PATHWAYS TO PRIMARY CARE

A Dissertation

Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

Daniel Marlowe Read, B.A., Ed.M.

September 2015

ABSTRACT

The composition of the physician workforce has immediate implications for the quality of accessibility of healthcare services in the US, and there is a growing concern that both the number of physicians trained and the specialties that they choose will not satisfy the needs of society. The production of primary care physicians is particularly worrisome, and it is suspected the pre-residency environment engenders negative perceptions of primary care which contributes to this problem. The purpose of this dissertation was to inspect the differences between groups of students who indicated a primary care preference, a non-primary care preference, or were undecided at matriculation into and graduation from an LCME-accredited medical institution. Using data provided by the Association of American Medical Colleges (AAMC), 31,425 medical graduates who took the Post-MCAT Questionnaire (PMQ) between 2001 and 2006 provided data for these analyses. Using undecided students as the referent group in multinomial logistic regression models, we found that a student's gender, pre-matriculation experience with science or medicine, and reasons for pursuing medicine as a career all strongly distinguished the three outcome groups. One's prior specialty preference also strongly predicted their specialty preference at graduation. Analyzing the data through a socialization framework, it was concluded that experiences that allow students to "try on" the physician role are particularly powerful motivators. For institutions with a primary care focus, these data may help admission and recruiting teams identify the characteristics of students who are likely to enter and graduate with a primary care orientation. Alternatively, these findings may help institutions develop structures that will persuade students to consider a primary care specialty. Collectively, these strategies can be used to intentionally design medical student experiences so that students make optimal career decisions that satisfy the needs of society.

CISE Department Curry School of Education University of Virginia Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

This dissertation, ("Pathways to Primary Care"), has been approved by the Graduate Faculty of the Curry School of Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Major Advisor, Robert Tai

Committee Member, Susan Pollart

Committee Member, Ji Hoon Ryoo

Committee Member, Frackson Mumba

Committee Member, Christine Liu

Date

DEDICATION

This dissertation is dedicated to all of my own communities of practice that have supported me for my thirty years. Thank you all.

ACKNOWLEDGEMENTS

Writing a dissertation, I have found, can be very lonesome and isolating. That is, until you realize that you really aren't doing it by yourself at all. This is at once a comforting and humbling feeling.

My family and friends deserve as much credit as I do. Anything I have achieved I attribute to their influence on me. You have all been just fantastic, especially my sister who was my peer editor.

I also think of my DPS team. Where would I be without your support and guidance?

There are so many lovely people in the lab, at UVA, and in Charlottesville. It really has been a wonderful place to spend a few years.

To my committee members: when I asked for your help, you all joined with enthusiasm and you have inspired me to raise the level of my work. This could be said of Donna Jeffe as well.

Finally, Robert Tai made all of this happen. Let's not forget that. ©

Daniel Marlowe Read 25, August 2015

TABLE OF CONTENTS

	PAGE
LIST OF TABLES	Iii
LIST OF FIGURES	Iv
CHAPTER 1: INTRODUCTION	1
Chapter Overview	1
Presentation of the Issue	2
The Historical Context and a Growing Shortage	2
Three Sources of Demand	6
A Restricted Supply	8
Concluding Remarks	12
The Current Study	13
Study Rationale	13
Original Contributions	16
CHAPTER 2: REVIEW OF THE LITERATURE	19
Chapter Overview	19
Guiding Frames	20
Primary Care Definitions	20
Outcome Operationalization and Definitions	23
Review of Existing Theoretical Frameworks	24
An Integrative Theoretical Framework	27
A Brief Review of Seminal Investigations	40
Intentions or Preferences at Graduation	41
Intentions or Preferences at Matriculation	53
Review of the Extant Literature	55
Major Limitations of the Literature	55
Synthesis and Conclusions of Relevant Predictors	58
Placing this Study	66
CHAPTER 3: DATA SOURCES AND METHODOLOGY	70
Chapter Overview	70
Data Sources	71
SRS (Student Records System)	72
PMQ (Post-MCAT Questionnaire)	73
MSQ (Matriculating Student Questionnaire)	73
GQ (Graduation Questionnaire)	74
Data Screening and Study Respondents	74
Outcome Variables	78
Specialty Preference at Matriculation	78

Specialty Preference at Graduation
Predictor Variables Predictive of Specialty Choice
Research Strategies
RQ #1: Motivational Factors
RQ #2: Specialty Preferences at Matriculation
RQ #3: Specialty Preferences at Graduation
CHAPTER 4: RESULTS
Chapter Overview
Descriptive Statistics
Matriculation Specialty Preference
Graduation Specialty Preference
RQ #1: Factor Analysis of Motivational Factors
Exploratory Factor Analysis Results
Confirmatory Factor Analysis Results
Factor Scores
RO #2: Matriculating Specialty Preferences
RO #3. Graduating Specialty Preferences
CHAPTER 5: DISCUSSION
Chapter Overview
Tools for Synthesizing the Results
Pathways in Preferences
A Second Pseudo-Longitudinal Approach
"Splitting the Difference"
Conclusions
Motivational Factors Conclusions
Multinomial Logistic Regression Conclusions
I imitations
Implications
Implications for Research
Implications for Policy
Implications for Practice
Dersonal Deflection
DEEEDENICES
ADDENIDIV A. SVNITAV
AFFENDIA A: 51 NTAA

LIST OF TABLES

TABLE		PAC
2.1	List of Seminal Studies and Timing of Specialty Preference or Plans Measurement	69
3.1	Cross-Tabulation of Race/Ethnicities and Racial/Ethnic Categories	107
3.2	Battery of Reasons for Pursuing a Medical Career from the MSQ	108
3.3	Cross-Tabulation of Parental Occupation	109
3.4	Cross-Tabulation of Professional Socialization Experiences	110
4.1	By Matriculating Specialty Preference, Student-Level Characteristics.	136
4.2	By Graduating Specialty Preference, Student-Level Characteristics	137
4.3	Descriptive Statistics of 22 Motivational Items	138
4.4	Exploratory Factor Analysis of 18 Motivational Items	139
4.5	Confirmatory Factor Analysis Goodness-of-Fit Statistics	140
4.6	Confirmatory Factor Analysis of Five Motivational Factors, Structural	
	Model	141
4.7	Confirmatory Factor Analysis of Five Motivational Factors,	
	Measurement Model	142
4.8	Factors Score Means and Standard Deviations for All Five	
	"Motivational Factors" by Matriculating Specialty Preference	145
4.9	ANOVA A Priori Contrast Results, Showing the F-Value for All Five	
	"Motivational Factors"	146
4.10 (A)	Multinomial Logistic Regression Results Comparing the Matriculating	
	Specialty Preference Option of Primary Care versus the Base	
	Outcome of Undecided	147
4.10 (B)	Multinomial Logistic Regression Results Comparing the Matriculating	
	Specialty Preference Option of Non-Primary Care versus the Base	
	Outcome of Undecided	148
4.11 (A)	Multinomial Logistic Regression Results Comparing the Graduating	
	Specialty Preference Option of Primary Care versus the Base	
	Outcome of Undecided	149
4.11 (B)	Multinomial Logistic Regression Results Comparing the Matriculating	
	Specialty Preference Option of Non-Primary Care versus the Base	
	Outcome of Undecided	150
5.1	Summary of the Multinomial Logistic Regression Results	195

LIST OF FIGURES

FIGURE

PAGE

1.1	Projected supply and demand for physicians from 2010 to 2025, showing workforce shortages of primary care physicians and non- primary care physicians
2.1	Bennett-Philips conceptual model of the process of primary care specialty choice.
2.2	Conceptual model of specialty plan-making among medical students
3.1	Data-screening process to identify the study sample and delimit the study sample
3.2	Survey question #15 and #15a from the 2005 MSQ asking students' plans about specialization
3.3	Questions from the 2014 GQ regarding students' specialization plans
4.1	Modified factor model
5.1	Flowchart of medical students with a primary care specialty preference at matriculation
5.2	Flowchart of medical students with undecided specialty preferences at matriculation
5.3	Flowchart of medical students with a non-primary care specialty preference at matriculation
5.4	Conceptual chart showing typical outcome specification shown with the solid line, and the outcome specification of this analysis shown with the dashed lines

CHAPTER 1: INTRODUCTION

Chapter Overview

While reading the *Twentieth Report of COGME*, I encountered the following curious passage:

"Students' future career choices are strongly shaped during medical school. While many students express interest in primary care when they first enter medical school, this interest may erode by the time they choose their graduate medical education specialty in their fourth year of training." (COGME, 2010, p7.)

It struck me for three reasons. First, it is to my knowledge the most succinct and articulate summary of the issue at hand in this dissertation. Second, it conveys a number of important points. For example, it correctly notes that the process of choosing a medical specialty is rooted in the ecology of the medical school environment, while simultaneously demonstrating that interest in primary care careers, in fact, diminishes while in this environment. Lastly, this passage struck me for the number of questions that it raises: *Why do medical students have interest in primary care at matriculation? What bodies of knowledge are available to them in the decision-making process? How can we assess and build policies to reduce this attrition?* For this researcher, this topic is as exciting as it is challenging. The goal of this chapter is a simple one: I aim to demonstrate that the historic trends of medical students moving away from primary care careers

warrants an investigation into how and why students prefer or choose primary care careers.

I first show that in the U.S. context, the balance of primary care physicians has been seriously recognized as problematic since the 1970s. I then discuss the current status of the shortage by demonstrating that the magnitude of the shortage of primary care physicians, specifically generalists, has undisputable consequences for the wellbeing of the U.S. healthcare system. I follow this presentation of the issue with the dissertation's rationale, its research questions, and its potential contributions.

Presentation of the Issue

The Historical Context and a Growing Shortage

Although this dissertation focuses on the current status of the primary care physician shortage, an awareness of the issue's legacy is essential. While a robust debate around workforce development had arisen as early as the 1920's (Ricketts, 2011), more serious approaches to understanding this workforce balance began as the number of subspecialties began to increase towards the middle of the twentieth century. Using models such as the physician supply model (PSM) or the physician requirements model (PRM), it became accepted that a glut of specialists would be accompanied by a shortage of primary care physicians (US Congress, Office of Technology Assessment, 1988; Council on Graduate Medical Education [COGME], 1992; Weiner, 1994; Greenberg & Cultice, 1997).

Moving toward the present, a workforce analysis by Cooper, Getzen, McKee & Laud (2002) projected a deficit of 200,000 physicians by the year 2020, with the generalists likely to be heavily and negatively impacted. While the importance of their conclusions is undeniable, it was their macroeconomic approach that proved to be seminal for two very important reasons. First, it set the tone for how the medical workforce would be analyzed by introducing the "trend model" which considered economic expansion, population growth, effort of physicians, and finally the related efforts of non-physicians. (p. 142). As I show in the next section, this approach is nearly universally applied in current workforce analyses. A second reason for this article's importance is that it provided definitive support for the position that there would be a future workforce *shortage* rather than *surplus*, more or less ending that debate.

In 2006, the Association of American Medical Colleges' *Statement on the Physician Workforce* anticipated an imminent physician shortage of physicians to service the U.S. population, concluding that "a shortage of physicians would undeniably make access to care more problematic for all citizens" (AAMC, 2006, p. 1). It appears that—at least in part—this prediction has been realized. Indeed, the U.S. healthcare system now faces an historic shortage of physicians to support a growing and aging population, and it is anticipated that this shortage will worsen in the foreseeable future, perhaps leading to an "emerging time bomb" (Frisch, 2013, p. 1). A recent review of state-level reports reveals similarly ominous forecasts such as "Extant Physician Shortage will Become More Severe," "Impending Physician Shortage in the State," and "Looming Shortage of Physicians" (Center for Workforce Studies, 2012).

At the national level, various methodologies have been used to estimate the magnitude of the predicted physician shortage (e.g. Colwill, Cultice, & Kruse, 2008; Dill & Salsberg, 2008; Center for Workforce Studies, 2010; HRSA, 2013) A commonly cited figure from the AAMC's *Center for Workforce Studies* is the projected shortage of 124,000 total physicians in the U.S. by 2025 (Dill & Salsberg, 2008). Using the Physician Supply and Demand Model (PSDM), these researchers found that between 2006 and 2025, the demand for physicians is projected to increase by about 8% while the per capita supply of physicians is projected to *decrease* by about 8% (from 228 to 210 physicians per 100,000 persons) in this same time period (p. 28). Alarmingly, when disaggregated by specialty group, their projections show that 37% of the physician shortage is attributable to primary care physicians alone, indicating a total shortage of 46,000 primary care physicians by 2025 (p. 26).

Preliminary updates of these figures as of 2010 painted an even more distressing portrait of this issue (Center for Workforce Studies, 2010). Taking into consideration physician utilization, working hours, new entrants to the physician workforce, and obesity figures, the revised figures showed a total physician shortage of 130,600 by 2025, of which 65,800 will be attributable to primary care physicians alone. (*Figure 1.1* displays these findings.) However, the most recent updates undertaken by IHS. found that

Onational demand for physicians outpaced the supply, but concluded that the 2010 figures mostly likely overstated the physician shortage, and significantly (2015). For example, IHS determined the primary care physician shortage to be between 12,500 and 31,100 by 2025, a non-trivial adjustment to this important figure. They accounted for this large discrepancy by citing revised implementation models of the ACA (Patient Protection and Affordable Care Act), by modeling to account for the workforce production of nurse practioners and physicians assistants, and by accounting for "retail clinics" which provide services that primary care visits would normal deliver. However, some scenarios represented in this update rely on drastic transformations of primary care service delivery.

While the AAMC is well placed to execute such analysis, other groups and researchers have also rigorously investigated this issue. For example, Colwell, Cultice, and Kruse (2008), using a baseline year of 2005, provide an estimated shortage of 35,000 to 44,000 generalists for adult patients. While they estimate a stable—but increasing—demand for generalist services, they note that increasing trends away from primary care residencies as well as subspecialization *within* primary care specialties suggests that a continual increase in the supply of generalists is an untenable supposition. Using an altogether different methodology to assess the primary care physician shortage, figures provided by the Health Resources and Services Administration (2013) suggest a smaller overall shortage of 20,400 primary care physicians by the year 2020, about 50% of the Center for Workforce Studies' estimated shortage for that same year. The difference between these estimates may be explained by the HRSA assumes yearly growth in the

supply of primary care physicians, or by the HRSA's categorization of geriatrics as a primary care specialty. Making sense of this nontrivial issue, Thomas Ricketts correctly notes that "our best way forward is to tighten our definitions and improve our data to have at least a good sense of where we are now with supply" (2011, p. 426). Thus, while there is some explainable reason for variation regarding the *magnitude* of the primary care physician shortage, there is little doubt that one exists and that the shortage is increasing. The next two sections detail more specifically its root causes.

Three Sources of Demand

How can we best make sense of this shortage and its root causes? Economic theory suggests that any shortage is, in its most elementary form, an imbalance between the supply and demand. Looking first at the demand for primary care services, researchers generally agree that population growth and aging will be primary stimulus of the demand for physicians. In terms of population growth, the U.S. Census Bureau estimates a 7.8% population increase from 2015 to 2025 while the proportion of the population 65 and older will also increase (2012). This is particularly problematic as older Americans utilize healthcare services at a significantly higher rate than younger Americans (Institute of Medicine, 2008, p.1, Health Resources and Services Administration [HRSA], 2008). Dill and Salsberg (2008) also concur, estimating a 26.3% growth in demand for physicians from 2006 to 2025, and that "most of this projected demand increase is attributable to the projected growth and aging of the population" (p. 20). This finding is consistent with other studies of workforce needs. One found that the U.S. healthcare system will require 52,000 *additional* primary care

physicians by 2025, the majority of which (33,000) will be due to population growth, while population aging will necessitate the training of 10,000 additional primary care physicians (Petterson, et al., 2012) while another reported that "aging and population growth are projected to account for 81 percent of the change in demand between 2010 and 2020" (HRSA, 2013, p. 1).

These healthcare demands are further problematized by recent legislation aimed at increasing the number of those with health insurance (Patient Protection and Affordable Care Act [ACA], 2010), through which an estimated 30-34 million Americans will gain access to healthcare insurance. These newly-insured individuals will be more likely to utilize the healthcare system than those without insurance (Buchmueller, Grumbach, Kronick, & Kahn, 2005; Freeman, Kadiyala, Bell, & Martin, 2008), signaling a greater need for physicians. However, this increase in utilization is predicated on the assumption that physicians are available and willing to accept new patients, which may not be a tenable assumption in the current market (Massachusetts Medical Society, 2013). That said, it is generally assumed that the increase in those with insurance will place further pressure on the healthcare system to provide adequate coverage, and that primary care physicians will experience a substantial increase in demand as they are positioned as the point of entry into the healthcare system (Goodson, 2010).

Although estimates vary, it is generally agreed that universal insurance coverage will increase the demand for physician service. One report estimates a 4% increase in demand for *all* physicians (Dill & Salsberg, 2008) and another by the Bureau of Health

Professions estimates a 5% increase in demand for all physicians (Health Resources and Services Administration, 2006), though it is possible that these estimates may not generalize to the primary care professions in particular. Therefore, looking only at primary care physicians provides more credible estimates of how increased coverage will impact demand. To that end, Hofer, Abraham, and Moscovice found that expanded coverage will necessitate approximately 4,300 to 7,000 additional primary care physicians to meet that demand by 2019 (2011), while Petterson et al. estimated an increased demand of 8,000 primary care physicians (3%) by 2025 (2012). Thus, while the expansion of insurance is germane to discussions about primary care physician demand, it is the aging and growth of the U.S. population that will primarily drive demand for their services in the coming decades. In turn, it is important to recognize that the geographical maldistribution of primary care physicians (for example in rural communities) is likely to further exacerbate the issue of physician availability in high needs communities (Petterson, Philips, Bazemore, & Koinis, 2013).

A Restricted Supply

The supply of primary care physicians also merits discussion. The AAMC's 2006 *Statement on the Physician Workforce* posited that increasing LCME-accredited U.S. medical school capacity and enrolment by 30% from 2003 to 2015 would help ameliorate the anticipated physician workforce shortage, and a recent survey suggests that this target will be met by 2018 (Center for Workforce Studies, 2014). And yet, it has become clear that the supply of primary care physicians will not substantially increase in the foreseeable future. As Bodenheimer and Pham state, "primary care faces a growing crisis,

in part because increasing numbers of U.S. medical graduates are avoiding careers in adult primary care" (2010, p. 799).

The authors of a recent study on preferences for a career in primary care come to a similar conclusion, finding that while about 60% of graduates from LCME-accredited, U.S.-based institutions were planning a career in primary care in 1997, only about 30% were making these plans by 2008 (Jeffe, Whelan, and Andriole, 2010, p. 948). Similarly, Lambert and Holmboe (2007) used the definition of controllable lifestyle specialties from Schwartz et al. (1990) and found that the proportion of graduates preferring "uncontrollable specialties"—most of which are in primary care—significantly decreased from 1990 to 2003. Moreover, suspicions that medical students lose interest in primary care careers because of medical school climates hostile to primary care careers has also been documented (Compton, Frank, Elon, & Carrera, 2008). Thus, it appears that many potential primary care physicians may be receiving signals throughout their medical school experience that primary care is a poorly perceived career choice (Erickson, Danish, Jones, Sandberg, & Carle, 2013).

While plans and preferences are important indicators of future career decisions, it is the Graduate Medical Education (GME) system that plays a more crucial role in determining the composition of the physician workforce. Many recent analyses and reports have likened the GME to a "bottleneck" as qualified medical school graduates are unable to match with a residency position that would allow them to acquire professional certification. For example, Jolly, Erickson, and Garrison state that "although both MD- granting and DO-granting medical schools have raised enrollments substantially, these changes cannot increase the number of practicing physicians without concomitant growth in the number of residency positions available for GME in teaching hospital" (2013, p. 468). Moreover, Dower and O'Neil, writing on behalf of the *Robert Wood Johnson Foundation*, note that "because of a ceiling on Medicare-funded graduate medical education residencies, any increase in U.S. medical school graduates will merely displace international medical school graduates in residency programs" (2011, p. 6).

Most locate the cause of this bottleneck in the federal funding restrictions mandated by the Balanced Budget Act of 1997, which in practice limited the number of federally funded residency positions to around 26,000 (Salsberg, Rockey, Rivers, Brotherton, & Jackson, 2008), and is of concern for medical schools who see their students struggle to find residency positions. (Center for Workforce Studies, 2014). Though some small growth (13.6%) in the number of residency positions has been reported from 2001 to 2010 (Jolly, Erickson, & Garrison, 2012, p. 470), there exists a competitive "matching" process where M.D. graduates contend with D.O. and other qualified individuals for the same number of clinical training opportunities. According to a recent report by the National Resident Matching Program (NRMP), the 2014 "match" saw 34,270 applicants and 26,678 Post-Graduate Year One (PGY-1) residency positions (NRMP, 2014). Importantly, when disaggregated by specialty, it is clear that primary care and non-primary care specialties have both reached nearly their maximum capacity. In the 2014 PGY-1 Match for example, family medicine filled 95.8% of its offered positions, internal medicine filled at 99.1%, pediatrics at 99.5%, while all PGY-1 offered

positions filled at 96.3% (p. 4). Thus, programs, policies, and incentives to increase the number of residency positions—such as the Primary Care Residency Expansion (PCRE) program and the Creating Access to Residency Education (CARE) act—are well placed to address the primary care supply bottleneck (Chen, Petterson, & Bazemore, 2014; AMA, 2014a).

Although these residency programs are at capacity, it is also important to note the unique subspecialization trends for each specialty. While most family medicine graduates continue into generalist careers, a significantly greater percentage of pediatric and internal medicine residents subspecialize and ostensibly—but not definitely—exit the primary care pipeline. The American Medical Association's (AMA) online system for tracking resident's career plans, FREIDA, shows stark differences between residents in these primary care specialties for the graduating year 2013. For family medicine residents, program directors indicated that 13.5% were planning on further training, while for pediatric and internal medicine residents, they indicated that 44.9% and 47.9% were planning on further training, respectively (AMA, 2014b), and Colwill, Cultice, and Kruse found an even sharper decline in interest amongst internist residents (2008). They found that 52% of internal medicine graduates had subspecialization plans in 1995 compared to 62% just ten years later, while Jolly, Erickson, and Garrison found that about 57% of first year internal medicine residents in 2010 would subspecialize (2012). Demonstrating the influence of GME's institutional mission, the estimates from the study by Chen and colleagues (2013) show some institutions retain less than 10% of their internal medicine residents in primary care fields. Moreover, the American Board of Internal Medicine

(ABIM) has recently added and is planning on adding new subspecialties (Cassel & Reuben, 2011), thus, it is reasonable to conclude that this will further decrease the proportion of primary care generalists from internal medicine.

As a result, while fewer and fewer medical students are planning careers in primary care during medical schools, the number of *actual* generalist residents has similarly decreased since 1998 (Colwill, Cultice, & Kruse, 2008, p. w232). Indeed, a recent study found that of residents who graduated between the years 2006-2008 only 25.2% graduated in primary care, and the authors estimate that even this figure may be an overestimation (Chen, Petterson, Phillips, Mullan, Bazemore, & ODonnell, 2013).

Concluding Remarks

In the nexus of these discouraging supply trends, it is clear that the U.S. physician training programs will not be able to sustain the workforce or ameliorate the anticipated increase in demand for primary care physician services. Recent comments made by the U.S. Secretary of Veterans Affairs Robert McDonald, suggest that these shortages are already problematizing health care delivery:

We are the canary in the coal mine. I was in Florida. I visited the University of Central Florida medical school, University of South Florida medical school. The people there told me Florida needs 17,000 more doctors. I was in California. Janet Napolitano, who leads the University of California, she said that California needs 22,000 more doctors. So we are demonstrating to the American public the problems that the American medical system has seen. We're just not producing enough doctors. (Siegel, 2014)

As medical workforce alignment with national needs remains problematic in the U.S., understanding why, how, and under which circumstances medical students choose

their specialty or sub-specialty has relevance for healthcare policy discussions. At the heart of this issue is the training and distribution of primary care physicians who are able to serve as frontline staff. This is especially important as primary health care systems are well placed to reduce healthcare costs and disparities, resulting in higher quality of patient care (Starfield, Shi, and Macinko 2005). As primary care physicians play a vital role in providing healthcare services at a relatively low cost, there is widespread consensus that addressing this immense workforce imbalance is an urgent task (Bodenheimer, Chen, & Bennett, 2009). Or, as Colwill, Cultice, and Kruse succinctly summarize the issue, "our goal should be to maintain or preferably increase today's availability of generalists. It is in the best interest of the public and of all physicians" (2008, p. w 239). Recommendations of great diversity for addressing this issue remain including restructuring Medicare's support for graduate training positions (Voorhees, Prado-Gutierrez, Epperly, & Derksen, 2013), shifting primary care services to other potential members of the medical community such as nurse practioners (Laurant et al., 2014), requiring GME programs to monitor residency trends and workforce adequacy (Josiah Macy Jr. Foundation, 2011), and redesigning the monetary and non-monetary incentives and motivations for preferring, choosing, and remaining in the primary care field (Dorsey, Jarjoura, & Rutecki, 2005; Dezee et al., 2011).

The Current Study

Study Rationale

This study looks at early career preferences amongst medical students. As such, the purpose of this study is to examine the matriculating factors that are associated with specialty preferences for a career in primary care, both at matriculation into medical school and at graduation from medical school. In this dissertation, these factors will collectively be known as a "pre-medical identity," a collection of three kinds of psychosocial, person-level variables.

First, the demographic profile of the respondent is considered. Second, I consider some of the pre-matriculation experiences that may have had a socializing influence on their interest in medicine. Third, I investigate a set of factors related to the respondent's motivational factors for choosing medicine as a career. Investigating these factors will help explain why students choose careers in primary care and may have policy implications. For example, conclusions from this study may aid pre-medical programs be it for high school or college age students— in better supporting interest in primary care careers. It may also present trends to medical school admissions teams who are seeking to recruit, support, and retain their students in the primary care pipeline.

To do this, I first created and validated a set of premedical factors using factor analysis. I then created a matriculation model and a graduation model and applied a confirmatory multinomial logistic regression to predict a three-category outcome (primary care preference, non-primary care preference, and an undecided specialty preference). For this dissertation, I recognize family medicine (including subspecialties), pediatrics (no subspecialties), and internal medicine (no subspecialties) as primary care specialties. This is consistent with the AAMC's stance, and further explanation of this rationale is provided in Chapter 3, where I discuss the outcome variables. All respondents who indicated that they had other specialty plans were categorized as non-primary care. Finally, those who indicated that they had no specialization plan or did not respond to the question about their career plans were labeled as undecided.

