implications of family experiences for occupational and educational attainments. Further, sampling methods matched the study goals: For example, Study 2 applied an ethnic homogeneous approach that allowed examining within-group variability among Mexican-origin youth and families, whereas Study 3 took the advantage of a large-scale, secondary dataset from a U.S. nationally representative sample with racial/ethnic diversity to build comprehensive and generalizable prediction models.

With respect to measurement, in both Studies 1 and 2, I used reports from both youth and the parent to study their dyadic relationship quality: Study 1 used latent variables in the structural equation model to capture the shared variance of youth’s and parents’ reports, and Study 2 used multivariate multilevel models wherein both youth’s and fathers’ reports were examined simultaneously. Study 3, as a cross-study synthesis, grounded the selection and measurements of all study variables in published studies. Further, Study 1, by showing that career adaptivity can be captured as a latent construct with observed components, extended extant literature on career adaptation processes that has predominantly treated career adaptivity as an umbrella concept with its components studied separately.

With regard to statistical analyses, Study 3, in particular contributed to the literature on family and career development by taking an innovative approach to synthesize previous studies with machine learning. Machine learning has recently garnered interest among developmental scientists given its capacity to use large-scale datasets, make predictions, and discover complex
patterns in development (Rosenberg, Casey, & Holmes, 2018). This study showcased the utility of machine learning in analytically synthesizing previous studies that used large-scale, secondary datasets, through which developmental scientists can build the next layer of knowledge and inform next steps to take for future research.

**Practical Implications**

Although this dissertation was not designed to develop policy or intervention/prevention programs, findings from the three studies have implications for practice. By illuminating the importance of family in youth career development processes and attainments, the study findings suggested potential feasibility and efficacy of family-based programs for promoting youth career achievements. In particular, findings from Study 3 provided models, using information about adolescents’ family experiences, to forecast their future educational attainment, and these findings can be used in practice to identify at-risk youth as targets. Also, the family experience factors identified as key predictors in this study can inform cost-effective interventions to target at important modifiable factors. Results of Study 1 suggested that adolescent career adaptivity can be an important proximal outcome to evaluate for programs promoting long-term career attainment. Further, Study 2, by showing implications of youth work experiences for their family relationships, suggested a feedback loop from career development processes to the family life. This loop can potentially be leveraged for establishing upward spirals that promote and sustain youth career development outcomes and family well-being (Fredrickson, 2013).

**Future Directions**

Based on the study findings, this dissertation directs attention to future directions to better understand the role of family in youth career development processes and attainments, beyond the directions that have been discussed separately in each chapter. Although the study findings, in
combination with extant literature, contribute to a better understanding of the integrative theoretical model in Figure 1.1 for family and career development, this model still needs further examination such as through testing additional dimensions of both youth career development and family systems.

One important direction is to examine the interplay among family factors and multiple dimensions of youth career development processes. For example, beyond adolescent career adaptivity, other components of the career adaptation process such as adaptability resources, adapting responses, and adaptation results also may mediate the effects of family experiences on future career attainment. In addition, dimensions of youth’s work experiences, such as autonomy and work pressure also may spill over to influence their family relationships, according to findings from the work-family literature among ethnic minority adults (Sun, McHale, Crouter, & Jones, 2017).

Also important is to extend current study findings, especially those from Studies 1 and 2, is to examine other aspects of family experiences. Studies 1 and 2 focused on parent-youth relationship quality and its role in career development, but as revealed in Study 3, family experiences include multiple dimensions and components beyond relationship quality. The important family processes identified in Study 3, such as parent educational expectations and family meals, may provide guidance for models from Studies 1 and 2 to test in future research.

Finally, in future research, attainment outcomes should be evaluated in terms of occupational prestige as well as work and career dimensions such as income and subjective career satisfaction (Ng, Eby, Sorensen, & Feldman, 2005). And, moving beyond Study 3’s focus on educational attainment, comprehensive prediction models based on adolescent family experiences need to be built for these other career development outcomes.
Conclusions

Taken together, this dissertation built on the theory of life-span career development to advance our understanding of the role of family in youth career development, especially by illuminating mechanisms that have been understudied in previous research. Study 1 contributed to the understanding of career development processes, such as adolescent career adaptivity, as mechanisms underlying (i.e., mediating) long-term family influences on career attainment. Study 2 targeted directional links that have been neglected in prior research but highlighted in theory—that is, from youth career development processes to family processes—by examining implications of youth work experiences for their relationship quality with fathers. Study 3 took an innovative approach to highlight the importance of adolescent family experiences to youth career development and reveal family system complexities in predicting future attainment. Overall, this dissertation provided evidence for the theoretical model in illuminating the mechanisms linking the family system and youth career development processes and career attainment, introduced innovative methods to research on family and career development, provided implications for family-based practice promoting youth achievement, and directed attention to future research efforts that can better our understanding of the mechanisms underlying family systems influences on youth career development.
References


VITA
Xiaoran Sun

Education
2019 Ph.D., Human Development and Family Studies, The Pennsylvania State University
2019 Graduate Minor, Social Data Analytics, The Pennsylvania State University
2016 M.S., Human Development and Family Studies, The Pennsylvania State University
2015 B.S., Psychology, Zhejiang University

Selected Traineeship and Awards
2017–2018 Traineeship, Big Data Social Science Training Program, National Science Foundation–IGERT (Grant DGE-1144860)
2015–2016 University Graduate Fellowship, The Pennsylvania State University
2015–2018 Morningside Cultural China Scholarship for Alumni Studying Abroad, Zhejiang University
2015 Outstanding Graduate of Zhejiang Province, Ministry of Education, China

Selected Research Experience
2018–2019 Project Coordinator to Innovative Family Prevention with Latino Siblings in Disadvantaged Settings, funded by NICHD (PIs: Kimberly A. Updegraff, Ph.D. & Adriana J. Umaña-Taylor, Ph.D.; Co-Is: Susan M. McHale, Ph.D., & Mark E. Feinberg, Ph.D., Arizona State University & The Pennsylvania State University)
2017–2018 Project Leader to Mining Sibling Interactions on Twitter, funded by NSF-IGERT

Teaching Experience
Summer 2019 Instructor, HDFS 129: Introduction to Human Development & Family Studies
Spring 2019 Teaching Assistant, HDFS 129: Introduction to Human Development & Family Studies
Spring 2017 Guest Lecturer, “Family as a context of young adult career development”, HDFS 433
Spring 2017 Teaching Assistant, HDFS 249: Adult Development and Aging

Publications