The main analytic output for this dissertation is independent odds ratios, which allow for an interpretation of the association between a single predictor and the outcome. Because a great deal of literature (reviewed here in Chapter 3) has already compared those with identifiable specialty preferences, the referent group in these models are those with an undecided specialty preference. This allows for an examination of how students with uncertain career plans compare to those who have already made future career plans. Comparing these groups can provide an indication of which particular experiences can be leveraged to support alignment between pre-medical programs and workforce needs. With this research agenda in mind, the specific research questions for this dissertation were:

What are the career and lifestyle motivations reported by matriculating medical students? And, are there differences between those with a matriculating preference for primary care specialty, those without a specialty preference, and those with a preference for a non-primary care specialty?

- Do these career and lifestyle motivations and other psycho-social factors differentiate those with a preference for a primary care specialty, those with an unspecified specialty preference, and those with a preference for a non-primary care specialty *at matriculation into* medical school?
- Do these career and lifestyle motivations, other psycho-social factors, and previous specialty preferences, differentiate those with preference for a primary care specialty, those with an unspecified specialty preference, and those with a preference for a non-primary care specialty *at graduation from* medical school?

Original Contributions

I recognize four potential original contributions of this study. First, I applied these methods to a large, de-identified, nationally-representative dataset (N=31,425) provided by the AAMC. This dataset allows for broad inferences to be made at the national level, which may inform policies aimed at addressing the anticipated primary care physician shortage.

Second, to my knowledge, there is no empirical study looking at recent trends. While two studies by Clinite and colleagues (2013, 2014) *do* provide data for medical school graduates as recent as 2013, their study included respondents from only 11 U.S. medical schools, and thus their results may not generalize to all contexts. Our dataset includes respondents who graduated from 2005-2011, allowing for both an analysis of recent graduates *as well as* an inspection of trends over this time period. Such a longitudinal approach may inform updates to workforce projections by groups such as the AAMC's Center for Workforce Studies or advocacy and intervention efforts in this field by groups such as the Robert Wood Johnson Foundation, the Robert Graham Center, and Primary Care Progress. Third, the interpretation of this study's results will help develop the conceptual basis and theoretical models that explain career decision-making amongst matriculating and graduating medical students. Rather than simply applying the same rationale and study design to an updated dataset (which *would* in fact provide some relevant information) this dissertation looks to build from a strong set of empirical studies (e.g. Colquitt, Zeh, Killian, & Cultice, 1996; Jeffe, Whelan, & Andriole, 2010) as well as building from a budding theoretical basis for understanding this issue(e.g. Bland, Meuer, & Maldonado, 1995; Bennett & Philips, 2010).

Fourth, and most importantly, this study specifically aims to understand the experiences, respondent characteristics, and pre-medical identities that lead to a preference for a career in primary care. Of the studies reviewed, only two (Babbott, Baldwin, Jolly, & Williams, 1988; Compton, Frank, Elon, & Carrera, 2008) looked at preferences both at matriculation *and* at graduation for the same population. As evidence suggests that a greater percentage of students matriculate into medical school with preferences for primary care than at graduation, this dissertation's assessment of preferences at both time points may help to explain factors associated with persistent inclinations for primary care careers. By investigating how one's pre-medical identity is associated with specialty preferences at these two time points, potential interventions can be developed to recruit and support the future primary care physician workforce.



Figure 1.1. Projected supply and demand for physicians from 2010 to 2025, showing workforce shortages of primary care physicians and non-primary care physicians. Data from "*The Impact of Health Care Reform on the Future Supply and Demand for Physicians Updated Projections Through 2025*" by the Center for Workforce Studies. (2010)

CHAPTER 2: REVIEW OF THE LITERATURE

Chapter Overview

Having defined the issue in the previous chapter, I continue the argument by reviewing the literature to demonstrate how this study will rely on existing research to provide an innovative analysis and discussion to the issue of primary care specialty preferences.

Approaching this topic can be aided by three important frames of reference. First, I show why investigations on primary care necessitate clarification of the very term *primary care* as it can be used both globally—to define *all* generalist and subspecialties—or used narrowly such as when the population of interest contains only one particular branch. Second, when reviewing the literature, it was necessary to distinguish between the outcomes of interest, generally called "primary care choice," which can be operationalized in multiple ways. Third, a sound theoretical frame of reference is required to guide the development of the study, such as the identification and operationalization of key constructs, and will allow for more robust interpretation of the findings. I review the existing models and explain how an integrative framework will guide this dissertation. I lay out these three frames, and follow it with a *concise* review of the literature and show how this study will contribute to this body of literature.

Guiding Frames

Primary Care Definitions

Approaching this topic necessitates defining which professions and specialties primary care includes. Indeed, a first distinction must be made between the *primary care workforce*—which includes physicians, physician assistance, nurse practitioners, pharmacists, and other medical assistants—and *primary care physicians or practitioners*—including those with a D.O. or M.D. degree (Bodenheimer & Pham, 2010). Although policy analysts agree that non-physicians and D.O. physicians represent a critical component of the primary care workforce (e.g. Carrier, Yee, & Stark, 2011), in this dissertation the focus is specifically those who are studying or have graduated from an M.D-granting, LCME-accredited U.S. medical school.

Even within the M.D. workforce, multiple definitions of primary care physicians exist in policy literature. Most common is the operationalization of general internal medicine, general pediatrics, and family medicine as the primary care specialties; however, these studies typically leave out subspecialties of internal medicine or pediatrics and focus on those with *generalist* careers. From a patient perspective, generalists are uniquely trained to provide a broad array of services such as chronic illness assessments

and preventive health care advice, while from a workforce perspective, these services are considered essential for the medical system to function properly (Colwill, Cultice, & Kruse, 2008). For these reasons, separating primary care generalists from specialists has a substantial record in studies on the primary care workforce (e.g. Starfield, Shi, & Macinko, 2000; Schwartz, Basco, Grey, Elmore, & Rubenstein, 2005; Robert Wood Johnson Foundation, 2003) and on specialty selection (e.g. Rosenblatt & Andrilla, 2005). However, other sensible definitions of primary care physicians exist. For example, the Health Resources and Services Administration (HRSA, 2014, np) state that primary care practitioners are "doctors of medicine (M.D.) and doctors of osteopathy (D.O.) providing direct patient care who practice principally in one of the four primary care specialities -general or family practice, general internal medicine, pediatrics, and obstetrics and gynecology." The inclusion of obstetrics and gynecology has been used in recent research to define and delineate primary care specialties (e.g. Jeffe, Whelan, & Andriole, 2010; Brotherton, Rockey, & Etzel, 2005), and suggests that its inclusion may be considered valid.

The empirical research on specialty preferences provides multiple ways of categorizing and analyzing primary care physicians. The most common definition includes general internists, family practitioners, and pediatric generalists, often analyzed with one single model (e.g. Erickson et al., 2013). On the other hand, some studies include both primary care generalists *and* some specialists (such as obstetricians and gynecologists) in similar models but separate analyses (Jeffe, Whelan, & Andriole, 2010; Lawson, Hoban, & Mazmanian, 2004), while others specifically investigate the selection

trends within each of family medicine (Senf, Campos-Outcalt, & Kutob, 2003), internal medicine (Hauer et al., 2008), and pediatrics (Frinter, Mulvey, Pletcher, & Olson, 2013). A number of combinations exists in the literature reviewed, and this diversity of definitions of primary care represents, therefore, a major limitation in this field. While analyzing these groups in one analysis *seems* like the most appropriate choice, some researchers posit that commonly-used predictors change signs and magnitudes based on the how primary care is defined, along with other inclusion criteria. For example, Senf and colleagues (1997, p. 530) posit that "these specialties [family medicine, general internal medicine, and general pediatrics] are competing with each other for students and that combining them will reduce the ability to understand what factors relate to choosing them individually." Lawson, Hoban, and Mazmanian (2004, p. S38) also note that researchers should "design studies with a specific field as a dependent variable." In sum, while it is possible to treat primary care as a composite variable, researchers generally agree that it is not a homogenous group. To that end, this review of the literature focuses on primary care *generalists*, and I delineate the specific branch when applicable.

With this in mind, and for the purposes of this dissertation which is framed around the current concerns of primary care training and distribution within the U.S. context, I categorized primary care generalist careers as internal medicine (no subspecialties), pediatrics (no subspecialties) and family medicine. Family medicine subspecialties are included since there are relatively few subspecialties for this field (Adolescent, Geriatric, Hospice and Palliative, Sleep, and Sports Medicine), few AGCME-accredited training positions for family medicine subspecialties, and a small number of residents planning such subspecialization training (American Board of Medicine Specialties, 2014; American Medical Association, 2015b). This categorization will best allow for an examination of the trends associated with those making career plans as a generalist physician.

Outcome Operationalization and Definitions

Having defined primary care itself, the timing and nature of the outcome must also be defined. In the literature the terms 'choice,' 'decision,' 'plan,' and 'selection' are all used in regards to the outcome variable, but these are not synonymous, and provide categorically different evidence. Moreover, the articles reviewed for this dissertation apply these terms to the outcome, when in fact, the questions ask about *preferences* or *plans*, which are of a categorically different nature. As Bland et al. succinctly state the issue, "interpretation is made even more difficult because several studies have found differences between a student's preference... a student's choice... and the student's career attainment" (1995, pp. 622-3).

While actual attainment is an indisputable fact, looking at medical students' perceptions and intentions before selection and attainment can indicate more about their processing of self-efficacy, expectations, and personal goals within the medical school environment. To that end, the terms *preference, intention, consideration,* or *plan* (a student's intended specialty goal, but also considering likelihood of achieving that goal) and *choice, selection, decision,* or *attainment* (the actual specialty that was chosen) are

used intentionally in this literature review. By and large, the outcomes reviewed in the literature focus on the former set of outcomes.

Review of Existing Theoretical Frameworks

A common challenge for medical specialty researchers—policy-makers, economists, meta-analysts included—is the lack of a clear theoretical basis to guide the already expansive volume of research on primary care specialty choice (Dohn, 1996). Although the groundwork for theoretical development began in the 1970's with Mitchell (1975) and Funkenstein's (1978) development of student-focused models predicting career-choices, their models include a noticeably narrow scope of predictors and are limited in their utility for research. However, the more recently developed theories of medical specialty preferences do provide a sound rationale for understanding these trends, especially in regards to primary care.

The recent theoretical literature has largely been developed off of secondary literature reviews. In this genre, the most commonly cited theoretical framework comes from Bland, Meurer, and Maldonado (1995), who reviewed 73 articles from 1987 to 1993. By incorporating various determinants of primary care specialty choice into a comprehensive schema, the authors created a typology of three sets of influential variables: student characteristics, student perceptions of the specialty, and medical school characteristics. The authors conclude with "the basic premise that career choice is accomplished as a student finds the specialty whose characteristics best match his or her own career needs" (p. 636). Thus, while students match the perceived benefits of a specialty to their own needs, school characteristics and faculty composition play a socializing role, as medical students are exposed to the dominant attitudes, values, and expectations of their institution. Through this seminal work, it became clear that students do not matriculate into medical school with unshakeable career intentions, but rather use this period to form opinions and investigate specialties that appeal to their career and personal goals and motivations. On this point, the authors also note that many students, through rotations and other learning opportunities, may construct inaccurate and distorted perceptions of specialties, which may not reflect its actual characteristics. As a consequence, they argue, matching one's career needs to an intended specialty is complicated by misunderstandings or poor experiences.

Building off the work of Bland et al., Reed, Jernstedt, and Reber (2001) applied subjective expected utility (SEU) theory to explain the rationale underlying medical student's specialty choices, with the goal of *improving* medical students' decisions rather than *predicting* them. Applying this theory to specialty planning, the authors systematically reviewed much of the same literature as Bland, Meurer, and Maldonado (1995), but examined the interaction of medical student's assets and career expectations specifically during the decision-making period. The authors' recommendation of making specialty decisions explicit during the student's development process (i.e. undergraduate medical education) suggests that gathering and incorporating knowledge—or even discarding knowledge—about specialties is a complex process where students are subconsciously assessing their fit with a potential career choice. Lawson and Hoban (2003) continued the theoretical development of primary care specialty choice by incorporating human capital theory into their framework in order to explore the rationale of medical students choosing a specialty. They determine that multivariate analyses hold the most potential for predicting primary care residency choice, as multiple factors (motivations, perceptions of specialty, demographics) predict specialty choice. In a subsequent article, these authors specify "a test of selected variables in the Bland-Meurer Model" using a single institution's dataset, and confirm that some of the demographic variables from the Bland-Meurer model did in fact influence primary care selection (Lawson, Hoban, Mazmanian, 2004, p. S36).

Of these and other similar theoretical frameworks, four immediate limitations should be noted. First, while most of these and other meta-analyses employ rigorous quality criteria for inclusion in their studies (e.g. Campos-Outcalt, Senf, Watkins, & Bastacky, 1995, p. 613), no empirical validation of their work is presented within them. Therefore, a limitation of conceptual models that *do* exist are often not tested, and therefore lack the validation that predicates their utility for policy and potential interventions. Second, the theoretical frameworks are often based on univariate primary analyses; therefore, these comprehensive schemas are based on narrowly-focused findings (Senf, Campos-Outcault, Watkins, Bastacky, & Killian, 1997). Third, because most these are frameworks rather than models, they generally overlook temporal dimensions of decision-making, such as how perceptions, attitudes, and ultimately actionable decisions are likely to change over the course of a medical student's educational trajectory as they experience different clinical rotations or have memorable interactions with faculty. Finally, they provide only cursory theory of actions for

explaining how and why a medical student commits to a primary care specialty. Stated otherwise, these frameworks answer the question "what matters?" but provide little response to questions such as "why does it matter?" "when does it matter?" or "how much does it matter?" As a result of these limitations, these frameworks do not allow for differentiation between students who confirm or change their perceptions based on these experiences; thus while comprehensive, this literature may simultaneously *overemphasize* the nature of influential sources and *underemphasize* their impact on medical students' psychological processes and subsequent career decision-making.

An Integrative Theoretical Framework

Responding to these limitations, I chose to incorporate socialization and career development theory within the most recent theoretical framework from the medical literature in order to preserve the localize context of medical schools, but interpret the results using a widely accepted theory for understanding graduate school experiences. To that end, Bennett & Phillips (2010) have proposed a theory of how initial preferences and environmental factors fully considers the temporal aspects of specialty choice, and incorporates it into their conceptual model. (See *Figure 2.1.*) Thus, it appears to be better suited to providing a structure to understand the decision-making processes, rather than just a decision itself.

This is important for at least two reasons. First, the commonly accepted idea of there being a "leaky" STEM pipeline is a curious one, since it fallaciously recognizes that individuals can exit such a trajectory, but not reenter it (Lewis, Menzies, Najera, & Page,
2009). While this may be appropriate for some populations where loss of interest or stagnation of skills would be detrimental, Bennett and Phillips' model uniquely theorizes that there are multiple entry points, or "pathways" into a primary care career before or during medical school. As an overview, these pathways are defined as:

- *PC Committed*, those with strong predisposition towards primary care at matriculation and at graduation;
- *PC Positive*, those with predisposition towards primary care at matriculation but who may or may not have interest by graduation;
- Undecided, those with no predisposition at matriculation but who may or may not have interest by graduation; and
- *Non-PC Committed*, those with no predisposition or interest in primary care at matriculation or graduation.

Second, being free from the restrictions of a single model predicting the trends of *all* students, Bennett and Phillips' model allows researchers to explore the multidimensional nature of particular determinants, and the potential impact it might have on particular subgroups. For example, in their model each student is influenced by their demographic background and predispositions, but the curriculum, economic incentives, and psychological processes that students encounter have a differential influence based on their initial proclivity towards primary care careers. To illustrate how this might occur, we only have to imagine the extreme opposites in this model: a *PC committed* and *non*- *PC committed* medical student with similar educational profiles who enter the same medical school.

For the former, we would expect to find an individual who matriculates with a planned interest in a primary care career path, based on family history or previous engagement with the medical community. This person engages with faculty role models who provide positive mentoring experiences in primary care fields, and their confidence in their commitment to primary care grows. Their role model validates their initial inclination, which in turn leads them to pursue other primary care opportunities throughout their medical school experience further solidifying their commitment and preparing them for a career in primary care. Conversely, the non-PC committed medical student enters with no established interest in primary care, given their career-related expectations such as prestige or income. For this individual, they have similarly positive role model experiences in primary care. However positive, the experience confirms their negative stereotypes about the primary care field and bolsters their decision to pursue a different career path. In this extreme exemplification, these two students' motivations and proclivities at matriculation correspond to how they might interpret the same experience. Other hypothetical students without such strong knowledge or motivations—the PC *Positive* and *Undecided* groups—would interpret this same role modeling experience with fewer predefined preferences, and thus their decision would be less predictable.

Socialization theory. Although Bennett and Philips' conceptual model holds immediate relevance for primary care career choices, it is important—particularly in

regards to the interpretation of the findings—that the theoretical framework has relevance for the medical literature, as well as a well-reasoned understanding of the socialization process that might explain these cognitive processes. To that end, I draw primarily from socialization theory for two intentional reasons.

The first reason is that socialization theory is consistent with the model proposed by Bennett and Phillips in a number of key areas, suggesting that integration is indeed possible. For example, both posit that pre-existing beliefs, motivations, and experiences shape the graduate environment, and each approaches this environment as a *subjective* one (Weidman, Twale, & Stein, 2001), which is especially germane for discussions around the so-called "hidden curriculum" that is theorized to instill the dominant cultural values of the medical community (Hafferty & Franks, 1994). Indeed, socialization theory is born from a constructivist approach, and assumes that unique thinking processes explain why individuals facing similar environmental circumstances make different reactions, behaviors, and career choices. Bennett and Phillips similarly posit that the environmental elements impact a student's specialty choice, since students "social norms may be defined or interpreted very differently based on a student's predisposition toward primary care or non-primary care." (p. S86). The idea of subjective perceptions and nonlinear development of these perceptions represents a fundamental trait of the Bennett-Phillips model, as they consider the perception of the specialty to be located within the experiences a student has, or the latent codes they might be receiving about it from other sources. To that end, both consider socio-environmental and cognitive elements of career decision-making as part of their comprehensive models. Indeed, the typologies of

variables from both models include elements from the person-environment structure (Osipow, 1990), in which the personality traits develop alongside their environment, where the student is an active agent, rather than passively being acted upon by their environment (Weidman, Twale, & Stein, 2001).

The second reason for integration is more straightforward: socialization theory has a historical record of use in explaining patterns of socialization in the graduate studies (Austin, 2002; Gardner, 2010; Boden, Borrego, & Newswander, 2011), and specifically in the STEM field (Austin et al., 2009; Cole & Espinoza, 2011) as well as in the medical education literature (Hafler et al., 2011; Merton, Reader, and Kendall, 1957; Hafferty, 2008). Thus, using a socialization framework will allow for the findings of this dissertation to be assimilated into the larger corpus of literature on this and similar populations.

Professional and academic socialization. Multiple models of socialization have been developed for graduate student populations. The most comprehensive of which comes from Weidman and colleagues (2001), who cite Brim's definition of socialization as "the process by which persons acquire the knowledge, skills, and dispositions that make them more or less effective members of their society" (1966, p. 3). In their framework for the socialization of graduate and professional students, they posits that the institutional culture and socialization processes are surrounded by ancillary environments (professional communities, personal communities, and novice practitioners) as well as an anticipatory environment (for prospective students). With this model, the authors propose that socialization is dynamic, non-linear, and influenced by formal and informal sources. That socialization is dynamic and non-linear represents an important and valuable aspect of this model, and the authors propose that there are four stages of socialization in graduate programs. In the *anticipatory* stage, when students are preparing for their program, they are becoming aware of procedures and behavioral expectations through anecdotal sources. Thus, the information they receive may be distorted, inaccurate or incomplete. During the *formal stage*, the student enters their apprenticeship and becomes familiar with the normative role expectations. The *informal* stage continues this trend, but is marked by the student learning the latent codes of conduct and communication. Finally, the *personal* stage denotes their internalization of the social norms such that their identity reflects a professional identity. Other graduate-specific models (e.g. Sandmann, Saltmarsh, O'Meara, 2008; Austin, 2002; Reybold, 2003) similarly focus on this final point; that is, identity formation represents a sort of "pinnacle" of graduate and professional socialization.

Participation and engagement within a community is intrinsically linked to the development of one's professional identity as well as the attitudes and beliefs that characterize it. In terms of academic or professional socialization, Lave and Wenger's (1991) and Wenger's (1998) conception of a "situated learning" locates learning and identity formation within social structures called a "communities of practice." Simply put, their theory posits that a shared and repetitive practice within a localized community influences one's attitudes, expectations, and ultimately, one's own identity. As the

authors acknowledge, the conception of a community of practice can be broadly thought of as a group of individuals who share a passion and interact regularly.

In this model, participation is not defined in any abstract terms; rather, it is seen as an active process through which individuals engage in a shared practice with members of their community. Three key elements define a practice (Wenger, 1998). First, practices within a community must be a mutual engagement through which members participate in activities and maintain the community structures. This brings the concept of *personal relationships* into the conceptualization of practice. Second, practice within a community is shaped by a joint enterprise of collectively negotiating a response to a situation or collectively developing group goals. Third, the practice must be defined by a shared and repeated repertoire, such as routines (behaviors), discourses (communication), or other community-specific interactions. However, actively practicing within a community does not itself explain how identities form as identities themselves are not static labels or traits, but rather the internalization of dynamic experiences and characterized by one's own meaning making of these experiences. Thus, as repeating a practice provides a rich set of experiences, one simultaneously learns the cognitive elements (i.e. content knowledge) and internalizes the social elements (i.e. social skills and attitudes) of their community and internalizes these forms of knowledge into their own professional identity. Identity is a constructed self-narrative (Linde, 1993; McAdams, 1993) dependent on how one interprets their "self" in relation to their lived experience, and can be represented by both one's value systems and motivations. A value system, in the case of professional identity, relates to the attitudes, ethics, and beliefs one maintains and

practices in their occupational role (Kanter, 1968). Closely aligned are motivations, representing the impetus for an action or engagement. Although value systems may be subsumed within one's motivational structure, these also include other intrinsic or extrinsic motives which are personal rather than normative.

For this dissertation, a community of practice can be represented by those kinds of recurrent experiences that medical students might have prior to matriculation, especially those which are situationally and contextually dependent (Weidman et al., 2001). These may take the form of academic or pre-medical clubs, physician shadowing opportunities, coursework, laboratory work, interactions with the medical community, among many other potential experiences. In essence, it is suspected that while pre-medical students are forming career plans, they are also actively engaged with related communities in which they are sharing information and undertaking interpersonal activities related to their future medical careers. As such, members of these communities of practice interact and learn in a shared and mutually constructed environment (Lave & Wenger, 1991).

While the term community of practice is not found in the literature on the medical school environment, the concept certainly is. We can locate elements of communities of practice in the well-known ritual that is the white coat ceremony, a traditional rite of passage in one's medical career, where the values and morals of the medical profession (through the reading of the *Hippocratic Oath*) and the dress (through the "cloaking") publically represents a transition from one identity to another. Particularly germane to this study is the "match day party" ritual that occurs each March, where students

publically gather to hear with which residency they have been matched. Certainly, we could imagine the pressure students must have to be accepted into a prestigious residency, and receive the respect and adulations from their closest peers. Because matching to a prestigious residency is nearly synonymous to being accepted into one's career specialty, this is no small matter and students, even early in their medical education, may be thinking about this moment with great anticipation.

Beyond these rituals, we should also consider the philosophy and visions of one of the founders of modern medicine, William Osler, who revolutionized the field by advocating for and formalizing the apprenticeship method, the ancestor of today's residency system (Osler, 1893; Dornan, 2005). His belief that students learn best at the bedside shows how intentionally medical education is situated in the professional working environment. To this end, communities of practice within the medical school environment are acknowledged to have strong socializing influences (Hafferty, 2008) as the "hidden curriculum" lays the foundation for not only professional standards, but also acceptability of choices, communication, and opinions. As Hafferty and Franks state, "although matters of technical information and the transmission of technical skills traditionally have been through to lie at the heart of the medical educational system, medical training at root is a process of moral enculturation" (1994, p. 861). From this perspective, the content knowledge is subsumed within a *context* knowledge, where students are expected to internalize the values, attitudes, and beliefs of their immediate culture. If we can reasonably assume that many of these aspects of a medical school

culture are not only hidden, but also *imbedded*, it becomes clear how the communities of practice and their powerful symbolic rituals can impact a student's career aspirations.

Researchers also recognize that these kinds of experiences influence career and specialty decision-making, especially in regards to student-faculty interaction, where faculty members may engender particular attitudes towards particular career preferences (Erickson et al., 2013), for example by explicitly providing advice and encouragement to enter a primary care field (Block, Clark-Chiarelli, & Singer, 1998) or by negatively portraying the primary care field (Scott et al., 2007). In these examples, the curriculum could correctly be called informal, rather than just "hidden," since they are tacit but not openly acknowledged. Finally, if this environment is indeed dynamic, then the faculty composition plays a substantial role in determining the attitudes, beliefs, and values (Haffler et al., 2011). Therefore, how a school structures the students' time and interaction with primary care physicians has non-trivial consequences for career decision-making, especially at the beginning of their professional careers, when interactions with faculty members may have a profound impact on how they internalize professional values and form career aspirations.

Parental and familial socialization. It is also important that socialization from parents and family members be taken into consideration, as it is sensible to presume that they influence the pre-matriculation plans, opinions, as well as the budding physician identities of medical students. Parental background and influence is an important, if understudied, aspect of career choices in medicine. Given that potential medical students have differing levels of economic and social capital to spark and sustain an interest in medicine (Grbic, Garrison, & Jolly, 2010; Jolly, 2008; Lentz & Laband, 1989), it is not surprising that the results from the 2012 Matriculating Student Questionnaire (MSQ) show that about 30% of medical students rated "Parent" as a factor for studying medicine (AAMC, 2012a). Despite this, the influence of parental occupation – especially for parents who are physicians – on their children's educational or career choices in medicine remains understudied even in the general medical student population.

With this in mind, Lentz & Laband's theoretical development of knowledge transfers (1989) between physician-parents and their children posits that formal and informal interactions possess motivational influences on the child's interest in medicine and also provide practical knowledge of how to secure admission. In that sense it is expected that they would be more likely to develop a physician identity and pursue a medical education; however, these authors also posit that physician-parents can provide career-relevant advice that other children would not easily acquire. Indeed, because these interactions so often occurred in the household, the cost of acquiring this knowledge is negligible or even inevitable for the children with physician-parents. In other words, it is suspected that physician-parents provide an environment that might lead their children to also pursue a medical degree. On this point, Nowacek and Sachs (1990) speculate that children of physician-parents are "thought to be more committed to medicine because he or she knows what it means to be a physician" (p.143), while Kleshinski, Khuder, Shapiro, & Gold note the "general perception among the medical community of the positive influence and environment provided by physician-parents" (2009, p. 76). In

regards to specialty preferences, Colquitt and colleagues (1996) claim that "physicians in practice may hold particularly negative perceptions about generalist–primary care careers, and they may convey these perceptions to their children" (p. 955). Thus, while it is hypothesized that parent-physicians may create an environment that supports interest in medicine, this environment may also create implicit or explicit incentives to choose a non-primary care career. Children of physician-parents may also have a better sense of what specialties exist, what the nature and prestige of the specialties are, and how the specialties align with their personal and career motivations. To exemplify, we might imagine the child of a physician-parent shadow their parents' colleagues, and be introduced to a variety of specialties beyond what might be learned through a typical visit to the doctor. Such knowledge may have long-term consequences for their career decision-making.

An integrative conceptual model. To conclude briefly on the integration of socialization theory into the Bennett-Phillips model, it is assumed that both professional and parental/familial interactions influence potential medical students to form an identity as a medical student or pre-physician. Although the strength of the identity may be tenuous, it is assumed that traits of their identity can be seen through motivations and values regarding their role as a physician, in essence, their predisposition or predilections. As this literature review will outline, these predispositions may be associated with lifestyle, economic, altruistic, or research and innovation motivations for becoming a physician, and furthermore, one's reasons for choosing a particular specialty.

By integrating parts of socialization theory into the Bennett-Phillips model, I have created a single conceptual model for the purpose of operationalizing variables and sensitizing the interpretation of findings in this dissertation. *Figure 2.2* graphically lays out this model and posits that one's demographic profile is filtered through lived experiences and socialization processes-informal and formal, at home and professional settings— which results in what I have called a 'pre-medical identity.' This identity is composed of predispositions such as values, attitudes, and skills related to the medical profession. In turn, these students attempt to match their identity to a career plan, specifically, their preferred specialty.

As one's identity is fluid, it is assumed that these dispositions are further molded during the medical school experience through professional socialization to influence their graduating preference. To that end, their identity and predispositions at graduation may be categorically different at matriculation, and therefore, their specialty plan may also be different. Finally, although beyond the scope of this study, it is important to note that preferences and career planning are of a different nature than actual career attainment, which is represented by choice and match of a residency and eventual licensure or certification. Thus, while one's values and motivations play a role in determining preferences during the medical school stage, we would expect that during in the residency choice process, competency—such as STEP 1 or STEP 2 scores—would play a larger role.

A Brief Review of Seminal Investigations

Based on this definition of primary care and noting the kind of outcome for each study, I focused this review of the literature on both broad and narrow definitions of "primary care" as used in the literature. It should be noted that this was a non-systematic review of MEDLINE, PhsycInfo, Academic Medicine, JAMA, and Dissertation Abstracts which used the terms 'specialty choice' and 'primary care,' 'internal medicine,' 'family medicine,' 'pediatrics' and other synonyms as the main search categories. Importantly, because so many systematic reviews and meta-analyses have been undertaken on this topic (Bennett & Phillips, 2010; Campos-Outcalt, et al., 1995, Lawson & Hoban, 2003; Senf et al., 2003; Reed, Jernstedt & Reber, 2001; Meurer, 1995; Bland, Meurer, & Maldonado, 1995), I summarize their findings when appropriate and juxtapose it with the more recent primary literature available. This is important because of significant changes to the U.S. and medical student population. For example, in light of the trend of women increasing their proportion of the medical school matriculant population—reaching parity (49.1%, 7,923/16,142) for the first time in 2007 *after* many of these systematic reviews had been published—and we must see if these trends still hold true (AAMC, 2008).

Thus, in this section I organize the literature according to what outcome the authors investigated. Through this, I present the major trends of student demographics and characteristics, student values and motivations, student perception of the specialties, and their relationships to these outcomes.

Intentions or Preferences at Graduation

Graduation Questionnaire (GQ) studies. By and large, the most common data source for graduate preferences comes from the AAMC's Graduation Questionnaire (GQ), which is used for program evaluation purposes (AAMC, 2015a). The survey is administered annually to all graduates of LCME accredited U.S allopathic medical institutions and is voluntary and confidential. Researchers commonly recognize three advantages of using this data source. First, because a large suite of questions are asked of graduate students (clinical experiences, student services, career intentions, finances, et cetera), researchers can analyze the data and respond to issues such as students' psychological development and human resource management. Second, researchers can also link these data to other AAMC data sources, such as the Post-MCAT Questionnaire (PMQ; AAMC, 2015b) and the Matriculating Student Questionnaire (MSQ; AAMC, 2015c), providing a deeper understanding of longitudinal processes and population changes. Third, although voluntary and thus does not have a 100% completion rate, the data is considered to be nationally representative (AAMC, 2014a). For this section of the literature review, the GQ items of interest ask the respondent if they are planning to becoming certified in a specialty, and if so, which specialty (AAMC, 2015a).

Looking at this specific outcome, two studies provide some insight into such population trends during the relatively uncertain 1990's, when primary care selection of great interest to researchers and policy-makers. Kassebaum and Szenas (1994) evaluated the "factors influencing the specialty choices of 1993 medical school graduates" (p. 163), and organized these factors into a hierarchy of importance on preferred specialty preference. By comparing these factors across four different specialty groups (primary care generalist, medical specialists including primary care subspecialties, surgical specialists, and support specialists), they found that generalists rated patient contact and patient education as strong influencers of their specialty preferences, while income prospects, lifestyle factors such as predictable working hours, and research opportunities were not rated as influencing their preference to choose a generalist career. While this research demonstrates trends related to important career factors show elsewhere in this review, this study's lack of any rigorous statistical analysis or control variables reduces the findings to only a cursory indication of influences on specialty preference at graduation.

Using the more robust approach of logistic regression to uncover the importance of debt on intended specialty choice, Colquitt, Zeh, Killian, and Cultice (1996) investigated the specialty intentions of 1991 and 1992 graduates, using an outcome of a career intention as a primary care generalist. The authors included important covariates at the institution level (ownership, percentage of graduates entering primary care) as well as at person-level (age, race/ethnicity, gender, MCAT-Science score, amount and type of debt), and importantly, included a student's "preference of the primary care specialty upon entry to medical school" (p. 402) and whether or not one of the student's parents was a physician, as this may have socialized them to eschew primary care fields. (The authors characterized the last two as "predispositions to specialty.") Analyzing general internal medicine, pediatrics, and family practice in separate models allowed for the trends to be explored and interpreted within each specialty, but in general the authors conclude that the most consistent predictor of the intention to specialize in primary care at graduation was in fact the intention to do so at matriculation. For example, the independent odds ratios for a matriculating preference for general internal medicine (2.59), family practice (3.70), and especially for pediatrics (5.35) imply a stability of preferences before, during, and upon graduation from medical school.

Lambert and Holmboe (2005) investigated the same GQ outcome of intended specialty for U.S. medical school graduates from 1990-2003 with a particular interest in gender differences over this time period. Looking at specialties with "uncontrollable lifestyle specialties," most of which were in primary care, these authors showed that women were more likely than men to prefer a specialty with an uncontrollable lifestyle. However, they also demonstrated a dramatic decrease by *both* genders preferring these specialties from 1996 to 2003, by about 20% over this period. Thus, rather than gender, this study shows that the kind of lifestyle (controllable versus uncontrollable) associated with the specialty explains career intentions, rather than the students simply moving away from only primary care. This finding appears to support the theoretical basis for the importance of perception of specialty characteristics (Bennett & Phillips, 2010) as well as matching one's own need to those characteristics (Bland, Meurer, & Maldonado, 1995). Although this univariate analysis (gender predicting type of specialty) omits important intermediating variables, this study's strength is that it provides a nationallyrepresentative analysis demonstrating that gender may in fact be a proxy for critical

motivations and lifestyle choices, important determinants of specialty preference as demonstrated by Kassebaum and Szenas (1994).

While these studies are somewhat limited in scope and somewhat outdated, Jeffe, Whelan, and Andriole (2010) provide a more comprehensive assessment of the factors associated with intended specialty at graduation, using the GQ data for 102,673 U.S. medical school graduates from 1997-2006. Looking at a suite of demographic information (student race/ethnicity, gender, debt amount, parental occupation), MSQ data elements (factors for choosing medicine; perception of medicine as a profession), and GQ data elements (serving in an underserved area; preferred career setting), the authors apply logistic regression methods separately to four preferred general primary care specialties (general internal medicine, family medicine, general pediatrics, and obstetricsgynecology).

Their analyses show that by and large, women had higher fitted odds than men to express an intended interest for a career in general internal medicine (*OR:* 1.78), family medicine (*OR:* 1.88), and general pediatrics (*OR:* 4.50). These large odds ratios support their conclusion that "the steady increase in the proportion of women among U.S. medical school graduates over the study period played a critical role in limiting the overall decline in proportions of graduates who planned generalist–primary care careers" (p. 953). In contrast to Colquitt and colleagues (1996) use of numerous race and ethnicities, these authors look specifically at how URM, Asian, and Other/unknown students compare to the White referent group and find that for pediatrics and family

medicine, White students have higher fitted odds of preferring these specialties. No universal trend was found for general internal medicine. Regarding reasons for pursuing medicine as a career, those who reported more altruistic beliefs about healthcare or gave more importance to the social responsibility of medicine had greater fitted odds of intending to specialize in general primary care. Consistent with the findings from Colquitt and colleagues (1996), these authors found that having at least one physicianparent was associated with a lower likelihood of choosing a general primary care career and state that "physicians in practice may hold particularly negative perceptions about generalist–primary care careers, and they may convey these perceptions to their children" (p. 955).

The time period for this analysis (1997-2006) corresponds to the decrease in primary care interest found by Lambert and Holmboe (2005) across the same outcome variable, and Jeffe et al.'s findings appear to support trends found in other investigations. Most saliently, this group of studies found strong evidence that the student's gender to be either independently or jointly predictive of their intention to pursue a primary care specialization. This suggests that women are still more likely to show preference for primary care specialties, though the differences may be smaller and vary by specialty. The findings also confirm that perceptions of medicine and lifestyle factors for choosing medicine as a career are relevant to specific career choices. For example while Kassebaum and Senzas (1994) found patient contact, lifestyle, income and prestige to be predictive of choosing a generalist career, Colquitt et al. (1996) noted that perceptions of having a controllable lifestyle appear to be important in the career decision-making process, and Jeffe and colleagues (2010) found similar trends by using attitudinal predictors such as social responsibility, importance of prestige, and altruistic beliefs.

Non-Graduation Questionnaire studies. While these analyses of GQ data provide strong evidence of trends in specialty choices, it should not be considered as the only source of evidence of career intentions at graduation. Indeed, other studies have used single-institution surveys to show different factors and demographics are associated with specialty preferences.

Looking at what predicts a medical student's consideration of family medicine at graduation, Scott, Gowans, Wrght, Brenneis, Banner, & Boone (2011) investigated demographic and attitudinal factors amongst 1,542 fourth year medical students at eight Canadian medical schools. Using an attitudinal battery that closely matches items from the MSQ, they created items such as 'prestige,' 'medical lifestyle,' and 'societal orientation' which match those found by some of the factors that Jeffe et al (2010) and Kassebaum and Senzas (1994) developed. Additionally, they found other reliable principal components such as 'varied scope of practice,' 'hospital orientation,' and 'role model.' As with other studies, these authors applied logistic regression to the outcome of considering a career in family medicine (versus other) and found that the factors 'societal orientation' and 'varied scope of practice' positively predicted interest in family medicine, while other potential socializing experiences such as having a 'relative or friend in medicine' or having a parent with postgraduate education negatively predicted it.

Matching the findings from Colquitt et al (1996), they state that "having family medicine as a career choice on entry was the most important predictor of a student's ultimate career choice" (p. E4), once again suggesting some stability of specialty consideration from matriculation to graduation. Thus, even if reported preferences change after impactful experiences, it appears that the stability of preference over time remains the most predictive factor of long-term preferences. Interestingly, the authors also investigated students who had changed to a preference for family medicine, and found that lower interest in research predicted this change; conversely, for students who changed *away from* a preference for family medicine having a higher research interest or relative or friend in medicine were also predictive of this change. Although this study is based on a Canadian population, their results substantiate three important trends found in the GQ-focused studies and the systematic reviews: first, socializing trends—such as having a relative (or parent) physician—appears to have a strong and negative association towards maintaining interest in primary care. Second, motivational and attitudinal factors also appear to explain both career choices and the development or changes in career choices.

Another specialty-specific study of this nature was performed by Hauer et al. (2008) who investigated the factors associated with a planned career in internal medicine amongst the class of 2007 medical students at 11 schools. Asking students to rate what influenced their specialty choice, the authors found that those with a preference for internal medicine rated 'intellectual challenge,' 'commitment to patient care,' 'personal

satisfaction,' higher than those with a different specialty preference. This contrasts previously-mentioned studies finding research opportunities to negatively predict an interest in primary care (e.g. Jeffe, Whelan, & Andriole, 2010), and suggests that internal medicine be perceived differently or attract a different population than family medicine or pediatrics. Additionally, while those with a preference for internal medicine rated the internist lifestyle more highly than others, the authors note that the data suggest a concern with the internist lifestyle across the whole population. They posit that students hold a negative perception of the internist lifestyle due to an expectation to work long hours in a hectic work environment.

Longitudinal or quasi-longitudinal studies. Studies such as Colquitt et al. (1996) and Scott et al. (2011) provide a very interesting analysis by including specialty preferences at matriculation to predict specialty preference at graduation. That said, these kinds of analyses are not truly longitudinal, and therefore limit themselves to identifying preferences at one time point rather than the optimal *two* time points. In my review of the literature, some studies have taken this approach and provide an indication of how preferences change over time and provide a stronger rationale for understanding how socialization impacts a student's career decision-making process.

The seminal longitudinal investigation conducted by Babbott and colleagues (1988) described the stability of specialty preference from pre-matriculation to graduation for a nationally representative sample of the class of 1983 graduates from U.S. medical schools. Dividing the 15 possible specialty preferences into primary care, specialty care,

and supporting services, the authors found that about 20% of graduates chose the same specialty as they did prior to matriculation. Looking specifically at primary care specialties, about 20% fewer men and 9% fewer women who had initially chosen a primary care specialty (including pediatric or internal medicine subspecialties) later reported a preference for a primary care specialty by graduation. Moreover, for those students with a preference for a primary care specialty at pre-matriculation, the researchers found that less than half reported a preference for *any* primary care specialty by graduation. However, the authors also found that students who initially preferred a non-primary care specialty had expressed a preference for primary care (notably, for internal medicine) by graduation. Thus, consistent with suppositions of the Bennett-Phillips model, some *Undecided* or *PC-positive* had shifted their career goals during their medical education.

While this confirms a general trend of students moving away from primary care during their medical education, we should note that class of 1983 graduates probably experienced a different academic and social environment than today's student would. Nonetheless, these findings provide a representative benchmark showing a more popular preference for primary care specialties at matriculation, and a trend for both men and women to abandon this preference by graduation. Importantly, the stability of preferences varied widely by specialty; for example, the stability of preferences for family practice (24.2%) and internal medicine (25.9%) were notably higher than for pediatrics (21.6%).

Of great interest to this dissertation is that of Compton, Frank, Elon, and Carerra (2008), who administered three surveys looking at career intentions (matriculation, entry to wards, and senior year) to class of 2003 students at 15 U.S. medical schools. Using the covariates of gender, desire for high prestige, and interest in prevention, the authors were able to model the stability of specialty preferences, defined as undecided, primary care, or non-primary care. Their analyses revealed that medical student's career preferences are relatively unpredictable over these three time points, even when considering only three possible outcomes. Changes were less prevalent for those who initially chose nonprimary care than for those who initially stated a preference for a primary care career, where about 47% of students maintained their preference for this career option by senior year. Moreover, they found that those who were undecided at matriculation were much less likely to report an interest in primary care by senior year. To that end, their finding that only about 30% of students who initially chose primary care also selected it at matriculation can serve as a kind of benchmark for the findings of this dissertation. Consistent with other studies reviewed here (e.g. Jeffe, Whelan, & Andriole, 2010), these authors found that a desire for high prestige was associated with non-primary care career intentions.

These findings confirm previously cited trends, for example, that medical students who change preference are more likely to move away from primary care specialties (Markert, 1991; Babbott, Bakdwin, Jolly, & Williams, 1988). Noting some possible social mechanisms for this finding, Compton et al. suggest that unfamiliarity with some specialties at matriculation and medical school culture may explain these changes. To that end, these findings also demonstrate the relative unpredictability of specialty preferences over the medical education period (Kassebaum & Senzas, 1995; Scott et al., 2011), and show some empirical support for the Bennett-Phillips model, specifically that stability of preference is impacted by gender and career motivations such as desire for a prestige career.

A quasi-longitudinal qualitative study by Scott, Wright, Brenneis, Brett-MacLean, and McCaffrey (2007) supports the trends found by Compton et al. (2008). Specifically, they found that amongst 33 students with a preference for family medicine at Canadian medical schools, socialization by faculty during their medical school experience lead to generally negative impression of family practice. For example, students in this study expressed a loss of interest in primary care due to reinforcement of negative stereotypes by faculty, or lack of faculty to act as role models in the first place. Indeed, students in this study reported little interaction with family medicine faculty. In their clinical phase, many students reported that the family physician role models they did have expressed their own frustrations with their work and lifestyle, but allowed them to see positive aspects of the role as well. Finally, students expressed that ease of match for family medicine represented a negative attribute. Thus, it appears that the perception of family medicine being an unprestigious career option was maintained at various points of students' medical education by a commonly perceived negative representation of family medicine, and is consistent with the finding from Babbott et al. (1988) that the greatest decline in interest of any specialty from matriculation to graduation was in family medicine ($\sim 60\%$).

Additionally, when discussing their interest in family medicine as it related to students' desired lifestyle, low income was discussed as a deterrent while work-life flexibility was discussed as an attractive aspect of being a family physician. In their estimation, clinical experiences and working with family physician role models allowed students experience the onerous, but rewarding lifestyle and match it to their own career and lifestyle preferences. Another finding from this study supported by the literature is that students enter medical school with limited assets or knowledge about specialties (Bennett & Phillips, 2010; Bland, Meurer, & Maldonado, 1995), or may have faulty information about their intended specialty. Thus, the clinical rotation period represented an important method of rectifying knowledge gaps and seeing how the characteristics of a specialty matched their own personality. Although this study only provided evidence for family medicine, a benefit of such a qualitative analysis is that it provides an indication of how students' motivations were impacted by socialization and interaction with specialty role models.

A final study in this genre compared the perceptions of a cohort of first year medical students to that of a cohort of fourth year medical students (Clinite et al., 2014), and demonstrated many of the same trends as found in other studies. For example, they demonstrated that fourth year students rated 'having time off of work,' 'having control of work schedule,' and 'financial compensation' as stronger lifestyle considerations than first year students; in contrast, first year students rated 'enjoying the kind of work I am doing' higher than fourth year students (2014, p.4). Speculating on the source of these

differences, the authors state that "as learners strive to become a part of a physician community and increase their professional role through clinical training, their views of physician lifestyle and the medical profession may be most influenced by their supervisors during training" (p. 5). In sum, the differences between first and fourth year medical students may be rooted in their professional socialization to medical culture and may suggest that the underlying rationale for a desired lifestyle may shift, and ultimately impact one's decision-making processes. With regards to those with a preference for primary care specialties, there was no difference between first and fourth year students' perception of the importance of social consciousness and working with underserved populations.

Intentions or Preferences at Matriculation

In the previous sections, I summarized the robust literature on graduates' specialty preferences, however, looking at specialty preferences at matriculation can also provide an indication of the reasons and motivations students have to enter medicine in the first place.

Of note is the relative paucity of literature looking a career choices and motivations upon matriculation. Indeed, this literature review found only one study that used career interests at matriculation as the primary outcome (though the evidence from quasi-longitudinal studies should be noted). This unique study by Clinite et al. (2013) undertook a cross-sectional study of 1,020 first-year medical students at 11 medical schools, looking specifically at student demographics, lifestyle preferences, specialty perceptions, and at the outcome which specialty they preferred first, preferred second, and preferred least. Based on the student's response, they were classified as *PC-first*, *PCsecond*, *PC-mixed*, *PC-least*, or *no opinion*, where primary care was defined as family medicine, general internal medicine, and pediatrics. Thus, in contrast to most studies which use a dichotomous outcome, this study takes from Bland, Meurer, and Maldonado (1995) and allows for a more nuanced categorization and allowed the researchers to distinguish students who may be sure of their interest in primary care, as compared with those who are slightly inclined.

Analyzing five kinds of desired lifestyle factors in separate one-way ANOVA models, only 'financial compensation' significantly distinguished the five groups, with *PC-first* rating it the lowest factor of a good physician lifestyle and *PC-least* rating it the highest. This appears to support the literature on graduate career preference already reviewed demonstrating prestige or salary has a negative influence on preferring a career in primary care (e.g. Kassebaum & Senzas, 1994; Jeffe, Whelan, & Andriole, 2010). While this difference distinguished matriculants with different primary care preferences, it was rated as the least important lifestyle characteristic. In turn, enjoyment of the work and work environment were rated as much more important. This confirms that while finances are an important motivator in the career decision-making process, controllability of work environment appears to hold more gravitas on that process (Lambert & Holmboe, 2005). These findings are also consistent with the literature showing that perceptions of specialty characteristics such as working with underserved populations and desiring a high salary distinguished those with different degrees of primary care preference at matriculation (Scott, et al., 2011). Research opportunities also distinguished these groups, substantiating the findings another large-scale study by Kassebaum, Szenas, Ruffin, and Masters (1995) showing limited career research ambitions among generalists. These findings also aligns with a study on premedical students' perceptions of primary care showing that potential matriculants already have negative associations with primary care such as lack of prestige, income, and responsibility (Gold, Barg, & Margo, 2014). This suggests that formation of career opinions and attitudes may already be formed even before matriculation.

While this study provides mostly consistent results with the graduates' career preferences, Clinite et al. (2013) also found that perceptions work characteristics did not distinguish medical school matriculates of any interest in a primary care career, while Kassebaum and Senzas (1994) did find that lifestyle characteristics significant distinguished those with career interest in general primary care from others. Explaining this inconsistency, Clinite et al. note that medical students having negative primary care clinical experiences may explain the trend of declining interest in the kinds of characteristics associated with primary care (Scott et al.; 2007; Compton et al., 2008; Lambert & Holmboe, 2007).

Review of the Extant Literature

Major Limitations of the Literature

Methodologically, many some studies used use robust sampling methods (e.g. Clinite et al., 2014) or nationally-representative datasets (e.g. Babbott et al., 1988), and many provide further evidence that their sample is representative of national trends (Scott et al., 2011). However, while these studies provide an important starting place for this dissertation they should also be evaluated in the context of their limitations.

First, it is important to consider that the data these studies investigate come mostly from surveys or questionnaires, in particular the AAMC's Graduate Questionnaire (GQ). Surveys, while practical, are by themselves inadequate for studying or understanding the intrapersonal cognitive processes or changes in emotions or feelings. As many of these studies *do* use attitudinal batteries to understand the relationship between lifestyle and career preferences, this issue is non-trivial. Additionally, being a survey, there is no guarantee that the respondents have been truthful or answered questions in a way that reflects their actual feelings or situations. Indeed, survey or respondent "fatigue" may play a role in the quality of the data provided in lengthy questionnaires, such as the MSQ and the GQ (Sharp and Frankel, 1983).

Second, researchers have commonly identified the problematic nature of "faulty information" that students might receive about the specialties before or during medical school (Bland, Meurer, & Maldonado, 1995; Bennett & Philips, 2010; Rogers, Creed, & Searle, 2009). Given the numerous specialties a matriculating medical student can choose from (Freeman, 2004), as well as the creation and recognition of new specialties (Cassel & Reuben, 2011), it is sensible to presume that they may not be aware of *all* the specialties which they could potentially pursue, and therefore have not evaluated their response with a complete body of information. For example, Compton et al. state that "students may enter medical school relatively familiar with some disciplines, such as pediatrics and surgery... but may be less familiar with other disciplines like psychiatry and preventative medicine" (2008, p. 1097-8). A third limitation is inherent to the very nature of we seek to study, as the rapidly changing landscape of medicine are likely to influence specialty preferences. Indeed, as the first chapter demonstrated, the consistent trend of students moving *away* from primary care residencies and careers suggests that specialty preferences is a non-static social phenomenon. Therefore, the utility of the results provided by previous studies—even high quality ones—depends on the timeliness of the data that was analyzed.

Finally, and of greatest importance to this dissertation, a great majority of these studies do not evaluate their results within a longitudinal framework, and thus it remains unclear as to how perceptions, motivations, and lifestyle inclinations are associated with a student's specialty preference. This is non-trivial, as all of these studies assume that there is a socializing structure that underlies students perceptions; for example, the physicianparent who exposes their children to some kind of professional knowledge or attitudes or even the way that clinical faculty and preceptors discuss the importance of specialty characteristics such as prestige, innovation, anticipated income, social responsibility, et cetera. Therefore, observing their career intentions (by way of specialty preference) at only one point gives no indication of how these socializing elements have impacted career preferences. Indeed, that only Compton et al. (2008) produces a model that addresses this issue serves as an indication of where scholarship has yet to explore.

Synthesis and Conclusion of Relevant Predictors

Considering these limitations, this review has also highlighted many salient trends which have immediate relevance for the scope of this dissertation. This is especially true for the selection and treatment of covariates that will be used to test the Bennett-Phillips specialty selection model. Thus, what has been exposed in this review of the literature is the complexity of their interaction, and the importance of timing as well as the specificity of the outcome.

The demographic factors featured in this review appear to be associated with differential specialty preferences and career interests. Of these, gender represents the strongest indicator of a commitment to primary care both at matriculation and at graduation (Jeffe, Whelan, & Andriole, 2011), but this depends on the specialty (Hauer et al., 2008; Scott et al., 2011). That said, gender was not universally found to be predictive across all outcome types and specialty preferences (Scott et al., 2011; Colquitt et al., 2008), though some suspect that this may be due to women reaching parity in the medical school population in recent years (Lambert & Holmboe, 2007).

The student's race and/or ethnicity was included as covariate in multiple studies, but no strong conclusion can be drawn from this synthesis of the literature. For example, Colquitt et al (1996) found no consistent association between race/ethnicity and a graduate's preference for general internal medicine, pediatrics, and family practice. Thus it appears that the connection between race/ethnicity and primary care preference may depend on inclusion of covariates or the composition of race/ethnicity variables, suggesting a complex interplay between demographic and psychosocial variables. That said, Jeffe, Whelan and Andriole (2010) *did* find that underrepresented minorities (Black, Hispanic, and Native American students) all had lower odds of preferring primary care specialties than White students, when controlling for a suite of other student characteristics. Hauer et al. (2008) also found that underrepresented minorities had about half the fitted odds of preferring internal medicine when compared with other students.

Parental backgrounds also played a role in explaining career preferences in many of the reviewed studies, with some using parental socioeconomic status (e.g. Colquitt et al., 1996) or parental occupation (Jeffe, Whelan, & Andriole, 2010; Scott et al., 2011). In terms of occupation, many studies treated the experience of having a parent-physician as a demographic variable, but also regard its socializing impact, and find that it has a negative association with a primary care preference. In the estimation of Colquitt et al., having "insider knowledge and the acquired experience of physician parents might influence student's specialty decisions" (1996; p. 409) while Jeffe, Whelan, and Andriole (2010) find a similar trend and speculate that physicians may harbor negative stereotypes of primary care specialties and express them to their children. Thus, it appears that having a physician parent—perhaps even one practicing primary care— socializes a medical student against pursuing a more demanding, less prestigious specialty with lower anticipated income.

In terms of psychosocial variables, what can generally be called career motivations have been well studied in this corpus of literature. Multiple studies found altruistic intentions (Jeffe, Whelan, & Andriole, 2010) or variables dealing with patient contact, education, and social responsibility (Kassebaum & Szenas, 1994; Haurer et al., 2008) to be positive associated with preference for a career in primary care or as a generalist. A factor produced by Scott et al. (2011) called 'societal orientation' provides a slightly different approach to altruistic intentions, focusing on relationships with patients and communities. Nonetheless, the authors found that prospective family physicians rated these items significantly higher than those with other specialty preferences. Additionally, the literature consistently presents evidence that research opportunities play an extremely strong role in pushing students away from primary care or generalist fields (Kassebaum & Szenas, 1994; Clinite et al., 2013). Jeffe, Whelan, & Andriole (2010), for example, found that the higher one rated research and innovation as a career motivation, the lower the fitted odds that they would prefer a primary care specialty by graduation. This suggests that intellectual demands and technological innovation are not identified as important for those who prefer a career in primary care; state another way, those medical student with career interests including technological and innovative opportunities do no, on average, have a strong interest in pursuing a career in primary care.

The importance of prestige is consistently associated with a negative influence on primary care preference; however, the definitions of prestige put a varying amount of weight on income, autonomy, leadership positions, and other elements that very well mark a "prestigious" career. For example, using matriculation data, Jeffe, Whelan, and Andriole (2010) created a broad prestige factor based on status, job security, authority, and income and found it to be a statistically significant predictor in separate models predicting a preference for family medicine, general internal medicine, and pediatric medicine. Similarly trends were found by others (Scott et al., 2007; Compton et al., 2008) but some put a greater degree of emphasis on high anticipated income (Scott et al., 2011). Indeed, anticipated income is typically associated with prestige, and has itself been shown to be a consistently negative predictor of primary care preference (Colquitt et al., 1996; Clinite et al., 2013). However, while Kassebaum and Szenas (1994) also noted that prospective generalists rated "income prospects" lower than other specialty groups, they also found that *all* groups rated it relatively low, suggesting that other aspects of prestige (job security, perceived status) may play a more influential role on career decisionmaking.

The lifestyle afforded by a particular specialty is a driving force underlying student's specialty preferences (Schwartz et al., 1989). For example, Dorsey, Jarjoura, and Rutecki (2005) found compelling evidence that perceptions of specialty lifestyles is *the* principal factor explaining current medical students' specialty plans. Moreover, many recent studies have found it to be a distinguishable trait between those with a preference for primary care and those with other career preferences. (Kassler, Wartman, & Silliman, 1991; Clinite et al., 2014). This trend was confirmed by Scott et al., (2011) where prospective family physicians aspects of a medical lifestyle, such as flexibility of hours and career options higher than others, while a qualitative study of perspective family

physicians confirmed these same findings (Scott et al., 2007). Additionally, Hauer et al. (2008) found that personal/professional satisfaction and lifestyle was more highly rated by prospective internists than those with other preferences for other specialties. Looking at all primary care specialties, Clinite et al. (2013) showed that those with a greater preference for primary care reported a greater interest in aspects of the physician lifestyle such as spending time with their family, having a good work/personal life balance, and having free time outside of work. However, Kassebaum and Szenas (1994) found a different trend altogether, in particular, that those with a preference to become a generalist rated predictable working hours, demand of time and effort, and lack of stress as less influential than those with other specialty preferences. On that note, Lambert and Holmboe (2007) took the definitions from Schwartz et al. (1989), deeming primary care specialties as having "uncontrollable lifestyles" and found that both men and women were actively moving away from these specialties from 1990-2003 to specialties with more stable lifestyles. Executing a similar analysis, Dorsey, Jarjoura, and Rutecki (2005) found that "by all measures examined, controllable lifestyle is an increasingly important factor in the specialty choices of both women and men" (p. 793) and later note that this could, in part, be explained by the general trend of students eschewing primary care career paths.

These findings around lifestyle predictors may suggest that future physicians are putting a greater emphasis on the kind of daily life the specialty affords, but the overall trend based on this review remains unclear. It is important to note that the different trends found in these studies could be explained by the fact that the *perceptions* of specialty characteristics were studies, and are rooted in the student's own experiences and impressions. To that end, the findings of Clinite et al. (2013) showed that amongst *matriculants*, "the importance of the five domains that defined a good lifestyle differed very little by interest in practicing PC" (p. 4), suggesting that the importance of lifestyle perceptions appear to evolve over a student's medical education. Phillips, Weismantel, Gold & Schwenk (2009) also found negative primary care lifestyle perceptions evolve during medical school, saying "although medical school does not create these negative views of primary care work life, it may reinforce them. The gap between perceptions of primary care and specialist work life appears to increase slightly over the course of medical school." (p. 11). Thus, it is reasonable to anticipate that lifestyle factors would play a differential role in explaining specialty preference at matriculation versus graduation.

In sum, this review of the literature demonstrates an overall consistency with previous reviews (Bennett & Phillips, 2010; Campos-Outcalt, et al., 1995, Lawson & Hoban, 2003; Senf et al., 2003; Reed, Jernstedt & Reber, 2001; Meurer, 1995; Bland, Meurer, & Maldonado, 1995) and shows the demographic and psychosocial predictors of career preferences. Specifically, the student's gender, parental background, career motivations (social responsibility/altruism, research and technology interests, anticipated income and prestige) as well as lifestyle predilections (flexibility of hours, demands, stress) appear to be justified in an investigation of specialty preferences and its evolution during their medical education. Importantly, the student's race/ethnicity was not found to be consistently associated with primary care specialty, however its inclusion is justified
in light of current interest in attracting underrepresented populations to the medical workforce (Marrast, Zallman, Woolhandler, & McCormick, 2014; Nivet, 2011).

Before continuing, two limitations of this review should be heeded. First, it should be noted that a large set of covariates—at the personal and institutional level—have intentionally been omitted in this review as they are outside the scope of this dissertation's research agenda. Indeed, while recent research has shown, for example, that a desire to practice in a rural location has historically strong association with preference for primary care (Click, 2013), this relationship has not been reviewed here. Other factors such as educational debt (Colquitt et al., 1996) or MCAT scores (Jeffe, Whelan, & Andriole, 2010) also play a role in how specialty preferences, but are similarly beyond the scope of this dissertation. Finally, a large body of literature focuses on institutional factors such as ownership, curricula, school mission, and faculty composition (e.g. Bland, Meurer, & Maldonado, 1995), and are recognized in this dissertation as critical, but ancillary, to the analysis.

Second, this review has focused on the important factors related to specialty *preferences* or *plans* during the medical school experience and not actual specialty match, choice, or career attainment. All of these provide important—albeit different— evidence about how medical students move through the physician pipeline. While the key variables previously mentioned are likely to influence their goals, preferences, and perceptions, this review of the literature only provides information up through graduation, and the trends mentioned are not expected to transfer to other segments of the physician pipeline. In

sum, it is sensible to heed the advice of Babbott et al. (1988) who state that specialty preferences "represent prospective certification plans that have yet to be validated by actual certification statistics" (p. 1975).

Placing this Study

As noted earlier, there is a substantial corpus of work related to career interests at graduation—and this is a very important measure indeed—but there exists an opportunity to replicate the methods and variable selections of these studies when looking at career interests at matriculation *as well as* graduation. Thus, if it is the case that students' career motivations and desired lifestyles evolve during various stages of their training, it is important to look at how predictors are uniquely associated with these outcomes. As *Table 2.1* demonstrates, there is a gap in the literature as most studies *only* investigate preferences at graduation (e.g. Hauer et al., 2008; Scott et al., 2011), or *only* investigate preferences at matriculation (Clinite et al., 2013), or investigate both time points with limited methods (Compton et al., 2008).

Additionally, it is sensible to ask if the same goals, motivations, and lifestyle predilections found in these studies are still germane to today's medical student, and whether they have the same impact on specialty preferences (Eckleberry-Hunt & Tucciarone, 2011), and these trends must be evaluated within an evolving social and demographic landscape. To sum the rationale for this study, I borrow from the conclusion of Clinite et al. who state that "understanding the role of lifestyle in students' specialty considerations at the beginning of medical school provides an opportunity for comparing

the role of lifestyle later, such as when they select their specialty at the end of medical school." (2013, p.6). This goal of this dissertation is to fill this gap in the literature with a robust analysis of recent data, and to provide an empirical test of the Bennett-Phillips specialty-choice conceptual model. Doing so may help uncover important linkages between physician recruitment, admissions, or training processes and the current primary care workforce issues.



Figure 2.1. Bennett-Philips conceptual model of the process of primary care specialty choice. Reprinted from *"Finding, recruiting, and sustaining the future primary care physician workforce: a new theoretical model of specialty choice process"* By K. Bennett and J. Phillips, (2010), Academic Medicine, 85(10), p. S84. Copyright 2010 by Wolters Kluwer Health Inc. Reprinted with permission.



Figure 2.2. Conceptual model of specialty plan-making among medical students.

Table 2.1.

Study	Preferences at					
	Pre Matriculation	Matriculation	1st Year	2nd Year	3rd Year	4th Year/ Graduation
Babbot et al. (1988)		X				Х
Clinite et al. (2013)		Х				
Clinite et al. (2014)			Х			Х
Colquitt et al. (1996)		Х				Х
Compton et al. (2008)		Х			Х	Х
Hauer et al. (2008)						Х
Jeffe, Whelan, & Andriole (2010)						Х
Kassebaum & Szenas (1994)						Х
Lambert & Holmboe (1996)						Х
Scott et al. (2007)		Х	Х	Х	Х	Х
Scott et al. (2011)						Х

List of Seminal Studies and Timing of Specialty Preference or Plans Measurement.

Note. X signifies the when the students were surveyed.

CHAPTER 3: DATA SOURCES AND METHODOLOGY

Chapter Overview

In this chapter, I demonstrate how I will answer a series of quantitative research questions aimed at understanding how psycho-social predictors are associated with specialty preferences at matriculation and graduation. These research questions are:

- What are the career and lifestyle motivations reported by matriculating medical students? And, are there differences between those with a matriculating preference for primary care specialty, those without a specialty preference, and those with a preference for a non-primary care specialty?
- Do these career and lifestyle motivations and other psycho-social factors differentiate those with a preference for a primary care specialty, those with an unspecified specialty preference, and those with a preference for a non-primary care specialty *at matriculation into* medical school?
- Do these career and lifestyle motivations, other psycho-social factors, and previous specialty preferences, differentiate those with preference for a primary

care specialty, those with an unspecified specialty preference, and those with a preference for a non-primary care specialty *at graduation from* medical school?

I begin this chapter by describing the datasets. I then describe respondents that were included, along with some exclusion criteria that were used to delimit that dataset. I also provide a rationale for the data elements that were explored in this dissertation and show my analytic strategy for each of these research questions.

Data Sources

All data used in this dissertation comes from the Project TrEMUR (Transitions in the Education of Minorities Underrepresented in Research), a large, mixed-methods study investigating the educational trajectories of biomedical professionals, specifically looking at the factors associated with decision-making during educational transitions. This investigation has four primary investigators (Robert Tai and Heather Wathington at the Curry School of Education at the University of Virginia as well as Donna Jeffe and Dorothy Andriole at the Washington University School of Medicine) and is funded by the National Institutes of Health (NIH-NIGMS 1 R01 GM094535-03). The qualitative data was undertaken by administering semi-structured interviews to 217 biomedical professionals (MD students, MDPHD students, as well as medical school and biomedical faculty members) from 2011 to 2013. These data, however, were not used in this dissertation. The institutional review board for social and behavioral sciences at the

University of Virginia and institutional review board at Washington University School of Medicine approved this study.

Forming the quantitative portion of Project TrEMUR, the primary investigators worked with the AAMC to secure the release of de-identified, student level data from their Student Records System (SRS), the PMQ, the MSQ, and the GQ for all PMQ respondents from 2001-2006. By combining these surveys using a unique student identifier into a single datasheet, it has been possible to explore the educational choices of medical students at multiple points in their educational trajectory. Additionally, there is a rich set of attitudinal factors in these surveys that allows us to formulate hypotheses regarding the rationale for these educational choices, and have allowed for a recent crosssectional analysis of MD-PHD matriculants (Jeffe, Andriole, Wathington, & Tai, 2014). Although Project TrEMUR primarily deals with trends amongst biomedical *researchers*, this large dataset also allows us to understand the career choices of the general medical student as well. I describe each of these datasets in full below.

SRS (Student Records System)

The AAMC's division of Enrollment Services uses a secured, regularly updated, web-based data warehouse called the SRS to collect and maintain student-level information of all medical students at LCME-accredited medical institutions, tracking them from matriculation to graduation (AAMC, 2015d). The SRS includes biographic information (gender, race/ethnicity) as well as programmatic information (year of matriculation, most recent status, graduation date), and is intended to be used by researchers to study enrollment and graduation trends.

PMQ (Post-MCAT Questionnaire)

The PMQ is a secured, web-based questionnaire administered each year to those who register for the MCAT (Medical College Admissions Test); its completion is voluntary. For the years included in this dataset, the PMQ was known as the *Pre*-MCAT Questionnaire, as students received it prior to their MCAT test date; however, it is now known as the *Post*-MCAT Questionnaire, as students complete it after completing the MCAT (AAMC, 2015b). The data provided by the PMQ is designed to allow researchers to understand the past experiences of potential medical students as well as their future career plans. The TrEMUR dataset includes PMQ data from 2001 to 2006.

MSQ (Matriculating Student Questionnaire)

Every year from June to September, first year medical students are invited by email to participate in the MSQ, a survey that collects information on students' premedical experiences, medical school selection processes, career interests, and educational debt (AAMC, 2015c). Their information is identifiable but confidential. These data are designed to be used both by researchers for investigational purposes as well as by medical school leadership to tailor their programming to the needs of their incoming students. While response rates vary from year-to-year and by medical institution, in 2013 the response rate was 72.4% (14,888/20,555) indicating that the data represents most LCME-accredited medical school matriculants (AAMC, 2013a). The TrEMUR dataset includes MSQ data from 2001 to 2011 for all respondents who completed the PMQ from 2001 to 2006.

GQ (Graduation Questionnaire)

Finally, the GQ is administered annually from February to June to all graduating medical students, and is completed online and on a voluntary basis (AAMC, 2015a). The information collected—such as clinical experiences, financial aid, career intentions—is identifiable but confidential. It is intended to be used for research and evaluation purposes, in order to identify issues of medical education in light of national policy goals (AAMC, 2014a). In 2013, the completion rate was 82.4%, again demonstrating the high representation of this survey. The TrEMUR dataset includes GQ data from 2005 to 2011 for all respondents who completed the PMQ from 2001 to 2006.

Data Screening and Study Respondents

The TrEMUR dataset included 262,815 records, though not all of them could be included in this analysis, which necessitated a comprehensive data screening process. This is graphically detailed in *Figure 3.1* and all syntax can be found in *Appendix A*.

First, two duplicate students were eliminated. Second, I created a variable to ensure that the respondent had matriculated into a medical school, which eliminated 162,700 respondents who had not. (It was not clear if they had not applied, were not

accepted, or did not enroll in an LCME-accredited medical school in this timeframe.) Then, I eliminated 27 respondents whose matriculation date was outside of the year 2001-2011, as well as 3,948 respondents who matriculated into MD/PhD programs because their timeframe is categorically different and because they might have categorically different motivations for pursuing a medical degree (Jeffe, Andriole, Wathington, & Tai, 2014), and therefore they might have unique specialty preferences (Brass, et al., 2008). However, other dual degree students (e.g. MD/MPH) were retained. Finally, I specified a variable for medical school graduation and eliminated 10,282 respondents who did not graduate from medical school, the majority of which matriculated in 2008 or after, and would not be reasonably expected to have graduated by 2011. Additionally, I eliminated 59 respondents who did had graduated, but not between the years 2005-2011. Based on these exclusion criteria, I identified 85,797 respondents as M.D. graduates of a U.S. based, LCME-accredited institution.

As stated above, however, the completion rate for the MSQ and GQ surveys is not 100%, and therefore further data screening was required. To be specific, the first for inclusion in this investigation, each student must have completed the MSQ between the years 2001 and 2011 as well as the GQ between the years 2005 to 2011. When cross tabulating the completers versus non-completers for each survey, I found that 12,132 (14.1%) completed neither survey, 20,268 (23.6%) had completed *only* the MSQ, 14,020 (16.3%) had completed *only* the GQ, and that 39,377 (45.9%) had completed both questionnaires. This last group is the population of interest for this dissertation.

Indeed, of these 85,797 medical school graduates, only 53,397 (62.3%) had completed the GQ in that timeframe. However, the reason for this lower-than-expected completion rate is one of timing: the basis for inclusion in the TrEMUR quantitative dataset was that the respondent had completed the PMQ from the years 2001 to 2006. Therefore, it is possible that many PMQ respondents in this timeframe had delayed their matriculation, and had not graduated by 2012. Or, it is possible that students had completed the PMQ in the year 2007 or after, or had not completed the PMQ at all, and would not have met the criteria to be included in this dataset.

To determine if completion of the survey was associated with a demographic bias, I examined the profiles of those who completed the GQ and those who did not. I found a small gender difference between GQ completers and non-completers as females (63.6%) were slightly more likely to complete the survey than males (60.4%). Additionally, there were differences in completion rates by race/ethnicity are apparent. For example, White students had the highest completion rate (65.0%), while Pacific Islanders had the lowest completion rate (49.7%), and the other race categories had relatively similar completion rates (56.3% to 59.2%). I repeated this demographic comparison between MSQ completers and non-completers. I found that females (71.3%) were again more likely than males (67.7%) to complete the MSQ. Native American/Native Alaskan (74.6%) and White students (71.5%) had the highest completion rates, while all other races/ethnicities had lower completion rates (63.7% to 69.8%).

At this point, I eliminated the GQ and MSQ non-completers (46,420/ 85,797, 54.1%), and noted that White students will be slightly overrepresented in this study's sample compared with other races/ethnicities; additionally, females will also be slightly overrepresented compared with males.

Having retained the 39,377 respondents with valid GQ data, it was also important for the analysis to identify the gender, race/ethnicity, parental occupation, factors for pursuing medicine for each respondent as well as the two outcomes of specialty preference and matriculation and specialty preference and graduation. The gender, race/ethnicity, parental occupation for each respondent as these came from the PMQ and SRS (data availability was 100.0%), but given that the factors for pursuing medicine and specialty preference at matriculation came from the MSQ, 231 respondents had partially missing data and that 7,721 respondents of the 2007 MSQ did not have the option to respond to the scale. As these factors represented important psychosocial data, I eliminated any respondent without complete data.

In sum, the sample population for this study includes 31,425 individuals who completed the PMQ from 2001-2006, completed the MSQ from 2002 to 2006 and completed the GQ from 2005 to 2011. To that end, all respondents in the study sample had complete data for the analysis and can be considered graduates from an LCME-accredited medical school.

Outcome Variables

Because this study deals with respondent's career plans at two points in time, it was necessary to investigate each outcome separately. In the following sections, I demonstrate how I operationalize student preferences at matriculation (from MSQ data) and at graduation (from GQ data).

Specialty Preference at Matriculation

In the MSQ, matriculants were asked if they were planning to become certified in a specialty. Of the 39,911 respondents in the sample, 60.4% (18,970) responded "Yes," 37.4% (11,749) reported that they were "Undecided," and 2.3% (706) reported that they were not planning on becoming certified in a specialty. For the purposes of this analysis, this last group was combined with the "Undecided" respondents. A similar grouping strategy has called this group the "no board-certification group" (Jeffe, Whelan, & Andriol, 2010). Only the respondents who indicated that they were planning on becoming certified received a follow up question asked them to indicate in which particular specialty they were planning to specialize from a list of specialties. *Figure 3.2* shows the survey prompt for the 2005 MSQ.

From the list of specialties, I recognized family medicine, internal medicine, and pediatrics as primary care specialties as consistent with the most common definition (Starfield, Shi, & Macinko, 2000; Rosenblatt & Andrilla, 2005; Clinite et al., 2014; Schwartz et al., 2005). Of the 31,425 respondents in the sample, 22.6% (7,103) could be

recognized as having a preference for a primary care career, 37.8% (11,867) could be recognized as having a preference for a non-primary care career, and 39.6% (12,455) reported no specialty preference at matriculation. These are the three important categories that serve as the outcome for this dissertation. Descriptive statistics comparing these three groups (along with a breakout for each of the three primary care specialties) can be found in the next chapter in *Table 4.1*.

Specialty Preference at Graduation

In the GQ, graduates were also asked if they had plans to specialize. From the same sample, the numbers reveal that a greater number graduates reported a plan to specialize than when they were matriculants, suggesting that career intentions had been influenced, shaped, or strengthened during medical school. In regards to having a specialty plan, 84.7% (26,604) responded "Yes," 9.1% (2,846) reported that they were "Undecided," and 6.3% (1,975) reported that they were not planning on becoming certified in a specialty. For the purposes of this analysis, this last group was combined with the "Undecided" respondents. Again, if respondents indicated that they were planning on becoming certified, a follow up question asked them to indicate in which specialty they were planning to specialize. See *Figure 3.3* for the GQ survey prompt from the 2014 survey.

Using the same criteria mentioned above, I categorized family medicine, internal medicine, and pediatrics as primary care specialties. However, it is important to note that the response set of the GQ had a greater number of specialty options than that of the

MSQ; additionally, a follow up question asked "If you are planning on specializing in Family Medicine, General Internal Medicine, Internal Medicine/Pediatrics or Pediatrics, do you plan on going into a subspecialty in that field?" (AAMC, 2014a, p. 31). Because Internal Medicine and Pediatric subspecialties have different training pathways and serve their populations in a categorically different way than generalists (American College of Physicians, 2014; Council of Pediatric Subspecialties, 2014), we decided to only include the graduates who reported no interest in subspecialization in these fields as having a primary care preference. However, this distinction was not assumed between family medicine generalists and subspecialists since there are relatively few subspecialties for this field (Adolescent, Geriatric, Hospice and Palliative, Sleep, and Sports Medicine), few AGCME-accredited training positions for family medicine subspecialties, and a small number of residents planning such subspecialization training (American Board of Medicine Specialties, 2014; American Medical Association, 2014b).

Thus, those with a preference for primary care constituted 23.6% (7,409) of the total sample, while students who reported that they were undecided or gave no response represented 15.3% (4,821) of the total sample, and 60.4% (19,204) reported a preference for a non-primary care specialization. Again, these are the three important categories that serve as the graduation outcome for this dissertation. Descriptive statistics comparing these three groups (along with a breakout for each of the three primary care specializes) can be found in the next chapter in *Table 4.2*.

Predictor Variables Predictive of Specialty Choice

In order to best explain the patterns of specialty preference, it is critical to include variables that have a demonstrable association with that outcome which might moderate relationships.

In the literature review, I demonstrated that females were more likely than males to indicate a preference one of the primary care specialties, especially pediatrics. Indeed, while the direction of the predictor was generally consistent across all three primary care specialties, their magnitude was not and suggests that females may matriculate with a differential preference for individual primary care specialties.

I drew the student's gender from the SRS and use the variable name 'Female' in this dissertation. When dealing with relative-risk-ratios (the main analytic output) a value *greater* than one indicates that females had greater fitted odds than males in regards of indicating a primary care specialty preference. 51.5% (16,220/31,425) of respondents self-identified as female and 48.5% (15,205/31,425) self-identified as male.

As shown in the literature review, there is no strong trend showing that race or ethnicity consistently predicts whether or not a matriculation or graduated will prefer a primary care specialty as a career choice. However, combining races and ethnicities into sensible groups may allow for broad inferences to be made about the medical student population. For example, while some of the seminal studies I reviewed considered all races/ethnicities as separate (e.g. Colquitt et al, 2010), others (e.g. Jeffe, Whelan, & Andriole, 2010; Hauer et al., 2008) showed that combining students of races/ethnicities historically underrepresented in medicine (Native American/ Native Alaskan, Hispanic, Black) had predictive validity of *not* preferring a career in general internal medicine, family medicine, or general pediatrics.

Following this grouping strategy, I drew the student's self-reported race/ethnicity from an eight-category variable found in the SRS which included White, Black, Hispanic, Asian (including Southeast Asian), Pacific Islander, Native American/Native Alaskan, Other/Multiple races, and No response. Then, I categorized students as:

- ✤ 'White,' including only students who self-identified as White;
- ✤ 'Asian,' including those who self-identified as Asian or Southeast Asian;
- 'Underrepresented Minorities (URM),' including those who self-identified as Black, Hispanic, or Native American/Native Alaskan, and Pacific Islanders; and,
- 'Other,' including those who self-identified as Other, Multiple races, and those who did not have a response.

A cross-tabulation of these race/ethnicities and categories is found in *Table 3.1*. I then created a dummy variable for each of these groups in order to treat them as categorically different; for analyses, I chose White as the reference group as they represented the most students and thus can be considered a kind of normative category.

While some studies in the academic medicine literature treat Pacific Islanders as Asians in such analyses (Andriole, Whelan, Jeffe, 2008; Jeffe, Yan, & Andriole, 2012), I chose to categorize them as underrepresented minorities based on a recent report from the AAMC's Analysis in Brief, which showed that "physicians of American Indian, Alaskan Native, or Hawaiian or Pacific Islander; black or African American; or Hispanic/Latino origins are more likely to practice primary care than white physicians, which suggests that increasing representation from these groups within the physician workforce may increase the number of physicians practicing primary care" (Xierali, Nivet, & Fair, 2014, para. 6). Additionally, the AAMC states that "underrepresented in medicine' means those racial and ethnic populations that are underrepresented in the medical profession relative to their numbers in the general population" (AAMC, 2004, p. 1). Looking at the U.S. matriculating class of 2010, 0.2% (103/42,741) were Pacific Islanders, (AAMC, 2012b), and census data shows that "Native Hawaiian and Other Pacific Islander" also represent 0.2% (540,013/308,745,583) of the U.S. population (Humes, Jones, & Rameriz, 2011). Conversely, Asian students represented 20.6% (8,787/42,741) of the U.S. matriculating class of 2010, but comprise only 4.8% (14,674,252/308,745,583) of the U.S. population. Therefore, while Pacific Islanders have reached proportional representation, it is clear they have categorically different matriculation patterns than other Asian students and should not be combined into one pan-Asian group. That said, the number of Asian Pacifiers in the final sample was 0.2% (51/31,911), suggesting that any categorization of Pacific Islanders is unlikely to change the results.

At any rate, any discussion regarding racial/ethnic composition of medical students or physician must also consider recent work on how to best define underrepresentation in medicine according to local contexts (e.g. Page, Castillo-Page, Poll-Hunter, Garrison, & Wright, 2014; Nivet, 2011), such changes to how racial/ethnic characteristics contributes to one's underrepresentation status is likely to evolve. In essence, the racial categorization for this dissertation represents my best estimation for how to analyze *current* national trends, and not necessarily future or local patterns of racial or ethnic underrepresentation.

The MSQ asks respondents to "indicate how important the following factors were in your choice of medicine as a career goal," using a 22 item battery. *Table 3.2* shows all of the items prompts, the variable names that will be used throughout this dissertation. These items represent a broad spectrum of important motivations related to student's interest in medicine. For example, Jeffe, Whelan, & Andriole (2010) used this battery to create an innovation and research factor (six items), a social responsibility factor (five items), and a prestige factor (four items); all of these factors were reliable. However, the authors used only 15 of the 22 items, potentially indicating that other important scales could be derived from the remaining items. In this dissertation, these items are used to create "motivational factors." This term was preferred not only for consistency with the literature, though I note their conceptual similarity to those termed "attitudinal factors" (Scott et al., 2011, Jeffe, Whelan, & Andriole, 2010) or "perception factors" (Hauer, et al., 2008). This battery used a five-point Likert scales from 0 ("Not at all important"), 1 ("Slightly important"), 2 (Somewhat important"), 3 ("Moderately important") and 4 ("Very important"), and all of the items were scaled in the same direction. A five-point scale is commonly considered to have the lowest number of responses to be considered as continuous, and I proceeded as such (Preston, & Colman, 2000). Observing the univariate histogram for each item, most items were found to be approximately normally distributed on this five-point scale with the notable exception of the variable 'Difference' which 90.0% of respondents responded "Very important."

Using response data from the PMQ, I developed a variable related to the parents' occupation role in order to account for any potential socializing differences between physician-parents, professional-parents, and parents with a different education. The prompt asked respondents to "Please indicate your mother's occupation during the major portion of the past year" and "Please indicate your father's occupation during the major portion of the past year."

From the response set, I created three categories of parental occupation for both the student's mother and father. Physicians represented the category "Physician (M.D., M.D./Ph.D. or D.O.)." Professionals were represented by choices including postbaccalaureate education such as lawyers, scientists, or parents with graduate work, professional school, or a doctorate. Importantly, this also included other healthcare professionals such as veterinarians, dentists, nurses, pharmacists, et cetera. Finally nonphysician or non-professional parents represented all other categories, included the responses "Do Not Know" and missing data. Only 750 (2.4 %) of mothers' and 747 (2.3%) of fathers' data was missing, and was not found to be systematically biased. Because the most common group was non-physician or non-professional, I chose this as the reference group for all analyses. *Table 3.3* shows a cross-tabulation of mother's and father's occupation.

Once student responses for their mother's and father's education had been categorized, I created variable to combine the two responses called "Highest Parent Education." This allowed me to take the same category types, and establish whether the highest education of either of the respondents' parents was a *physician*, a *professional*, or *non-physician/professional*. It is sensible to think of a professional as having "more" education than someone without a professional degree as it requires additional years of scholarship, however, I decided that all students with at least one physician-parent should be included in the *physician* category, as there might be a unique parent-child dyad that would influence their children's career decision-making in a systematic way.

Providing the clearest perspective on this issue, Lentz and Leband (1989) theorize that children of physician-parents are likely to receive intergenerational knowledge transfers, as career relevant information is informally passed to the child. Thus, children of physicians *may* have less difficulty navigating the application process and may receive some first-hand advice of how to choose a specialty. Additionally, physician-parents may encourage their children to enter a particular specialty as "physicians in practice may hold particularly negative perceptions about generalist–primary care careers, and they may convey these perceptions to their children" (Colquitt et al., 1996, p. 955). Thus, even if other parents can provide the same kind of experiences and formal educational opportunities, it is presumed that physician-parents can pass professional knowledge to their children or socialize them to have particular career interests.

Based on this coding schema, 16.9% (5,319) of respondents were deemed to have *Physician-Parents*, 39.1% (12,278) had *Professional-Parents*, and 44.0% (13,828) to have *Other Profession-Parents*. This last category also represented the 233 students whose parents' occupation could not be identified. From these categories, I created three dummy variables called 'Physician-Parent,' 'Professional-Parent,' and 'Other-Parent.'

As we focused on the socializing impact of medically-oriented experiences that students might have received prior to matriculating into medical school, we found that one PMQ item was directly relevant to this study. This item asked students to "mark any category below that describes your extracurricular activities and/or work experiences," and we focus on the response option "paid or volunteer work in hospitals, medical clinics, or labs" for its theoretical importance and potential contribution for understanding how early interacting with the medical community may play a role in long-term career plans.

While this variable may represent a "humanist" experience—where the student interacted directly with community members—the research reviewed in Chapter 2 provided no indication of how this variable might interact with the outcome; however, it is sensible to hypothesize that those with early contact with the physicians, their role, and their place of work would help solidify student career decision. Therefore, students who experience work-life in a hospital during college can be thought to be more likely to have a preference—be it for a primary care or non-primary care specialty—than to be undecided. Importantly, no follow-up question asked them to detail the kind or nature of the experience, therefore we cannot assess in which capacity the student worked, their interaction with physicians and patients, or their subjective experiences. As seen in *Table 3.4*, 70.9% responded that they had experienced some hospital work during college, while 9,145 (29.1%) responded that they had not. As there were only two options, this variable was dummy coded and called 'Hospital Work.'

We also wanted to focus on other experiences that may facilitated a career in medicine, such as involvement in scientific research during college. We drew data for this experiential indicator from the MSQ item asking students to specify whether or not they had been involved in a "laboratory research apprenticeship for college students." Again, we do not know the nature of these experience; however, the review of the literature does suggest that those with a proclivity towards or interest in research are more likely to prefer non-primary care specialties (Kassebaum & Szenas, 1994; Scott et al., 2011; Jeffe, Whelan, & Andriole, 2010), though we should note that proclivity towards research is of a different nature than having a research experience. That said, it appears sensible to include an indicator or college research experiences, especially laboratorybased experiences—in this analysis. This variable was binary, and 16,489 (52.5%) indicated that they *had* some kind of college laboratory-based research experience, while 14,936 (47.5%) indicated that they did *not* have this kind of experiences. This variable was dummy coded and called 'College Research.'

Finally, for the third research question, which looks at specialty preferences at graduation, it was necessary to control for prior interest in a specialty, which was shown to significantly predict future career plans (Colquitt et al., 1996; Babbott et al., 1988; Compton et al., 1988; Scott et al., 2011). Thus, the outcome variable 'Specialty Preference at Matriculation' listed above as an *outcome* will serve as a predictor in the final research question, where its categorical operationalization remains: primary care, undecided, and non-primary care.

Research Strategies

The analysis for all three research questions was undertaken in Stata 13.1 (StataCorp., 2013) with the exception of the confirmatory factor analysis, which was undertaken in MPLUS 5 (Muthen & Muthen, 2011). Annotated syntax is provided in *Appendix A*.

Research Question #1: Motivational Factors

This research question deals exclusively with the 22 item battery from the MSQ that asks students to rate factors that lead them to choose medicine as a career goal. Importantly, while the data these 22 variables provide are theoretically meaningful to the outcomes in the subsequent research questions, it is preferable to reduce the number of items to a reasonable number of so called "motivational factors." To accomplish this according to best practices (e.g. Thompson, 2004; Reis & Judd, 2000), a confirmatory approach followed an exploratory one.

By its definition, Likert data are ordinal, however as previously mentioned, there is conceptually validity in treating it as continuous and linear, provided there are at least five categories and the data does not depart from a normal distribution (Preston, & Colman, 2000). The most important univariate assumptions are normality and outliers (Stevens, 2012), which were assessed statistically (through skewness and kurtosis statistics). The skewness and kurtosis of each item were evaluated to be normal if they were in the range -2.0 to +2.0 (George & Mallory, 2010). I chose not to use a formal normality test such as Shapiro-Wilk or Kolmogorov–Smirnov as these would be too sensitive for such a large sample size and such a small response set; however I inspected each item's histogram to determine how the data might depart from assumptions of normality. At this point, items with serious univariate violations were excluded.

All respondents were then randomly split into two equal cohorts to allow for a "split-sample design" where one group's data is analyzed for a *possible* factor structure, and the second is used to confirm or reject that possible structure. This is considered to be a sensible approach for testing exploratory models, as exploratory models suffer from factor indeterminacy (Reis & Judd, 2000). I assessed group differences to determine if the random sorting was biased.

For the first randomly selected respondents, I applied an exploratory factor analysis (EFA). Creating a more parsimonious structure reveals their unifying themes, conserves degrees of freedom, and reduces the likelihood of a Type II error (Thompson, 2004, p. 5). EFA was preferred over principal components analysis (PCA) because the review of the literature suggested that common themes (latent "motivational factors") would emerge. Then, based on the factor structure that emerged from the EFA, a *confirmatory* factor analysis (CFA) was undertaken on the second sample in order to assess that the exploratory structure fit the data (Muthen, 2003). From the final CFA model, I created the motivational factor scores which were used in the predictive models in *Research Question #2* and *Research Question #3*.

Exploratory Factor Analysis. The process of applying EFA to the data necessitated a testing of multivariate assumptions, especially in regards to potential for factorability. First, I created bivariate cross-tabulations between variables with apparent similarities to see if there were consistent response patterns. Second, I inspected the interitem correlation matrix to see if there were a number of large correlations, indicating that the items would be suitable for factorization; generally there should be some values greater than 0.5. Third, I inspected the Kaiser-Meyer-Olkin (KMO) statistic for each item to ensure that the sample is of sufficient size to factor well and the resultant factors will have informative power (Hutcheson and Sofroniou, 1999). The item level and overall KMO score should be above 0.6; however the large sample size immediately suggested that this would not be an issue. Finally, the Bartlett Test for sphericity tests the null hypothesis that the correlation matrix is an identity matrix; in other words, the non-zero correlations are due to sampling error (Bartlett, 1950). Thus, if the test is significant null hypothesis is rejected, then patterns are expected to be due to some underlying reason rather than chance. Further, this suggests potential for reducing the number of dimensions.

Once questions of univariate and multivariate were addressed, and any significantly inoperative item was eliminated, I proceeded with the EFA procedure using a principal components factorization (PCF) extraction method in order to assess the number of factors that should be retained. I chose this method because it is generally used as a data reduction technique, and it was important to provide an indication of the nature of the factors underlying reason for choosing medicine as a career. As first step, I confirmed the number of factors retained using the following decision rules:

- First, eigenvalues should have a value greater than 1.00 (Kaiser, 1958), since they represent the amount of variance accounted for by factor, which is itself a latent representation of multiple items. (In other words, if the eigenvalue is *less* than 1.00, then the factor is capturing less variance than the original variable would have.)
- I also visually assessed the scree-plots (Catell, 1966), which shows the rank of the eigenvalues of the principal components (large to small) and the actual eigenvalue as a scatter plot. By inspecting at the slope from the point-to-point, looking for changes from very steep slopes to more gradual ones ("elbows"), one can demonstrate a subjective delineation between meaningful factors and trivial ones.

- Horn's parallel analysis (Horn, 1965; Dinno, 2009), a Monte Carlo simulation of uncorrelated data, can also be used to determine the number of factors to retain. If the ordered eigenvalues from the principal components extraction is greater than its ordered pair from Horn's analysis, then this factor was retained. In general, this simulation is agreed to be more rigorous than the previous Kaiser's or Catell's method (Ledesma & Valero-Mora, 2007), and thus it will be used to determine the number of factors if there is any ambiguity.
- It is also important to assess the amount of variance accounted for by each factor. Although no commonly agreed-upon rule has been established, the cut-off for a meaningful factor was set at 0.05, or 5% of the total variance. Similarly it is important that the cumulative amount of variance accounted for by all retained factors should be substantial.

Once the number of factors to be retained had been determined, I explored the various orthogonal rotations options (varimax, quartimax, equamax, and parsimax) as well as oblique rotation options (promax, oblimax, and quatimax) in order to assess the best fit to the data, and ensuring that the solution converged and was not a Heywood case. This necessitated an iterative approach, as the inclusion or exclusion of some items may influence how other load onto factors. Thus, a first decision rule for selecting the final solution was that each item should load at the 0.45 level on one and only one factor; this level was considered appreciable for this study, although many deem 0.3 or 0.35 as the minimum value to be considered as appreciable. If items were not appreciably loading, or were cross loaded, they were temporarily eliminated, and the model was specified with

the reduced number of items. I also explored the correlations between factors from the oblique models and looked for values greater than .32, which would indicated that there is at least 10% shared variance between the factors, and would favor an oblique rotation over an orthogonal one (Tabachnick & Fiddell, 2007).

After this iterative process revealed a preferred simple structure, I assessed each factor in two ways. First, each factor's reliability must be greater than 0.70 using Cronbach's alpha (Nunnally & Bernstein, 1994) in order to indicate that each item is measuring a similar construct. Second, each factor should have a sensible interpretation consistent with the literature or guiding theory; in essence, the solution should result in a "simple structure" that facilitates interpretation of the factors. Co-authors were encouraged to provide insight and interpretation of the factors. Once I was satisfied that the factors had met these two final criteria, I gave each of them an appropriate name and created a factor score using the regression method (Thompson, 1951). Finally, because all respondents had been screened for complete data, no missing data techniques were necessary.

Confirmatory Factor Analysis. The simple structure of motivational factors that emerged from the exploratory factor analysis was a data driven approach; therefore, a theory was produced from the data provided by randomly chosen respondents in the EFA analysis. The purpose of a confirmatory factor analysis (CFA) was to evaluate whether this hypothesized motivational factor structure fit the data provided by other respondents of the same sample. Since there are a host of initial factor solution, rotation and extraction methods in EFA, it was important to see if *our* simple structure was a suitable one. Because CFA is considered to be a more rigorous evaluation of a model's factor structure (Greenbaum & Dedrick, 1998; Van Prooijen & Van der Kloot, 2001; Bollen, 1989), it allows for a more robust assessment of how data and theory drive approaches compare. To do this, we followed a standard confirmatory factor analysis approach (Muthen & Muthen, 2009), where we put parameter restrictions on the model using the MLR option in Mplus 5 (Muthen & Muthen, 2007). This model specification was chosen because it uses the maximum likelihood estimator preferred for this method, but uses robust standard errors to account for non-normal data (Yuan & Bentler, 1998).

It has been recommended to use multiple measures of fit were to post-evaluate the model (Tanaka, 1993). Measures for model fit test include the χ^2 likelihood ratio test, the comparative fit index (CFI), the Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA). Based on the results of the EFA model, it is expected that the χ^2 likelihood ration test will be statistically significant; therefore, this indice does not reveal how the model fit the data. Bentler & Bonett (1980) have advised that for CFI and TLI should have a value of greater than or equal to .90 to indicate acceptable model fit. However, more recently Hu and Bentler (1999) have recommended a more conservative threshold of .95. Therefore, the CFI and TLI estimates range between 0 and 1 with higher value indicating better fit. For the purposes of evaluating the model, a slightly more liberal cutoff of .90 was established as acceptable fit, while .95 or higher was regarded as good fit. In contrast, smaller values indicate good fit for RMSEA estimates. There is different advice for thresholds of RMSEA have been recommended. Browne and Cudeck

(1993) suggest value of .05 or less; Hu & Bentler (1999) suggest that value of 0.06 or less indicate good fit; MacCallum, Browne and Sugawara (1996) have used 0.01, 0.05, and 0.08 to indicate excellent, good, and mediocre fit, respectively.

Analysis also included an inspection of the structural pathway coefficients between the observed and latent variables (Schrieber, Stage, Barlow, & King, 2006). This was done by examining the z-values for each of the pathways and looking for significance at the p < .05 level, which indicates a statistically significant factor loading. The standardized coefficients were also inspected, as standardized loadings greater than .70 indicate a strong relationships. Additionally, the residual covariance matrix (i.e. the table of modification indices) was inspected for any potential areas where model fit could be improved by correlating the residuals. However, I only specified a residual correlation if the observed variables loaded onto the same factor, otherwise the practical interpretation and theoretical soundness could be lost, and the chance for replicability in a different sample would be lost (Thompson, 2004). At this point I evaluated the changes between models using the χ^2 likelihood ratio test to determine if the modified model had statistically improved the model's fit (Schrieber et al., 2006); however, it is important to recognize that data-driven modifications have moved us from the entirely confirmatory to the slightly exploratory (Ullman, 2001). That said, these slight respecifications that improve the model without changing the fundamental nation of the model or its hypotheses is considered acceptable (Stage, 1990). Once the final model had been determined, motivational factor scores for all respondents were computed.

Factor scores. To explore group difference in a way that would sensitize the findings for the following analyses, I computed one-way ANOVA models for each of the motivational factors that emerged in order to compare the differences between the primary care, undecided, and the non-primary care groups at matriculation. Thus, there were three independent groups, and the significance level was set at α =0.05/k, where k is the number of emergent factors. Because it was assumed that a difference would appear between undecided and the other two groups, an *a priori* linear contrast was undertaken to see if how these groups differed, first by computing a linear contrast between the undecided and the primary care groups and then between the undecided and the primary care groups and then between the undecided and the primary care groups differed. Looking at the undecided versus other contrasts would indicate whether this group behaved closer the primary care or non-primary care groups when looking solely at one of the motivational factors.

Research Question #2: Specialty Preferences at Matriculation

In this research question, I explore the how the motivational factors and demographic profile at matriculation differs between the three groups of interest: those with a matriculating preference for primary care and those who were undecided. To investigate how the predictors are independently associated with the outcome of specialty preference, the requisite method must be able to account for the following two data structures. First, the method must account for discreet (categorical) outcomes with three groups (i.e. k>2), as there are three different outcome groups (Primary Care, Non-Primary Care, and Undecided). Second, it must deal with both categorical and continuous

predictor variables, as the attitudinal factors are treated as continuous while demographic variables are treated as categorical.

While commonly used methods such as multivariate analysis of variance (MANOVA) accounts for the first structure, it cannot handle categorical variables such as gender, race/ethnicity, or experiential indicator dummy variables (Stevens, 2012). Conversely, dichotomous logistic regression only allows for two categorical outcomes, but *does* allow for both kinds of predictor variables (Pedhazer, 1997). Based on these limitations, I chose to analyze the data using the more appropriate multinomial (polychotomous) logistic regression, with a maximum likelihood (ML) estimation, which does allow for both of these data structures (Hosmer & Lemeshow, & Sturdivant, 2013). Multinomial rather than ordinal logistic regression was chosen as it is assumed that there is no order between the three outcome groups. Analyses were performed in Stata 13.1 using the *mlogit* commands and post-estimation commands. (See *Appendix A*.)

By modelling the data using multinomial logistic regression, the goal is to create a set of associated probabilities between the predictor variables and specialty preference at matriculation. In order to create and assess these probabilities, I developed the following *confirmatory* multinomial logistic model which included the person-level demographics as well as their matriculating career and lifestyle factors as predictors:

$$\begin{split} f(x) &= \log \left(\frac{P(Y_1 = Primary\ Care[0]|\ Non - Primary\ Care\ [2])}{P(Y_1 = Undecided\ [1])} \right) \\ &= \beta_0 + \beta_1\ Gender\ + \beta_2\ Asian + \beta_3\ URM + \beta_3\ Other \\ &+ \beta_5\ Physician\ Parent\ + \beta_6\ Professional\ Parent \\ &+ \beta_7\ Hospital\ Work\ + \ \beta_8\ College\ Research \\ &+ \ \beta_9\ Motivational\ Factor\ 1 + \beta_{10}\ Motivational\ Factor\ 2\ \dots \end{split}$$

Where:

- ✤ Gender is a dummy variable coded Female=1, Male=0;
- Race/Ethnicity is composed of four dummy variables where the Asian, URM, and Other groups are compared to the referent group (White), which is omitted from the model;
- Parental Occupation is composed of three dummy variables where Physician-Parent and Professional-Parent are compared to the referent group (Other-Parent);
- ✤ Hospital Work was dummy coded as Yes=1, No=0;
- ✤ College Research was dummy coded as Yes=1, No=0; and,
- ★ Each of the continuously scaled factors from *Research Question #1* are included.

I will also draw the reader's attention to how I have structured the outcome (*Y*). Using this method necessitates choosing the referent outcome group to which the others will be compared. As stated earlier, there is no hypothesized order amongst the Primary Care, Non-Primary Care, and Undecided groups. However, as many prior analyses have investigated the differences between Primary Care and Non-Primary Care groups, looking at how undecided students compare to them across recognized predictors could shed a new light on this topic. Therefore, the Primary Care and Non-Primary Care groups are compared with the Undecided group, the referent group, in this model.
Predictor level assessment was done by examining the strength of the associated probabilities were assessed using each predictor's relative risk ratio (RRR), which is simply the exponentiation of the coefficient's slope. (That is, e^{β} .) I also investigated the transformed standard error and the 95% confidence interval for each predictor before interpreting the coefficient. Thus, the null hypothesis for each predictor is that the coefficient is non-zero in the population (e.g. $\beta_x=0$) while the alternative hypothesis is that there is a difference (e.g. $\beta_x\neq 0$). Stated differently, if the 95% confidence contains the value [0], then we cannot reject the null hypothesis, and assume that there the predictor is not associated with a difference in the population.

Model fit was assessed using the same assessment methods for dichotomous logistic regression. That is, a -2 *Log Likelihood* (-2*LL*) χ^2 test was executed to ensure that a full model fit the data better than a null model. The test was performed by taking: [-2*LL (null model) -2*LL (full model)] and evaluating that difference on a χ^2 distribution where the degrees of freedom is equal to two times the number of additional predictors (Hosmer & Lemeshow, 2000). (It is multiplied by two since there are two equations in each model.) We interpreted statistically significant values as an indication that the included variables improved the model's fit, and is therefore predictive of specialty preference. This method of fit is preferred over McFadden's pseudo-R², since the relationships are non-linear and thus this statistic is less interpretable as compared to ordinary least squares (OLS) regression (Willett, 2011). Thus, looking at both the

coefficient *and* the change in model fit provides a multifaceted approach for assessing a predictor's value to the model.

Based on the review of the literature, some hypotheses about the predictors were made. It is assumed that the female gender will be positively associated with the outcomes of interest; that is, the relative risk ratio for β_1 (Gender) will be greater than one. For the student's race/ethnicity, it is assumed that when compared with White students, students who self-identify as a URM and those who did not provide their race/ethnicity will have a lower fitted probability of preferring a primary care specialty at matriculation. For Asian students, the literature provides no clear trend, and the population proportions in *Table 3.1* similarly provide no consistent trend across the three specialties. Therefore, we did not create a hypothesis for this variable. Based on the review of the literature and the theoretical framework, it is assumed that when compared with parents without a medical or professional degree, the predictor Physician-Parent will be negatively associated with the outcome of preferring a primary care specialty. The same hypothesis was made for Professional-Parents. I hypothesized that hospital work would be positively associated with *both* primary care and non-primary care preferences, as students may have had additional opportunities to make career plans. Additionally, it was hypothesized that college research would negatively predict a preference for primary care. Finally, hypotheses about the factors will be made once they are validated.

Research Question #3: Specialty Preferences at Graduation

101

As noted above, there are two outcomes of interest in this dissertation: specialty preference at matriculation and at graduation. While the former was addressed in the previous research question, this research question allows for a similar investigation at a second time-point. Indeed, this research question mirrors the previous one both in scope and methods; however, some subtle distinctions should be made.

In order to create and assess these probabilities between the psychosocial predictors and specialty preference at graduation, I reapplied the same confirmatory multinomial logistic model, however, I also chose to include two dummy variables related to prior specialty preferences. Thus, students' matriculating preferences are now included in the model as a *predictor* rather than an outcome. This model is specified as:

(3.2)

$$\begin{split} f(x) &= \log \left\langle \frac{P(Y_1 = Primary\ Care[0]|\ Non - Primary\ Care[2])}{P(Y_1 = Undecided\ [1])} \right\rangle = \beta_0 + \beta_1\ Gender\ + \beta_2\ Asian\ + \\ \beta_3\ URM\ + \beta_3\ Other\ + \ \beta_5\ Physician_{Parent}\ + \ \beta_6\ Professional_{Parent}\ + \\ + \beta_7\ Hospital\ Work\ + \ \beta_8\ College\ Research\ + \ \beta_9\ Motivational\ Factor\ 1\ + \\ \beta_{10}\ Motivational\ Factor\ 2\ ...\ + \ \beta_{11}\ Matriculating:\ Primary\ Care\ + \\ \beta_{12}\ Matriculating:\ Non\ - \ Primary\ Care\) \end{split}$$

Where:

- ✤ Gender is coded Female=1, Male=0;
- Race/Ethnicity is composed of four dummy variables where Asian, URM are compared to the referent group (White) which is omitted from the model;
- Parental Occupation is composed of three dummy variables where Physician-

Parent and Professional-Parent are compared to the referent group (Other-Parent);

Each of the continuously scaled factors are included;

- ✤ Hospital Work was dummy coded as Yes=1, No=0;
- ✤ College Research was dummy coded as Yes=1, No=0;
- Each of the continuously scaled factors from *Research Question #1* are included; and,
- Matriculating Specialty Preference is composed of three dummy variables: prior interest in a primary care specialty, prior interest in a non-primary care specialty, and undecided prior interest, which served as the referent group and omitted from the model.

As with the procedure specified in *Research Question #2*, the predictor level assessment was done by examining the strength of the associated probabilities were assessed using each predictor's relative risk ratio (RRR) and their 95% confidence interval. Model fit was assessed using the -2 *Log Likelihood* (-2*LL*) χ^2 test.

The same hypotheses from *Research Question #2* are adhered to in these model; however, the additional dummy variables representing prior specialty preference also require interpretation. Because the undecided specialty dummy represents the referent group, we assume that those who had a matriculating preference for a non-primary care specialty will have a lower associated probability than those with an undecided preference. However, we assume that those who had a primary care matriculating preference will have a greater associated probability than those with an undecided preference, as their preferences and proclivities towards specialty selection is more fixed.



Figure 3.1. Data-screening process to identify and delimit the study sample. (Note that the arrows are proportionally scaled.)

 Are you planning to become certified in a specialty/subs Yes - Go to question 15a 	specialty?
 No - Go to question 16 Undecided - Go to question 16 	
15a. Choice of specialty/subspecialty:	Specialty/Subspecialty

Figure 3.2. Survey Question #15 and #15a from the 2005 MSQ asking students' plans about specialization.

Are you planning to become certified in a specialty?

O Yes O No

O Undecided whether to seek certification in a specialty

Specialty and Career Plans (continued) Choice of specialty: O Allergy and Immunology O Anesthesiology O Colon and Rectal Surgery O Dermatology O Emergency Medicine O Family Medicine O General Surgery O Internal Medicine O Internal Medicine/Pediatrics O Medical Genetics O Neurological Surgery O Neurology O Obstetrics and Gynecology O Ophthalmology O Orthopedics O Otolaryngology O Pathology O Pediatrics O Physical Medicine and Rehabilitation O Plastic Surgery O Preventive Medicine O Psychiatry O Radiology O Radiation Oncology O Thoracic Surgery O Other

Figure 3.3. Questions from the 2014 GQ regarding students' specialization plans.

	Racial/ Ethnic Categories			
	White	Asian	URM	Other/Unknown
White	21,320 (100.0)	0	0	0
Black	0	0	1,770 (43.7)	0
Hispanic	0	0	2,125 (52.4)	0
Asian	0	5, 120 (100.0)	0	0
Pacific Islander	0	0	50 (1.2)	0
Native American/Alaskan	0	0	107 (2.6)	0
Other/multiple races	0	0	0	816 (87.5)
No response	0	0	0	117 (12.5)
Total No.	21,320	5, 120	4,052	933

Cross-Tabulation of Race/Ethnicities and Racial/Ethnic Categories

Note. Percentages represent column proportions

Item Name	Item Prompt
Intellectual	Being a physician is one of the most intellectually challenging professions.
Research	Profession provides opportunity for research.
Expertise	Profession provides opportunity to develop expertise in a specialized area.
Technology	Physicians employ advanced diagnostic and treatment technologies.
Evaluation	Physicians use critical thinking to evaluate medical findings.
Educate	Physicians can educate patients about health promotion and disease prevention.
Social	Profession provides opportunity to exercise social responsibility.
Difference	Profession provides opportunity to make a difference in people's lives.
Contact	Physicians can have continuing contact with their patients.
Interact	Medicine provides opportunity to interact with other health professionals.
Autonomy	Physicians can choose career directions that allow significant autonomy.
Innovate	Profession provides opportunities for innovation.
Decision	Medicine affords opportunities for decision-making under pressure.
Dexterity	Profession provides opportunity to use manual dexterity skills.
Leadership	Doctors enjoy opportunities for leadership.
Status	Doctors enjoy high status and prestige.
Income	Profession offers possibility of high income.
Security	Medical practice affords high job security.
Authority	Profession provides opportunity for authority.
Control	Physicians can have significant control of their work hours.
Stress	Doctors can find practices that limit the amount of work stress.
Patient Care	Dealing with the behavioral/psychological aspects of patient care is rewarding.

Battery of Reasons for Pursuing a Medical Career from the MSQ

Participants were responding to the prompt: "*Indicate how important the following factors were in your choice of medicine as a career goal*..." and was scaled 0 ("Not important at all") to 4 ("Very important").

		Father's Occupation		
		Other-Parent Professional-Parent Physician-Parent		Physician-Parent
Mother's Occupation	Other-Parent	13,828 (44.0)	2,776 (8.8)	2,653 (8.4)
	Professional- Parent	6,490 (20.7)	3,012 (9.6)	1,444 (4.6)
	Physician- Parent	367 (1.2)	134 (0.4)	721 (2.3)

Cross-Tabulation of Parental Occupation

Note. Percentages represent cell proportions.

		College Research Experience	
		No	Yes
Hospital Work	No	4,941 (15.7)	4,204 (13.9)
	Yes	9,995 (31.8)	12,285 (39.1)

Cross-Tabulation of Professional Socialization Experiences

Note. Percentages represent cell proportions.

CHAPTER 4: RESULTS

Chapter Overview

The focus of this dissertation is to understand how experiences, motivations, and personal characteristics are associated with specialty preferences in multivariate models. However, given that the TREMUR dataset is highly representative of the national population of medical students, I first present the descriptive statistics first disaggregated by specialty preference at matriculation, and then by specialty preference at graduation. I then provide the analytic results for each of the research questions.

Descriptive Statistics

Matriculation Specialty Preference

While the inspection of independent associations (that is, population proportions) between predictors and outcomes should not be interpreted inferentially, looking at group differences provides an indication of the results of the more complex analytic models. Additionally, owing to the large sample size and representativeness, these results may facilitate policy discussions in and of themselves. Keeping this in mind, I present in *Table 4.1* the descriptive statistics according to the respondent's specialty preference at matriculation. Looking first at population proportions, we see that of the final sample,

7,103 (22.6%) indicated a preference for a primary specialty versus 12,455 (39.6%) who were undecided and 11,867 (37.8%) who indicated a non-primary care specialty.

Looking first at the respondent's personal characteristics, we note a significant difference between the three main groups. For example, amongst all respondents who indicated a preference for a primary care specialty, only 37.3% were male as compared the undecided and non-primary care groups which were 46.8% and 56.6% male, respectively. Although there is considerable racial/ethnic variation amongst the three primary care specialties, the proportions of students who are White, Asian, and Other/Unknown are relatively similar for the primary care combined, undecided, and non-primary care groups. Only for URM respondents, who represented 12.9% of the sample, is there a group difference, as a greater proportion of URM students indicated a preference for primary care specialties and for non-primary care specialties than for having an undecided specialty choice.

Three categorical variables were used to account for potential socialization experiences: parental occupation, hospital work, and college research. In the combined sample, we see that 39.1% have a parent with a professional occupation and 16.9% have a parent who is a physician, while 44.0% do not have a parent with one of these occupations. Group differences emerge, and show that those with a preference for primary care are less likely to have a physician-parent, but more likely to have a parent without a professional degree. In *Table 4.1*, I also demonstrate that a greater percentage of those with hospital work experience preferred a primary care specialty than undecided and non-primary care, while the opposite trend occurred for those with a college research experience.

Two salient trends emerge from these descriptive statistics. First, that roughly 40% of medical students were undecided about their specialty preferences at matriculation is notable, and suggests that many respondents have not made firm plans by this point. Second, we note that the sample statistics for those with a preference for primary care differ from the other two groups, most notably in regard to gender, parental occupation, and hospital work, while slight differences were also found for college research and the respondent's race/ethnicity. This suggests that these personal characteristics and professional socialization experiences are likely to be important predictors in the analytic models.

Graduation Specialty Preference

The same set of descriptive statistics is presented in *Table 4.2* according to the respondent's *graduation* specialty preference. Here we note that about the same percentage of respondents indicated a preference for primary care (23.6%) as at matriculation (22.6%), however, each specialty shows a unique trend. For example, there was little difference in those with a preference for family practice while there was a large net migration *away* from pediatrics but a large gain for general internal medicine. We also see that a much smaller proportion of respondents indicated they were undecided at graduation, while the percentage of those with a preference for a non-primary care career greatly increased from matriculation to graduation.

Looking at the personal characteristics of the three groups, we see that a greater percentage of females have a preference for primary care (59.5%) or have an undecided specialty preference (63.4%), while only 45.6% of those with a preference for a nonprimary care specialty were female. The racial/ethnic differences between groups, although small, should be noted. For example, a greater percentage of White respondents preferred non-primary care specialties. Asian respondents had a greater preference for primary care, but we note that most Asian students with this preference indicated that they were interested in general internal medicine. URM and Other/Unknown respondents had relatively stable proportions across the three groups.

In terms of parental occupation, the group differences appear to be smaller at graduation than at matriculation. Of particular interest, a greater percentage of those with physician-parents indicated a preference for non-primary care than for primary care or undecided specialty preferences. A greater percentage of respondents with hospital work experience showed a preference for primary care specialties, while college research experience was more prevalent amongst those with a preference for primary care or nonprimary care specialties, potentially indicating that college research facilitates career preferences.

Finally, accounting for prior specialty preferences is an important part of this dissertation's theoretical framework, and we note that about a third of those with a matriculating preference for primary care maintained that interest at graduation. Those with an undecided preference at matriculation had strong representation in each of the three groups, but made up 47.2% of those with an undecided preference at graduation.

114

Lastly, most of those with a preference for non-primary care specialties at matriculation maintained that preference, perhaps suggesting a more robust preference.

RQ# 1: Factor Analysis of Motivational Factors

Univariate normality was inspected for all of the 22 items using the full dataset of 31,425 respondents, and the descriptive statistics are found in *Table 4.3*, where I present the mean, standard deviation, skewness, and kurtosis statistics for each item. Although there are many approaches and "rules-of-thumb" for how to interpret skewness and kurtosis in particular, the small scale for these items necessitated a relatively liberal cut-offs. Therefore, I used the range -2.0 to +2.0 for skewness and for kurtosis which is also recommended (George & Mallory, 2010). Using these ranges, I noted some violations and immediately discarded the item 'Difference' ("*Profession provides opportunity to make a difference in people's lives*") as its extremely large kurtosis value of 28.00 suggested that its inclusion would pose a serious statistical violation. I visually assessed other problematic items such as 'Evaluation,' 'Educate,' and 'Intellectual' for their non-normal attributes, however, their violations were relatively minor and so these items were included. Thus, it appeared that on the whole, the battery of 21 acceptable items was suitable for factorization.

At this point I used a random-number generator in Stata 13.1 (*Appendix A*) to divide the total sample of 31,245 into two groups for the cross-validation procedure (Reis & Judd, 2000). The first, 'Group 1- EFA,' comprised the first 15,712 respondents while the second, 'Group 2- CFA,' comprised the remaining 15,713 respondents. I inspected all

of the sample statistics to account for any bias between the two groups, however, the two groups were found to be extremely balanced in terms of all covariates listed in *Chapter 3*.

Exploratory Factor Analysis Results

I began by performing a series of preliminary analyses on the data from the 15,712 respondents in 'Group 1- EFA' in order to assess the potential for a favorable factor structure and to assess multivariate normality. As a first step, I inspected the pairwise and inter-item correlation tables to see if there were consistent response patterns. Using the literature as a guide (primarily, Jeffe, Whelan, & Andriole, 2010), I found that many bivariate relationships in this sample were strong and consistent with previous research, that is, many pair-wise correlations were above .50, a favorable indication for successful item clustering. I further inspected many of these strong bivariate relationships using a scatterplot as a visual assessment. Based on these promising bivariate trends, I also inspected the Kiaser-Meyer-Olkin (KMO) statistic to ensure that the multivariate relationships were adequate for this factorization procedure. Indeed, the overall KMO statistic was 0.84, well above the recommended minimum of 0.6 (Hutcheson and Sofroniou, 1999). Additionally, the item-level KMO values ranged from 0.75 to 0.91, again, well above the minimum cutoff. Finally, multivariate normality was assessed through Bartlett's Test for sphericity (Bartlett, 1950). For the 21 items, Bartlett's statistic was extremely significant (p < .001), suggesting that patterns within the data are due to underlying factors and not stochasticity.

At this point, I executed a principal components factorization (PCF) analysis for all of the 21 items. The first step was to determine the number of factors. A preliminary analysis revealed five factors with eigenvalues greater than the benchmark of 1.00 (Kaiser, 1960), suggesting that these factors accounted for at least the variance of a single item. The scree plot also indicated that a five-factor solution would be sensible (Catell, 1966), with a notable "elbow" between the fifth and sixth factors. To further sensitize these results, I performed Horn's Parallel Analysis, which is a Monte Carlo simulation of uncorrelated data used to determine the number of factors to retain and found that, indeed, the five factor solution was the most sensible (Horn, 1965; Ledesma & Valero-Mora, 2007). In determining the number of facts to be used, this is considered to be the superlative indicator (Thompson, 2004, p. 34). Collectively, these five factors accounted for 58.2% of the total observed variance. Factors one through five accounted for 25.6%, 12.7%, 9.1%, 5.6%, and 5.2% of the total observed variance, respectively.

Once it was clear that a five factor solution was preferable, I inspected the factor loadings table indicated that a rotation and found that it would be necessary in order to better disperse the significant factor loadings onto the second to fifth factors. For this design, a factor loading of .45 was considered significant, and according to the benchmark, appreciable cross-loadings were noted. Thus, I began by executing orthogonal rotations (varimax, quartimax, equamax, and parsimax) and oblique rotations options (promax, oblimax, and quatimax) to search for a sensible and statistically viable factor structure. Even when changing the rotation method, this iterative process revealed some consistent theme (i.e., similar items loaded significantly on the same factor) and also showed that certain items consistently cross-loaded or did not load on any of the five factors significantly. The most problematic items were 'Autonomy' ("*Physicians can choose career directions that allow significant autonomy*"), 'Interact' ("*Medicine provides opportunity to interact with other health professionals*"), and 'Innovate' ("*Profession provides opportunities for innovation*").

Eliminating these items, a second iteration followed using the same procedure to determine the number of items for the reduced dataset. Again, the Kaiser's test, the scree plot, and Horn's Parallel Analysis all confirmed a five-factor solution for the remaining 18 items. I then executed an unrotated PCF analysis on the reduced number of items, and collectively these five factors accounted for 62.1% of the total observed variance. Factors one through five accounted for 25.9%, 13.7%, 10.3%, 6.5%, and 5.6% of the total observed variance, respectively. Rather than exploring all orthogonal and oblique rotations as in the previous procedure, I opted for an oblique rotation rather than an orthogonal one, as I assumed that the factors were correlated in the population. Additionally, I used the *promax* rotation because it is considered to have the benefit of being conceptually straightforward (Abdi, 2003). Indeed, using the same criteria of .45 as a minimum factor loading and the absence of any appreciable cross-loadings created the most comprehensible simple structure.

From this extraction method, the factor loadings and reliability statistics (Chronbach's alpha) are presented in *Table 4.4*. Together with the co-authors, we reached agreement that this data-driven approach had produced interpretable factors. In terms of

labeling the factors, we decided that Professional Benefits, Intellectual Intrigue, and Doctor-Patient Relationship appear to be consistent with extant factors in the literature, but in factor development it is important to not relay on the literature for factor names, as different items are used to develop factors in each study (Thompson, 2004, p. 97). While we could not find in the literature a similar factor as Lifestyle Management or Command & Control, these names appear to encapsulate the essence of their loaded items.

In sum, applying an oblique rotation to 18 career motivational items produced five factors. Three of these factors enjoy a historical record in the literature, and two represent new scholarship for career decision-making amongst medical students in the especially important areas of lifestyle balance and work-life management. Additionally, it appears that these data-driven factors represent important and categorically unique latent motivations, as evidenced by the conceptual distinction of their loaded items.

Confirmatory Factor Analysis Results

As stated in the previous chapter, the goal of this split sample design was to determine if a *possible* factor structure could fit the data of a different group of individuals. Because the groups were randomly divided, and no systematic bias was found, it should be possible to evaluate whether the 18 items could be formed into five factors as found in the previous section. In sum, in this section I demonstrate that a *data-driven* approach could be confirmed by a *theory-driven* approach.

The data from 'Group 2- CFA,' comprised the other 15,713 respondents in the sample, was extracted from Stata and entered into MPlus (*Appendix A*), where I executed a confirmatory factor analysis using the MLR as previously described. I first started by examining the exact model produced from the EFA analysis; however, it was immediately clear that the fit indices for this model were below even the minimum recommendations. For example, the CFI value of .87 and the TLI value of .85 are both below the recommendation of .90 provided by Hu and Bentler (1999). Additionally, the RMSEA value for the hypothesized model of .67 would fall into the "mediocre" category suggested by MacCallum, Browne, and Sugawara (1996). Thus, there was sufficient evidence that this hypothesized model did not fit the data for 'Group 2- CFA.'

To address this issue and improve model fit, I executed two strategies. First, I inspected all of the factor loadings both for statistical significance (i.e. a *p*-value less than .05) and for weak standardized factor loading coefficients. Although there were some moderate to weak loadings in the .50 range, they were extremely statistically significant at the p<.001 level. Thus, I decided to move to a second strategy for improving model fit, which was to inspect the residual covariance matrix via the modification indices. Using these indices, I found that there were many potential residual correlations that would improve the model fit. In particular, I found that most of the potential correlations with the highest EPC (expected parameter change) were items that loaded onto the same factor. (As previously mentioned, this is important to note for the practical interpretation of the factor scores.) In a step-by-step process, I found the largest EPC values in the

modification indices, manually correlated the indicated residual, and re-inspected the goodness of fit statistics.

In total, I specified six residuals: one in the Command & Control factor, two in the Intellectual Intrigue factor, and three in the Professional Benefits factor. (The final specified correlations are found in *Table 4.6* under the "Crossloading Estimates" section.) Although additional modifications could have been made, I followed the accepted rule that modifications should be few in number, should not drastically change the model parameters, and should be theoretically justifiable. Regarding the last point, I felt that the modifications I made were between items that were closely and conceptually linked, even within the latent factor. Although these *post hoc* modifications do in fact represent violations of the confirmatory framework, the application was carefully specified so that the subsequent analyses could apply beyond just this subset of students (Thompson, 2004; Schrieber et al., 2006).

Thus, the structural model consisted of the five factors from the EFA analysis with all of their loaded items—and six correlated errors between items loading on the same factor. The final changes are represented in *Figure 4.4* which displays the final modified model. For comparative purposes, measures of model fit are shown in *Table 4.5* for both the hypothetical and modified models. Indeed, simply creating modifications would of course improve the model fit according to the χ^2 test given that all of the EPC values were well above 3.84, but improved model fit should be assessed through the more global goodness of fit tests. Indeed, looking at the change from the hypothesized to the modified models, we see that the RMSEA value decreased to .56, suggesting that model fit improved, but was still in the "mediocre" to "good" range. Additionally, we see that the CFI and TLI values improved to .93 and .91 respectively, suggesting that the fit was "acceptable." (Additionally, we should note that there were some moderate correlations between the factors, which is unsurprising given that the EFA extraction method was oblique. For example, in *Table 4.5*, we see that Professional Benefits and Lifestyle, Professional Benefits and Command & Control, as well as Command & Control and Intellectual Intrigue all had moderate to strong correlations.)

Based on the results from this confirmatory model, we can now answer the question: *What are the career and lifestyle motivations reported by matriculating medical students?* As shown, Professional Benefits, Intellectual Intrigue, Doctor-Patient Relationship, Lifestyle Management as well as Command & Control all appear to be important factors for choosing medicine as a career goal amongst for medical students upon matriculation. To briefly explicate each factor:

- The Professional Benefits factor comprises two items related to the physician's social status ('Status' and 'Authority') as well as two items related to their job benefits ('Income' and 'Security'). This could broadly be thought of as an external motivation due to its focus on financial rewards and social recognition.
- Lifestyle Management comprises only two items, 'Stress' and 'Control.' These items deal with how physicians can limit their work stress, while the second and hours. Thus, it holds conceptual similarity with Professional Benefits as it relates

to the professional benefits of the physician, yet it is unique because it focuses on *balance* rather than solely the benefits.

- Command & Control is also related to professional benefits, since it comprises
 'Decision,' 'Dexterity,' and 'Leadership,' which all deal with a hands-on,
 leadership style familiar in many aspects of medicine. This factor is unique in that
 it reflects both internal and external motivations, complicating its interpretation.
- By contrast, Intellectual Intrigue appears to be a wholly internal motivation.
 Indeed, the items 'Expertise,' 'Technology, 'Evaluation,' 'Research,' and 'Intellectual' each demonstrate how using knowledge can help solve medical problems. Collectively, they appear to indicate a fascination of how knowledge and practice intersect.
- Finally, Doctor-Patient Relationship is made of four items related to the social interaction of the patient and the physician, 'Educate,' 'Contact,' 'Social,' and 'Patient Care.' This factor represents a personal satisfaction of seeing patients and how the profession allows for opportunities to improve patients' health.

Factor Scores

The final phase of this research question is to determine whether there were statistical differences between medical students with a matriculating preference for primary care specialty, those without a specialty preference, and those with a preference for a non-primary care specialty. Given the results from the CFA analysis, I used the factor structure from the modified model to create factor scores (using the regression method where the mean is 0 and the standard deviation is 1) for all 31,425 respondents in the sample. The descriptive statistics are presented in *Table 4.8*. Then, to determine whether or not a group difference existed for each of these five factors, I specified a one-way ANOVA model with an *a priori* linear contrast, where the undecided group was compared with the primary care group as well as the non-primary care group. For all factors, the level of significance was set at .01, to account for the five independent inferential tests.

Each of the one-way ANOVA models resulted in an extremely significant *F*-value. Starting with the motivational factors with a history in the literature, I found that there was a statistically significant effect of matriculating specialty preference on Doctor-Patient Relationship [F(2, 31,422)= 292.96, p<.001], Intellectual Intrigue [F(2, 31,422)= 315.95, p<.001], and Professional Benefits [F(2, 31,422)= 280.49, p<.001]. I also found a statistically significant effect of matriculating specialty preference on Lifestyle [F(2, 31,422)=14.70, p<.001], and on Command & Control [F(2, 31,422)= 315.95, p<.001]. Thus, all of the one-way ANOVA tests demonstrated a significant difference between how students perceived medicine as a career goal and their specialty preference at matriculation.

Because I confirmed that there was a statistical difference between groups on all five motivational factors, I then proceeded with the *a priori* contrasts, which are presented in *Table 4.9*. Here, it is noted that all of the contrasts were found to be statistically significant at the p<.01 level, and nearly all were significant at the p<.001 level. We also see in this table that the *F*-values for each motivational factor appear to

foreshadow the multivariate analysis results. For example, the contrast results for Intellectual Intrigue, Professional Benefits, and Command & Control suggest that undecided matriculants have similar perceptions as primary care matriculants, as evidenced by the smaller *F*-value. However, the opposite appears to be true for the Doctor-Patient Relationship factor, where undecided matriculants have similar perceptions as *non*-primary care matriculants. Finally, the results from the contrast for the Lifestyle motivational factor show that undecided matriculants have relatively similar perceptions as both primary care and non-primary care matriculants.

Taken together, this suggests that the motivational factors I have developed may play a significant role in explaining how and why medical students form interests in particular specialties. Thus, in answering the question, "*are there differences between those with a matriculating preference for primary care specialty, those without a specialty preference, and those with a preference for a non-primary care specialty?*," it appears that the five motivational factors will play a role in explaining specialty preferences.

RQ# 2: Matriculating Specialty Preferences

The goal of this research question is to examine the independent relationship between the hypothesized influential variables and the outcome of specialty preference at matriculation. To accomplish this, I specified a confirmatory multinomial logistic regression model, through which a set of fitted probabilities was produced to compare undecided matriculants with primary care matriculants, as well as undecided matriculants with non-primary care matriculants. Again, this outcome structure was chosen because the relationship between primary care and non-primary care preferences is already well established in the literature. For this research question, the full model is defined as:

(4.1)

$$\begin{split} f(x) &= log \left(\frac{P(Y_1 = Primary\ Care[0]|\ Non - Primary\ Care\ [2])}{P(Y_1 = Undecided\ [1])} \right) \\ &= \beta_0 + \beta_1\ Gender\ + \beta_2\ Asian + \beta_3\ URM\ + \beta_3\ Other\ Parent \\ &+ \beta_5\ Physician\ Parent\ + \beta_6\ Professional\ Parent \\ &+ \beta_7\ Hospital\ Work\ + \ \beta_8\ College\ Research \\ &+ \ \beta_9\ Lifestyle\ Management\ + \ \beta_{12}\ Doctor\ - Patient\ Relationship \\ &+ \ \beta_{13}\ Intellectual\ Intrigue \end{split}$$

Although it is common in the medical literature, it is not sensible to begin by interpreting the final model's predictor coefficients. Instead, I specified a procedure of comparing a taxonomy of nested models, which the -2 Log Likelihood (-2LL) χ^2 test was used as a goodness of fit statistic (Hosmer & Lemeshow, 2000). Rather than inserting one predictor into the model at a time, I began by inserting "variable clusters" according to their conceptual characteristics and according to their importance to this study. For that reason, the first cluster included gender and race/ethnicity; the second cluster included parental occupation and the experiential variables hospital work and college research; and the third cluster included the set of five motivational factors. Thus, all models are nested, and can be evaluated with the -2LL statistic.

Looking at the -2LL value for each model in the final rows of *Tables 4.11(A)* and 4.11(B), it is clear that each of the clusters of variables statistically significantly improved

model fit. For example, from the Model #1 to Model #2, the change in the -2LL value was -913.4, far greater than the χ^2 critical value of 30.72 (3.84 [the value of χ^2 with one degree of freedom at p=.05] times four predictor variables [Female, Asian, URM, Other/Unknown] times 2[the number of sub-models]). Based on these values, it is possible to categorically reject the null hypothesis that the predictors gender and race/ethnicity are unrelated to the outcome, specialty preference at matriculation. Stated differently, gender and race/ethnicity *jointly* have a statistically significant relationship with the outcome, and as a result, improved the fit of the model.

I applied this same test to evaluate Model #3, and found that parental occupation and the experiential variables hospital work and college research *jointly* had a statistically significant relationship with the outcome. Similarly, applying this test to Model #4 showed that the five motivational factors *jointly* had a statistically significant relationship with the outcome. I emphasize the word jointly to highlight the fact that this does not necessarily indicate that the individual predictor coefficients will be statistically significant. It is, however, a promising sign that the indicators were well chosen, and will help makes sense of the conceptual model.

Based on this model fitting procedure, I then turned to interpreting the predictor coefficients in Model #4, which can be considered as the final model for this research question. To assess each predictor, I inspected the relative risk ratio (RRR), and its *p*-value where the level of significance is p=.05. Because an RRR value of 1.00 indicates that there is no population difference for a given predictor, a handy rule of thumb for

interpreting the coefficients is that if the predictor's coefficient is greater than one, then it positively predicts a primary care preference in in *Table 4.11(A)* and a non-primary care preference in *Table 4.11(B)*. Conversely, if the predictor's coefficient is less than one, then it positively predicts an undecided preference at matriculation.

Starting with the variables from the first cluster, it is clear that gender plays a role in predicting specialty preferences at matriculation. In terms of a primary care preference, the coefficient for Female was 1.25 (p<.01) suggesting that for every four males with a preference for primary care, five females would have this preference. Nearly the opposite is true for non-primary care, where the coefficient for Female was .81 (p<.01). Thus, females had higher independent fitted odds of indicating a preference for primary care than for an undecided preference, while they had higher fitted odds of being undecided than having a preference for non-primary care.

For race, all predictors are compared with the reference category of White. In the final model, Asian students had lower fitted odds (.89, p<.01) than white students for preferring non-primary care over undecided, but their coefficient was non-significant in the final primary care model (.95, p=.23). The coefficients for URM students was extremely statistically significant in both the final primary care model (1.46, p<.01) and the non-primary care model (1.74, p<.01), suggesting that they had much lower odds of being undecided at matriculation than White students. Indeed, when compared with White students, they were simply more likely to have a specialty preference of any kind. Finally, students of an Other/Unknown race/ethnicity had lower fitted odds than White

students of having a preference for primary care over being undecided (.78, p<.01), but the coefficient in the final non-primary care model was non-statistically significant (.97, p=.71).

In terms of the chosen socialization experiences, some interesting trends around parental occupation appeared. In the final primary care model, respondents of Professional-Parents did not have different odds than respondents of Other-Parents for preferring primary care over being undecided (.95, p=.09); however, respondents of Physician-Parents *did* have lower fitted odds (.70, p<.01) of preferring primary care versus being undecided. In terms of non-primary care, respondents of both Professional-Parents and Physician-Parents had lower fitted odds of preferring non-primary care specialties versus being undecided, .93 (p<.01) and .85 (p<.01), respectively. Taken together, these unexpected findings suggest that respondents of higher parental education were more likely to be undecided at matriculation.

For the predictor Hospital Work, the coefficients in both the final primary care model (1.21, p<.01) and in the final non-primary care model (1.11, p<.01) suggest that those with previous hospital work experience had higher fitted odds of preferring either primary or non-primary care specialties than being undecided. The College Research predictor was non-significant in the final non-primary care model (1.04, p=.43), but in the final primary care model, respondents with college research experience had lower fitted odds (.93, p=.01) of preferring a primary care specialty over being undecided.

Finally, the interpretation of the five motivational factors is necessarily different, as these are continuous variables. Thus, while we do interpret their significance in the same way, the coefficient is interpreted as an increase (or decrease) in fitted odds for each unit increase on the predictor's continuous scale. For example, while Professional Benefits was not significant in the final primary care model, for a unit increase in Professional Benefits in the final non-primary care model, the fitted odds were .88 (p < .01) lower for preferring a non-primary care specialty. In other words, the higher a respondent indicated that Professional Benefits were important to them, the lower the odds were that they would prefer non-primary care over being undecided. In the same way, higher scores for Intellectual Intrigue positively predicted a non-primary care preference (1.35, p < .01), but negatively predicted a primary care preference (.80, p < .01). The same trend was true for Command & Control; however, we found the opposite trend for Doctor-Patient Relationship, where higher scores negatively predicted a non-primary care preference (.43, p < .01), but positively predicted a primary care preference (2.41, p < .01)p < .01). Finally, for Lifestyle Management, higher scores negatively predicted a primary care preference (.93, p < .01), but was non-significant in the final non-primary care model (1.01, p=.86).

RQ# 3: Graduating Specialty Preferences

Nearly identical to the previous research question in conceptualization and methods, the goal of this research question is to examine the independent relationship between the hypothesized influential variables and the outcome of specialty preference at *graduation*. Thus, I will be focusing on the independent relationships between a set of predictors and preferring a primary or non-primary care specialty, compared with the base outcome of being undecided. The model is stated as:

$$\begin{split} f(x) &= log \left(\frac{P(Y_1 = Primary \, Care[0]| \, Non - Primary \, Care[2])}{P(Y_1 = Undecided[1])} \right) \\ &= \beta_0 + \beta_1 \, Gender + \beta_2 \, Asian + \beta_3 \, URM + \beta_3 \, Other \, Parent \\ &+ \beta_5 \, Physician \, Parent + \beta_6 \, Professional \, Parent \\ &+ \beta_7 \, Hospital \, Work + \, \beta_8 \, College \, Research \\ &+ \, \beta_9 \, Lifestyle \, Management + \beta_{10} \, Command \, \& \, Control \\ &+ \, \beta_{11} \, Professional \, Benefits + \, \beta_{12} Doctor - Patient \, Relationship \\ &+ \, \beta_{13} \, Intellectual \, Intrigue + + \beta_{14} \, Matriculating: Primary \, Care \\ &+ \, \beta_{15} \, Matriculating: Non - Primary \, Care \end{split}$$

Most of these predictor variables were also specified in the previous question, and so this model is structured in a very similar way. For example, the first thirteen predictor variables are clustered in the same way for the step-wise nested model-building procedure are the same, where the first cluster contains gender and race/ethnicity; the second cluster contains parental occupation, hospital work, and college research; while the third cluster contains the set of five motivational factors. Additionally, two binary variables were included to account for prior specialty preference, and are added to the model as the fourth and final variable cluster. Again, all models are nested into higher order ones, and can be evaluated with the *-2LL* statistic.

As seen in *Tables 4.11 (A)* and *4.11 (B)*, which show the model indices and predictor coefficients for primary care versus undecided and non-primary care versus undecided, respectively. First, I inspected the change in the *-2LL* value from Model #1

4.2

(null) to Model #2, Model #2 to Model #3, and so on. Each change in the -2LL value well exceeded the respective χ^2 critical value; from this, I concluded that each variable cluster was jointly associated with the outcome of specialty preference at graduation. In turn, I considered Model #5 to be the final model.

Assessing the variable coefficients for this final model, I found that being female was associated with statistically significant lower odds (.84, p<.01) of preferring primary care at graduation, and even lower odds (.60, p<.01) of preferring non-primary care specialties at graduation. This suggests that females were much more likely than males to be undecided at a critical juncture of their career. Looking at student's race/ethnicity, I found the same trend to be true for Asian students (when compared with White students), as their coefficient (.87, p<.01) negatively predicted a primary care preference *and* a nonprimary care preference (.76, p<.01). In the final model, there was no statistically significant coefficient for URM students, suggesting no population difference for which kind of specialty they preferred at graduation. Finally, for students of Other/Unknown race/ethnicity had lower fitted odds of preferring primary care specialties (.80, p<.01), but the coefficient in the non-primary care model was not significant (.90, p=.23).

The parent-occupation socialization variable demonstrated a different relationship with the outcome in each particular model. For example, in the primary care model, respondents of Physician-Parents were just as likely as respondents of Other-Parents to prefer primary care, but respondents of Professional-Parents had slightly higher fitted odds (1.10, p<.01) of preferring primary care specialties at graduation. The opposite was true for the non-primary care model, where respondents of Professional-Parents were just as likely as respondents of Other-Parents to prefer non-primary care, but I found that respondents of Physician-Parents had slightly higher fitted odds (1.10, p<.01) of preferring primary care specialties at graduation. Thus, in the final model, parental occupation played a mixed role. That said, the Hospital Work and College Research variables positively predicted a preference for non-primary care and primary care in their respective model, suggesting that these pre-medical school experiences shape career preferences.

On the whole, the five motivational factors were not statistically associated with the outcome of preferring primary care over being undecided. (That is, the coefficients were non-statistically significant.) The one exception was Intellectual Intrigue (1.11, p<.01), where for those who had a higher score for this factor had higher fitted odds of preference primary care. While this finding may *seem* inconsequential, it should not be overlooked as it essentially shows that students with no specialty preference and those with a primary care specialty preference expressed similar motivations at graduation. For the non-primary care model, the Professional Benefits and Lifestyle Management variables did not have explanatory value, however Intellectual Intrigue (1.50, p<.01) and Command & Control (1.22, p<.01) positively predicted a preference for non-primary care, as respondents with higher scores on these factors were more likely to choose nonprimary care preferences over being undecided. In contrast, the Doctor-Patient Relationship variable (.58, p<.01) significantly negatively predicted a preference for nonprimary care over being undecided. In the nexus of these findings, it is clear that in terms of key motivations, respondents who were undecided at matriculation were more similar to those who preferred primary care specialties than to those who preferred non-primary care specialties.

Finally, the primary contribution of this research was to show if and how specialty preferences changed from matriculation into medical school to graduation from it. Thus, the final variable cluster was prior specialty preferences coded as undecided, primary care, and non-primary care. Consistent with the previous research question, undecided respondents formed the referent category. For the primary care model, I found that those having a prior primary care specialty preference had about 1.5 times the fitted odds of previously undecided respondents for choosing primary care at matriculation. Similarly, those with a prior *non*-primary care preference had about 1.2 the fitted odds of previously undecided respondents for choosing primary care at matriculation. In combination, these data show that those having a prior specialty preference *of any kind* were more likely than previously undecided respondents to prefer primary care at graduation. This may suggest that non-primary care and primary care are less different poles on a continuum of preferences, and might be thought of as different *kinds* of choices on a continuum showing the *strength* of preferences.

Finally, for the non-primary care model, I show that those having a prior primary care preference had lower fitted odds (.79, p<.01) than previously undecided respondents to prefer a non-primary care specialty at graduation. In other words, for every five undecided respondents who switched their preference to non-primary care at graduation,

only four respondents with a previous primary care preference did so. Unsurprisingly, those with a previous preference for non-primary care had significantly higher fitted odds 2.05, p<.01) than previously undecided respondents of preferring non-primary care again at graduation.
By Matriculating Specialty Preference, Student-Level Characteristics

	Total		Primary C	Care Specialties		Other Specia	Other Specialty Choice	
	Total	Family Medicine	Pediatrics	General Internal Medicine	Primary Care Combined	Undecided	Non-Primary Care	
No. Respondents	31,425 (100.0)	1,612 (100.0)	3,183 (100.0)	2,308 (100.0)	7,103 (100.0)	12,455 (100.0)	11,867 (100.0)	
Gender								
Male	15,205 (48.4)	670 (41.6)	817 (25.7)	1,163 (50.4)	2,650 (37.3)	5,831 (46.8)	6,724 (56.6)	
Female	16,220 (51.6)	942 (58.4)	2,366 (74.3)	1,145 (49.6)	4,453 (62.7)	6,624 (53.2)	5,143 (43.3)	
Race/Ethnicity								
White	21,320 (67.8)	1,194 (74.1)	2,108 (66.2)	1,402 (60.8)	4,704 (66.2)	8,632 (69.3)	7,984 (67.3)	
Asian	5,120 (16.3)	159 (9.9)	497 (15.6)	491 (21.3)	1,147 (16.2)	2,204 (17.7)	1,769 (14.9)	
URM	4,052 (12.9)	222 (13.8)	495 (15.6)	354 (15.3)	1,071 (15.1)	1,211 (9.7)	1,770 (14.9)	
Other/ Unknown	933 (3.0)	37 (2.3)	83 (2.6)	61 (2.6)	181 (2.6)	408 (3.3)	344 (2.9)	
Parental Occupation								
Physician-Parent	5,319 (16.9)	176 (10.9)	383 (12.0)	414 (17.9)	973 (13.7)	2,341 (18.8)	2,005 (16.9)	
Professional-Parent	12,278 (39.1)	679 (42.1)	1,265 (39.7)	907 (39.3)	2,851 (40.1)	4,877 (39.2)	4,550 (39.3)	
Other-Parent	13,838 (44.0)	757 (47.0)	1,535 (48.2)	987 (42.8)	3,279 (46.2)	52,37 (42.1)	5,312 (44.8)	
Hospital Work								
Yes	22,280 (70.1)	1,168 (72.5)	2,377 (74.7)	2,377 (74.7)	5,236 (73.7)	8,628 (69.3)	8,426 (70.9)	
No	9,145 (29.1)	(444 (27.5)	806 (25.3)	806 (25.3)	1,867 (26.3)	3,827 (30.7)	3,451 (29.1)	
College Research								
Yes	16,489 (52.5)	659 (40.9)	1,658 (52.1)	1,258 (54.5)	3,575 (50.3)	6,512 (52.3)	6,402 (54.0)	
No	14,936 (47.5)	953 (59.1)	1,525 (47.9)	1,050 (45.5)	3,528 (49.7)	5,943 (47.7)	6,465 (46.1)	

Note. Percentages represent column proportions

By Graduating Specialty Preference, Student-Level Characteristics

	Total		Primary C	Care Specialties		Other Speci	alty Choice
	Totai	Family Medicine	Pediatrics	General Internal Medicine	Primary Care Combined	Undecided	Non-Primary Care
No. Respondents	31,425 (100.0)	1,675 (100.0)	2,427 (100.0)	3,307 (100.0)	7,400 (100.0)	4,821 (100.0)	19,204 (100.0)
Gender							
Male	15,205 (48.4)	670 (40.0)	581 (24.0)	1,744 (52.8)	2,995 (40.5)	1,767 (36.7)	10,443 (54.4)
Female	16,220 (51.6)	1,004 (60.0)	1,844 (76.0)	1,557 (47.2)	4,405 (59.5)	3,054 (63.4)	8,761 (45.6)
Race/Ethnicity							
White	21,320 (67.8)	1,244 (74.3)	1,711 (70.6)	1,969 (59.7)	4,924 (66.5)	3,126 (64.8)	13,270 (69.1)
Asian	5,120 (16.3)	123 (7.4)	325 (13.4)	824 (25.0)	1,272 (17.2)	910 (18.9)	2,938 (15.3)
URM	4,052 (12.9)	240 (14.3)	325 (13.4)	443 (13.4)	1,008 (13.6)	625 (13.0)	2,419 (12.6)
Other/ Unknown	933 (3.0)	67 (4.0)	64 (2.6)	65 (2.0)	196 (2.7)	160 (3.3)	577 (3.0)
Parental Occupation							
Physician-Parent	5,319 (16.9)	161 (9.6)	319 (13.2)	687 (20.8)	1,167 (15.7)	736 (15.8)	3,389 (17.7)
Professional-Parent	12,278 (39.1)	717 (42.8)	1,022 (42.1)	1,247 (27.8)	2,986 (40.4)	1,842 (38.2)	7,450 (38.8)
Other-Parent	13,838 (44.0)	796 (47.6)	1,084 (44.7)	1,367 (41.4)	3,247 (43.9)	2,216 (46.0)	8,365 (43.6)
Hospital Work							
Yes	22,280 (70.1)	1,239 (74.0)	1,772 (73.1)	2,368 (71.7)	5,379 (72.7)	3,337 (69.2)	13,564 (70.6)
No	9,145 (29.1)	435 (26.0)	653 (26.9)	933 (28.3)	2,021 (27.3)	1,484 (30.8)	5,640 (29.4)
College Research							
Yes	16,489 (52.5)	729 (43.5)	1,237 (51.0)	1,963 (59.5)	3,929 (53.1)	2,359 (48.9)	10,201 (53.1)
No	14,936 (47.5)	945 (56.5)	1,188 (49.0)	1,338 (40.5)	3,471 (46.9)	2,462 (51.1)	9,003 (46.9)
Prior Specialty Pref.							
Primary Care	7,103 (22.6)	638 (38.1)	1,132 (46.7)	873 (26.5)	2,643 (35.7)	1,365 (28.3)	3,095 (16.1)
Undecided	12,455 (39.6)	641 (38.3)	854 (35.2)	1,434 (43.4)	2,929 (39.6)	2,286 (47.2)	7,240 (37.7)
Non-Primary Care	11,867 (37.8)	395 (23.6)	439 (18.1)	994 (30.1)	1,828 (24.7)	1,170 (24.3)	8,869 (46.2)

Note. Percentages represent column proportions.

Item Name	Mean	S.D.	Skewness	Kurtosis
Intellectual	3.23	.92	-1.22	4.16
Research	1.73	1.22	.20	2.07
Expertise	2.80	1.04	71	2.98
Technology	2.60	1.11	57	2.67
Evaluation	3.33	.82	-1.27	4.63
Educate	3.53	.75	-1.76	6.16
Social	2.25	.87	-1.40	4.73
Difference	3.87	.44	-4.47	28.00
Contact	3.23	.92	-1.20	4.02
Interact	2.70	1.04	62	2.87
Autonomy	3.01	1.04	96	3.32
Innovate	2.98	.96	-1.22	4.12
Decision	2.58	1.18	0.58	2.53
Dexterity	2.23	1.30	27	2.00
Leadership	2.81	1.06	74	2.97
Status	1.86	1.25	.07	1.97
Income	2.05	1.14	21	2.19
Security	2.85	1.08	80	2.96
Authority	1.71	1.25	.12	1.97
Control	2.19	1.25	20	2.04
Stress	1.91	1.26	.02	1.97
Patient Care	3.01	1.09	97	3.13

Descriptive Statistics of 22 Motivational Items

Note. Participants were responding to the prompt: "*Indicate how important the following factors were in your choice of medicine as a career goal...*" and was scaled 0 ("Not important at all") to 4 ("Very important").

	Motivational Factor						
Item Name	Professional	Lifestyle	Command &	Intellectual	Doctor-Patient		
	Benefits	Management	Control	Intrigue	Relationship		
	(α=.84)	(α=.84)	(α=.71)	(α=.74)	(α=.67)		
Status	.89	.06	.14	.10	.02		
Income	.83	.20	.03	.12	07		
Authority	.74	.19	.23	.10	.06		
Security	.60	.32	.05	.16	.04		
Stress	.27	.87	.02	.04	.14		
Control	.32	.83	.07	.04	.10		
Decision	.13	01	.82	.29	.09		
Dexterity	.12	.07	.86	.28	.04		
Leadership	.12	.03	.57	.23	.12		
Expertise	.18	.08	.13	.72	.04		
Technology	.14	.08	.21	.65	.08		
Evaluation	.05	.01	.21	.61	.16		
Research	02	03	.03	.76	.02		
Intellectual	.16	.00	.11	.61	.12		
Educate	.21	.12	.02	.06	.75		
Contact	07	.03	01	.03	.70		
Social	.04	03	.09	.12	.74		
Patient Care	.04	.13	.02	.04	.61		
Note. All darkly-	shaded cells repr	esent the items as	sociated with the	factor in the colu	mn header.		

Exploratory Factor Analysis of 18 Motivational Items (N=15,712)

	X^2	df	X^2/df	RMSEA	CFI	TLI
Model 1 (Hypothesized)	16,868	125	<.001	0.67	0.87	0.85
Model 2 (Modified)	11,852	119	<.001	0.56	0.93	0.91

Confirmatory Factor Analysis Goodness-of-fit Statistics (N=15,713)

	Unstandardized (SE)	Standardized (SE)
Factor Correlations		
Professional Benefits with		
Lifestyle Management	0.43 (.01)	0.57 (.01)
Intellectual Intrigue	0.19 (.01)	0.39 (.01)
Command & Control	0.31 (.01)	0.62 (.01)
Doctor-Patient Relationship	0.02 (.01)	0.06 (.01)
Lifestyle Management with		
Intellectual Intrigue	0.14 (.01)	0.20 (.01)
Command & Control	0.20 (.01)	0.28 (.01)
Doctor-Patient Relationship	0.08 (.01)	0.15 (.01)
Command & Control with		
Intellectual Intrigue	0.31 (.01)	0.69 (.01)
Doctor-Patient Relationship	0.11 (.01)	0.33 (.01)
Intellectual Intrigue with		
Doctor-Patient Relationship	0.12 (.01)	0.36 (.01)
Factor Variance		
Professional Benefits	0.52 (.01)	1.00 (.00)
Lifestyle	1.07 (.01)	1.00 (.00)
Command & Control	0.47 (.01)	1.00 (.00)
Intellectual	0.45 (.01)	1.00 (.00)
Altruism	0.26 (.01)	1.00 (.00)

Confirmatory Factor Analysis of Five Motivational Factors, Structural Model (N=15,713)

	Unstandardized (SE)	Standardized (SE)
Factor Loadings		
Professional Benefits by		
Security	1.00 (.00)	0.67 (.01)
Status	1.27 (.02)	0.73 (.00)
Income	1.04 (.01)	0.65 (.01)
Authority	1.43 (.02)	0.83 (.00)
Lifestyle Management by		
Stress	1.00 (.00)	0.82 (.00)
Control	1.08 (.01)	0.89 (.00)
Command & Control by		
Dexterity	1.00 (.00)	0.53 (.01)
Leadership	1.14 (.02)	0.74 (.01)
Decision	1.04 (.01)	0.60 (.01)
Intellectual Intrigue by		
Expertise	1.00 (.00)	0.65 (.01)
Technology	1.06 (.01)	0.64 (.01)
Evaluation	0.78 (.01)	0.63 (.01)
Research	0.79 (.01)	0.43 (.01)
Intellectual	0.74 (.01)	0.54 (.01)
Doctor-Patient Relationship by		
Contact	1.00 (.00)	0.56 (.01)
Social	1.07 (.02)	0.62 (.01)
Patient Care	0.97 (.02)	0.45 (.01)
Educate	1.04 (.02)	0.70 (.01)
Intercept Means		
Authority	1.72 (.01)	1.38 (.01)
Contact	3.23 (.01)	3.50 (.02)
Control	2.18 (.01)	1.75 (.01)
Decision	2.52 (.01)	2.14 (.01)
Dexterity	2.22 (.01)	1.71 (.01)
Evaluation	3.33 (.01)	4.06 (.03)
Expertise	2.78 (.01)	2.68 (.01)
Income	2.05 (.01)	1.79 (.01)
Intellectual	3.23 (.01)	3.49 (.02)
Leadership	2.82 (.01)	2.65 (.01)
Educate	3.54 (.00)	4.70 (.04)
Patient Care	3.02 (.01)	2.77 (.01)
Research	1.74 (.01)	1.41 (.01)
Security	2.85 (.01)	2.65 (.01)
Social	3.35 (.01)	3.84 (.03)
Status	1.87 (.01)	1.48 (.01)
Stress	1.91 (.01)	1.51 (.01)
Technology	2.60 (.01)	2.35 (.01)

Confirmatory Factor Analysis of Five Motivational Factors, Measurement Model (N=15,713).

Crossloading Estimates		
Dexterity With Decision	0.33 (.01)	0.32 (.01)
Income With Security	0.23 (.01)	0.34 (.01)
Income With Status	0.29 (.01)	0.39 (.01)
Authority With Security	0.14 (.01)	0.26 (.01)
Expertise With Technology	0.18 (.01)	0.27 (.01)
Evaluate With Intellectual	0.11 (.01)	0.21 (.01)
Residual Variances		
Authority	0.48 (.01)	0.31 (.01)
Contact	0.60 (.01)	0.67 (.01)
Control	0.33 (.01)	0.21 (.01)
Decision	0.89 (.01)	0.64 (.01)
Dexterity	1.22 (.01)	0.72 (.01)
Evaluation	0.40 (.01)	0.60 (.01)
Expertise	0.63 (.01)	0.58 (.01)
Income	0.74 (.01)	0.57 (.01)
Intellectual	0.61 (.01)	0.71 (.01)
Leadership	0.52 (.01)	0.46 (.01)
Educate	0.29 (.01)	0.50 (.01)
Patient Care	0.94 (.01)	0.79 (.01)
Research	1.22 (.01)	0.81 (.01)
Security	0.63 (.01)	0.55 (.01)
Social	0.47 (.01)	0.61 (.01)
Status	0.74 (.01)	0.47 (.01)
Stress	0.52 (.01)	0.33 (.01)
Technology	0.72 (.01)	0.59 (.01)



Figure 4.1. Modified factor model.

	Primary Care	Undecided	Non-Primary Care
Doctor-Patient Relationship	.10 (.36)	01 (.42)	55 (.46)
Intellectual Intrigue	08 (.59)	05 (.57)	.10 (.56)
Professional Benefits	14 (.85)	05 (.83)	.14 (.85)
Lifestyle Management	04 (.94)	01 (.92)	.03 (.95)
Command & Control	09 (.59)	06 (.59)	.12 (.58)

Factors Score Means and Standard Deviations for All Five "Motivational Factors" by Matriculating Specialty Preference

	Linear Contrast (Undecided	Linear Contrast (Undecided
	vs. Frimary Care)	vs. non-r rimary Care)
Doctor-Patient Relationship	262.76 ***	89.92 ***
Intellectual Intrigue	11.62 ***	447.58 ***
Professional Benefits	49.55 ***	308.02 ***
Lifestyle Management	6.30 **	11.22 ***
Command & Control	12.67 ***	598.51 ***

ANOVA A Priori Contrast Results, Showing the F-Value for All Five "Motivational Factors"

Note. ***p<.001, **p<.01

Table 4.10 (A)

		Relative Risk Rat	io (Standard Erro	<i>r</i>)
	Model #1 (Null)	Model #2	Model #3	Model #4
Gender				
Male ^R		1.00	1.00	1.00
Female		1.45 (.04) ***	1.43 (.04) ***	1.25 (.04) ***
Race/Ethnicity				
White ^R		1.00	1.00	1.00
Asian		0.94 (.04)	0.98 (.04)	0.95 (.04)
URM		1.56 (.07) ***	1.55 (.07) ***	1.46 (.07) ***
Other/ Unknown		0.79 (.07) ***	0.80 (.07) **	0.78 (.07) ***
Parental Occupation				
Other-Parent ^R			1.00	1.00
Professional-Parent			0.94 (.03)	0.95 (.03)
Physician-Parent			0.69 (.03) ***	0.70 (.03) ***
Hospital Work				
No ^R			1.00	1.00
Yes			1.23 (.04) ***	1.21 (.04) ***
College Research				
No ^R			1.00	1.00
Yes			0.91 (.03) ***	0.93 (.03) ***
Motivational Factors				
Professional Benefits				1.04 (.04)
Intellectual Intrigue				0.80 (.02) ***
Command & Control				0.89 (.05) **
Doctor-Patient Relationship				2.41 (.10) ***
Lifestyle Management				0.93 (.02) ***
-2LL	67,292	66,379	66,231	64,210

Multinomial Logistic Regression Results Comparing the Matriculating Specialty Preference Option of Primary Care versus the Base Outcome of Undecided

Table 4.10 (B)

		Relative Risk Rat	tio (Standard Erro	or)
	Model #1 (Null)	Model #2	Model #3	Model #4
Gender				
Male ^R		1.00	1.00	1.00
Female		0.65 (.02) ***	0.65 (.02) ***	0.81 (.02) ***
Race/Ethnicity				
White ^R		1.00	1.00	1.00
Asian		0.88 (.03) ***	0.88 (.03) ***	0.89 (.03) ***
URM		1.66 (.07) ***	1.65 (.07) ***	1.74 (.07) ***
Other/ Unknown		0.94 (.07)	0.94 (.07)	0.97 (.07)
Parental Occupation				
Other-Parent ^R			1.00	1.00
Professional-Parent			0.92 (.03)	0.93 (.03) **
Physician-Parent			0.86 (.03) ***	0.85 (.03) ***
Hospital Work				
No ^R			1.00	1.00
Yes			1.10 (.03) ***	1.11 (.03) ***
College Research				
No ^R			1.00	1.00
Yes			1.08 (.03) ***	1.04 (.03)
Motivational Factors				
Professional Benefits				0.88 (.06) ***
Intellectual Intrigue				1.35 (.08) ***
Command & Control				2.01 (.07) ***
Doctor-Patient Relationship				0.43 (.02) ***
Lifestyle Management				1.01 (.01)
-2LL	67,292	66,379	66,231	64,210

Multinomial Logistic Regression Results Comparing the Matriculating Specialty Preference Option of Non-Primary Care versus the Base Outcome of Undecided

Table 4.11 (A)

	Relative Risk Ratio (Standard Error)						
	Model #1 (Null)	Model #2	Model #3	Model #4	Model #5		
Gender							
Male ^R		1.00	1.00	1.00	1.00		
		0.85 (.03)	0.84 (.03) ***	0.86 (.03) ***	0.84 (.03) ***		
Female		***		~ /			
Race/Ethnicity		1.00	1.00	1.00	1.00		
White K		1.00	1.00	1.00	1.00		
Asian		0.89 (.04) **	0.88 (.04) **	0.87 (.04) **	0.87 (.04) ***		
URM		1.04 (.06)	1.06 (.06)	1.05 (.06)	1.02 (.06)		
Other/ Unknown		0.78 (.09) **	0.78 (.09) **	0.79 (.09) **	0.80 (.09) **		
Parental Occupation							
Other-Parent ^R			1.00	1.00	1.00		
Professional-Parent			1.10 (.04) **	1.10 (.04) **	1.10 (.04) **		
Physician-Parent			1.05 (.06)	1.05 (.06)	1.08 (.06)		
Hospital Work							
No ^R			1.00	1.00	1.00		
Yes			1.17 (.05) ***	1.17 (.05) ***	1.16 (.05) ***		
College Research							
No ^R			1.00	1.00	1.00		
Yes			1.18 (.05) ***	1.17 (.05) ***	1.17 (.05) ***		
Motivational Factors							
Professional Benefits				1.04 (.04)	1.04 (.04)		
Intellectual Intrigue				1.09 (.06)	1.11 (.06) ***		
Command & Control				1.02 (.07)	1.02 (.07)		
Doctor-Patient Relationship				1.05 (.06)	1.05 (.06)		
Lifestyle Management				0.94 (.03) **	0.95 (.03)		
Prior Specialty Pref.							
Undecided ^R					1.00		
Primary Care					1.53 (.07) ***		
Non-Primary Care					1.17 (.06) ***		
-2LL	58,392	57,616	57,532	56,732	55,405		

Multinomial Logistic Regression Results Comparing the Graduating Specialty Preference Option of Primary Care versus the Base Outcome of Undecided

Table 4.11 (B)

	Relative Risk Ratio (Standard Error)						
	Model #1 (Null)	Model #2	Model #3	Model #4	Model #5		
Gender							
Male ^R		1.00	1.00	1.00	1.00		
Female		0.48 (.02) ***	0.48 (.02) ***	0.58 (.02) ***	0.60 (.02) ***		
Race/Ethnicity							
White ^R		1.00	1.00	1.00	1.00		
Asian		0.78 (.03) ***	0.76 (.03) ***	0.76 (.03) ***	0.76 (.03) ***		
URM		0.98 (.05)	0.99 (.05)	1.04 (.05)	0.97 (.05)		
Other/ Unknown		0.91 (.08)	0.90 (.08)	0.90 (.08)	0.90 (.08)		
Parental Occupation							
Other-Parent ^R			1.00	1.00	1.00		
Professional-Parent			1.05 (.04)	1.06 (.04)	1.06 (.04)		
Physician-Parent			1.16 (.05) ***	1.15 (.05) ***	1.15 (.05) ***		
Hospital Work							
No ^R			1.00	1.00	1.00		
Yes			1.09 (.03) ***	1.10 (.03) ***	1.10 (.03) ***		
College Research							
No ^R			1.00	1.00	1.00		
Yes			1.21 (.03) ***	1.14 (.03) ***	1.13 (.03) ***		
Motivational Factors							
Professional Benefits				0.94 (.04)	0.96 (.04)		
Intellectual Intrigue				1.58 (.08) ***	1.50 (.08) ***		
Command & Control				1.37 (.06) ***	1.22 (.06) ***		
Doctor-Patient Relationship				0.49 (.01) ***	0.58 (.01) ***		
Lifestyle Management				1.00 (.02)	1.00 (.03)		
Prior Specialty Pref.							
Undecided ^R					1.00		
Primary Care					0.79 (.03) ***		
Non-Primary Care					2.05 (.08) ***		
-2LL	58,392	57,616	57,532	56,732	55,405		

Multinomial Logistic Regression Results Comparing the Matriculating Specialty Preference Option of Non-Primary Care versus the Base Outcome of Undecided

CHAPTER 5: DISCUSSION

Chapter Overview

The goal of this, the final chapter, is to explore the meaning of the results of this study. Because there were essentially four models which included many predictors, I start with three tools to facilitate the interpretation and to clarify the major findings. I then discuss my conclusions and bring the dissertation to a close with an examination of the many implications this study may have for the field.

Tools for Synthesizing the Results

The results of the models from Research Questions #2 and #3 were broad, and certainly, there are various approaches to interpretation. In this section, I briefly discuss three ways to digest the results from *Chapter 4*.

Pathways in Preferences

First, looking at *Figures 5.1* to *5.3*, I present the changes in specialty preferences according to each matriculation group. (In fact, these are simple visual representations of the same data found in *Table 4.2.*) In my sense, two possible interpretations arise from these flowcharts. First, we see that those with a non-primary care preference at matriculation were more likely to express the same preference than those who were undecided at matriculation, and *much* more likely to express the same preference than those who had a primary care preference at matriculation. Thus, the difference between matriculating groups in their "stability" of preferences should be noted.

A second interpretation is that by and large, each matriculating specialty group gravitated towards a non-primary care preference by graduation. For example, we see in *Figure 5.1* that 43.6% of those with a primary care preference at matriculation moved to a non-primary care specialty preference by graduation. For those with an undecided preference at matriculation, we see that 58.1% move to non-primary care specialties. Finally, 74.4% of those who initially preferred a non-primary care preference did so at graduation. These flowcharts, while solely focused on the outcome, will help make sense of the interpretation in the following sections.

A Second Pseudo-Longitudinal Approach

I used the models from *Research Questions #2* and *#3* to explore trends at two critical junctures in a medical student's career: first, at matriculation when career expectations are ostensibly malleable and then at graduation when a body of evidence has informed these expectations. While the models are not linked in any mathematical or statistical sense, there is certainly conceptual value in looking at the matriculation and graduation models side by side. I do this in *Table 5.1*, though for the sake of interpretation, I do not reproduce the fitted odds, but use the symbols [+], [-], [?], and [R] to denote a statistically positive relationship with the specialty outcome, a statistically negative relationship with the specialty outcome, no clearly identified relationship with the specialty outcome, and the reference category, respectively.

This table can be used for the remainder of this dissertation as a shorthand tool for interpreting the many results and relationships explored in *Chapter 4*. For example, we see from the [+] symbol that having previous Hospital Work experience proved to be a positive predictor of preferring either primary or non-primary care over being undecided in all models.

"Splitting the Difference"

What strikes one as curious when they read the academic literature on the primary care field is the notion that the *non*-primary care group of students is generally the most commonly used reference group. This should be curious for at least two reasons. First, if it is indeed assumed that those with an interest in primary care careers are truly distinguishable from those with a non-primary care career interest, then what value is there in comparing groups that are *known* to be different in measurable ways? In essence, are we proving what we already know?

Second, as this dissertation has shown, there is no conceptual or data driven reason to specify a binary outcome for this research. About 80% of those undecided at matriculation indicated a specialty preference by graduation. This suggests that medical schools and policy groups should invest in understanding these undecided students, as there may be more leverage to impact their final specialty decisions. In *Figure 5.4*, I show how I have "split the difference" by examining how undecided medical students compare with both groups. By doing so, it is possible to see how a relatively understudied population compares with the *two* kinds of specialty groups, in a way that represents new scholarship on this topic.

Conclusions

With these three guiding frames, I show how the results of this study furthers our understanding of students' specialty preferences during medical school, and what factors may explain the changes in these preferences. Ultimately, it is important to discuss what new knowledge was generated from this study.

Motivational Factors Conclusions

Albanese and colleagues asked the question, "What constitutes a compelling personal characteristic, and which is/are most compelling?" (2003, p. 316). This dissertation has added to the discussion around which motivations and attitudes are most

compelling. And methodologically, this study has shown the advantage of a robust data analysis technique not commonly applied in this field's literature.

In developing an attitudinal instrument, it is important to know its psychometric properties; however, most of the personal and motivational factors reviewed in the literature were developed from an exploratory factorization of a battery of items. This data-driven method does not provide rigorous indicators of a strong instrument; however, according to Wetzel (2012), who reviewed factor analysis methods in the medical education literature, this analytical method is common. For example, Scott et al. used a 27-item battery in a principal components analysis to develop a six factor attitudinal predictor set (2011) while Hauer et al. used the same approach on a 32-item battery to determine the influences on career choice (2008). The fact that these studies—and many others like them—used exploratory factorization methods to develop a parsimonious variable set is laudable, but we should be aware of its limitations. To highlight the most important limitation, exploratory methods can be used to develop a *possible* factor solution, but it should not be used to determine whether this solution is optimal, or to gauge replication in future studies.

Thus, an original contribution that this study has provided to the literature is a rigorously tested set of motivational factors. Using the EFA/CFA split sample approach appears to be a novel methodological application of developing an attitudinal or motivational battery, and it would be sensible for future researchers to apply this method

(Reis & Judd, 2000). The results would gain additional credibility, and more consistent inter-study patterns would emerge.

Nonetheless, my goal in *Research Question # 1* was to use the MSQ's 22-item battery to explain how career motivations were related to specialty decision-making. The factors that I developed appear to enjoy consistency with the many of the factors in the reviewed literature. Without analyzing each of the included items, I noted that my Professional Benefits connected closely with other recent studies (Scott et al., 2007; Compton et al., 2008; Jeffe, Whelan, & Andriole, 2010), as did Intellectual Intrigue (Kassebaum & Szenas, 1994; Clinite et al., 2013), and Doctor-Patient Relationship (Jeffe, Whelan, & Andriole, 2010; Haurer et al., 2008; Scott et al., 2011).

Unsurprisingly, this study had much in common with Jeffe and colleagues study as it used the same MSQ battery. However, their factorization led me to suspect that they may have underfactored the battery, and indeed, I found that two lifestyle related motivations could be extracted and confirmed. This is critical, given Dorsey, Jarjoura, and Rutecki's (2005) assessment of the relationships between perceptions of specialty lifestyles and specialty plans. First, the Lifestyle Management factor was not a strong predictor in the final models, and this may be due to the fact that only two items loaded significantly on it. That said, it does respond to the research gap of how students might balance career and personal desires and sensitizes other, more robust findings of lifestyle balancing and decision-making (Clinite et al., 2014; Scott et al., 2007). The second factor, Command & Control, represents a conceptually novel motivational factor in the specialty preference literature. It also demonstrates how student perceptions of the dayto-day routines of a physician might influence future specialty choices.

(However promising these results, it should be noted that current versions of the MSQ do not include this 22-item battery. Rather, a 7-item *Tolerance of Ambiguity* scale, a 14-item *Interpersonal Reactivity Index*, a 6-item *Quality of Life* scale, and a 4-item *Perceived Stress Scale* are all used to measure matriculants' personal characteristics an and motivations [AAMC, 2014b].)

Multinomial Logistic Regression Conclusions

The purpose of this paper was to demonstrate how key characteristics differed between three specialty preference groups, using undecided students as the referent group. In this section, I show what we have learned from the four multinomial logistic models and how the results of this study support, refute, or add new knowledge to previous findings on this issue. Doing so will provide a strong body of evidence for determining their implications.

I used the Phillips-Bennett model (*Figure 2.1*) and the review of the literature to identify the predictors used in this study. Repeatedly, the student's gender appeared to have a strong connection with primary care preferences, as self-reported females had higher fitted odds of preferring primary care specialties over non-primary care specialties. The results of this study show that female students had lower fitted odds of preferring a non-primary care specialty than males at both time points, suggesting that gender plays a

role in preferring more technical specialties, and confirms what has been demonstrated in the literature (Jeffe, Whelan, & Andriole, 2010; Bland, Meuer, & Maldonado, 1995; Senf, Campos-Outcalt, & Kutob, 2003). However, when comparing undecided students to those who preferred primary care specialties, the signs of the fitted odds flipped between the two time points. That is, at matriculation females had higher fitted odds than males (1.25) of preferring primary care than being undecided, but at graduation, they had lower fitted odds than males (.84) of preferring primary care than being undecided. This suggests that gender is an important aspect of one's pre-medical identity. It also categorically confirms that the medical school environment plays a differential role between males and females on career decision-making (Babaria, Bernheim, & Nunez-Smith, 2011). To summarize bluntly, the results of this study support the notion that choosing a specialty is a gendered process.

The second demographic element that I incorporated into the pre-medical identity was the student's race/ethnicity. The literature provided many possible race/ethnicity permutations, and I was sensitive to choosing the categorization that would uncover any population differences while also allowing for results that would inform current diversity efforts in medical education (Nivet, 2011). Looking specifically at graduation, the results show that when compared with White students, Asian, URM, and students of Other/Unknown race/ethnicity had either higher fitted odds of being undecided or the relationship was not determined. For example, in *Table 5.1*, that self-reported Asian students had lower fitted odds than white students of indicating a preference for either primary or non-primary care specialties, suggesting that on average Asian students were

more likely to be undecided than White students. (Additionally, these findings must be evaluated in light of the fact that these two demographic groups made up 86.2% (22.4% Asian; 63.8% White) of the 2014 matriculating class [AAMC, 2014b].) This may suggest that White students, on average, may have more defined specialty preferences, or have access to experiences that help them define these preferences during medical school. While the definition of diversity is likely to evolve with population changes (Page et al., 2013), it is important to recognize that one's race/ethnicity appears to play an minor role in specialty preferences. And ultimately, we should note, these racial/ethnic differences will play a role in determining the primary care physician composition (Xierali, Nivet, & Fair, 2014).

Parental occupation was treated as a socializing variable, because it was assumed that key career related knowledge would be acquired by children of physicians through years of informal knowledge transfers. These children would, therefore, be more likely to express career intentions. It was considered to be a kind of community of practice where children would indirectly experience their parent's workplace through informal conversations and through spontaneous interactions with their parents' physician acquaintances (Lentz & Leband, 1989). As a result, it was assumed that children of physician-parents are "more committed to medicine because he or she knows what it means to be a physician" (Nowacek & Sachs, 1900, p.143). For children of Physician-Parents, the literature also supported the notion that they would carry negative perceptions of the primary care (Jeffe, Whelan, & Andriole 2010). In essence, the premedical identity of children of Physician-Parents was thought to include career-related information and attitudes around medical specialties.

In fact, the results of the parental occupation variable in this study were nuanced and intriguing. For example, children of Physician-Parents were more likely to be undecided at matriculation than children of Other-Parents, which suggests that a greater prior knowledge of medical specialties was associated with a higher prevalence of being undecided. While this makes sense in the primary care model, that they were also more likely to be undecided in the *non*-primary care model was surprising. This may suggest that children of Physician-Parents knew that in medical school, they would be able to try on "trying on possible selves" through clinical experiences (Burack et al., 1997, p. 534). One could also speculate that despite having a greater knowledge of the specialties, these respondents were just as likely to be undecided as children of other parents. That is to say, they had greater knowledge, but ultimately, similar patterns of preferences as respondents of parents with different occupations.

The second set of socialization variables were the professional experience predictors Hospital Work and College Research. The former was found to be a positive predictor in all of the final models, suggesting that those with experience in a hospital were more likely to express a specialty preference (of any kind) than those who didn't have that experience. Although it is not possible to know the nature of the hospital work (patient or physician-related, lab-based, paid or volunteer position, et cetera) it does suggest that having a hospital-based experience plays a role in forming specialty

160

preferences, which is supported by the literature on physician shadowing and physician interactions (Kitsis & Goldsammler, 2013). This also supports the notion that communities of practice that are situated in the professional context have a particularly influential role on how individuals form professional identities (Osler, 1893).

Research interest variables have played a key role in studies of specialty preferences (Hauer et al., 2008; Scott et al., 2011; Jeffe, Whelan, & Andriole, 2010), however the results of this study expand on these findings by demonstrating that actual experience plays a role in determining specialty preferences as well. In particular, respondents with college research experience were found to have the same fitted odds of preferring non-primary care as those without a college research experience, but interestingly students with these experiences were more likely to be undecided than to prefer a primary care specialty. In sum, having a college research experience appears to have a primarily negative influence on students' perceptions or interest in primary care. This could be due to the fact that non-primary care specialties are seen as more technical (Kassebaum & Szenas, 1994), but without knowing the exact nature of the experience, we can only speculate that laboratory based communities of practice seem to lead students away from primary care specialties because of perceptions there were exposed to through these experiences.

The novel most novel aspect of this study was that I inspected specialty preferences at two time points. Doing so allowed for an inspection of the degree to which specialty preferences change during medical school (*Figures 5.1* to *5.3*), as well as an

analysis of whether prior specialty preferences were strong predictors of one's graduating specialty preference (*Table 5.1*). In my review of the literature, only two studies specifically looked at this trend (Babbott et al., 1988; Compton et al., 2008). Compton et al. found that about 47% of those who initially preferred primary care specialties did so at graduation, compared with 37.2% in this study's sample (2008, p. 1099). They also found that 28% of undecided respondents changed to a primary care preference, while in this sample, only 23.5% did so. Together, this suggests this this sample's specialty preferences are not as stable, though it confirms Compton et al.'s finding that about a quarter of initially undecided students moved to a primary care preference by graduation.

In my theoretical framework, these undecided students have a pre-medical identity that is not as well informed by previous socializing experiences that would provide them with a knowledge base or strong specialty inclinations. It was assumed that the medical school environment played a stronger role on their specialty choices than the other "decided" groups, and this appears to be the case. This was confirmed by the large and statistically significant fitted odds in the final models. Looking at *Table 4.11 (A)*, we see that those who expressed a specialty preference at matriculation, no matter what the preference was, had higher fitted odds than undecided students of preferring primary care. *Table 4.11 (B)*, we note that those who expressed a non-primary care preference had about twice the fitted odds of previously undecided students for preferring non-primary care at graduation, while previously undecided students were more likely to choose non-primary care at graduation than those who had previously expressed an interest in primary care. Taken together, I suspect that any movement away from primary care is

actually done so by previously undecided students. This has undeniable implications for the field, as I discuss in the following sections.

The results from this "longitudinal" approach have also allowed us to examine any "slide" of student interest away from primary care (Erickson et al., 2013; Jeffe, Whelan, & Andriole, 2010). In terms of the methods used, we should consider whether a survey is truly the best way to gauge a student's preference for a specialty. On the one hand, social desirability bias to indicate an interest in primary care may have played a role in this study, as respondents may have felt that there was a "correct" answer. However, we should additionally question whether this "slide" could be due to students indicating a primary care specialty preference simply because they were familiar with those specialties. For example, looking at the list of specialties presented to students in the 2014 MSQ, we could conclude that a truly undecided student might choose a specialty that they were familiar with to appear knowledgeable (AAMC, 2014b). If this were the case, then the proportion of truly PC-committed students at matriculation would be overestimated, and the "slide" would be less extreme.

Finally, this study's development of five motivational factors showed that there are methodological areas of improvement for the field, and their inclusion in the final multinomial logistic regression models can also inform the literature. The three factors commonly found in the literature— Professional Benefits, Intellectual, and Doctor-Patient Relationship—produced very compelling results. First, many studies have found "prestige" or "financial motivations" to be a driving factor in distinguishing those with a

163

primary care preference versus those with non-primary care preference (Scott et al., 2007; Compton et al., 2008; Jeffe, Whelan, & Andriole, 2010), and the ANOVA results (*Table* 4.9) to confirm this as well. However, when comparing *undecided* students to these two groups, I did not find a statistically significant relationship. In other words, having greater financial and social status motivations was not associated with a greater likelihood of preferring primary care or non-primary care specialties. This may suggest that students with unformed specialty preferences have pre-medical identities that are less structured by social and financial career goals. Alternatively, we should consider whether "prestige," especially its social elements, is interpreted the same way by all students. For example, there is a connection between expressing a primary care interest and expressing an interest in practicing in a rural location (Click, 2013), and we may wonder if students see being a primary care physician in these areas to be prestigious.

The Intellectual Intrigue factor was formed by items related to technical and analytical expertise, and as in the literature, it seems to have an influence on student's specialty preferences, where those with a non-primary care preference ascribe greater importance to intellectual challenges (Jeffe, Whelan, & Andriole, 2010; Scott et al., 2012). In this study, I found that those with greater Intellectual Intrigue factor scores had greater fitted odds (1.35) of indicating a non-primary care preference at matriculation and at graduation (1.50). Unlike Professional Benefits, this factor clearly distinguishes undecided students from those with a non-primary care preference. Interestingly, I found diverging relationships in the two primary care models. For example, at matriculation, students with higher Intellectual Intrigue factor scores had lower fitted odds (.80) of preferring primary care specialties. At graduation, however, higher factor scores were positively associated with a primary care preference. This suggest that students enter medical school with the perception that primary care specialties are not as challenging, but throughout their medical school experiences gain different perception.

By far the most interesting factor, and most original contribution to the literature came from the Doctor-Patient Relationship, where students' scores strongly distinguished undecided students from those with a non-primary care preference at matriculation (.43) and at graduation (.58). To state it bluntly, indicated a stronger "service orientation" was associated with significant lower odds of preferring non-primary care specialties. Other researchers who have compared primary care to non-primary care preferences have found similarly alarming odds-ratios (Jeffe, Whelan, & Andriole, 2010; Hauer et al., 2010; Kassebaum, Senzas, & Schuchert, 1996). As these researchers concluded, this factor highlights that "altruistic" motivations are clearly not as pronounced for those with a nonprimary care preference. Comparing undecided students to those with a primary care preference revealed equaling interesting results. At matriculation, higher Doctor-Patient Relationship factor scores were associated with extremely higher fitted odds (2.41) of preferring primary care specialties. Conversely, at graduation, no distinction could be made between these two groups. This suggests that while undecided students may have weaker "altruistic" career motivations than those with a primary care preference at matriculation, the two groups' altruistic motivations were similar by graduation.

While the review of the literature revealed the importance of lifestyle related factors in determining specialty preferences (Dorsey, Jarjoura, & Ruteckis, 2005; Scott et al., 2012; Hauer et al., 2008; Lambert & Holmboe, 2007), the Lifestyle Management factor did not distinguish specialty preference groups. This lack of predictive power is probably due to the fact that it contained only two factors, and not due to the true underlying motivations of medical student.

Conversely, the Command & Control variable did strongly distinguish undecided students from those with a non-primary care specialty preference. For example, the fitted odds for Command & Control factor scores at matriculation were 2.01 in the matriculating model and 1.22 in the graduating model. These results suggest that undecided students enter medical school with different career motivations than those with a non-primary care preference. Importantly, the items in the Command & Control factor were related to the day-to-day operations of physicians, which may suggest that the latent motivations were about perception of the physician role as much as they were about career motivations.

Limitations

There are a number of questions raised, but left unanswered, by this study. Some of these questions stem from the limitations of this study.

We should first consider the study sample: while some 31,425 medical school graduates responded to the PMQ, MSQ, and GQ, censoring of the full TrEMUR dataset

was required to reach this sample. In particular, we must consider that students who indicated that they were matriculating into an MDPhD were not included as their career motivations were anticipated to fundamentally skew their specialty choices (Jeffe, et al., 2014; Brass, et al., 2008). Additionally, students who did not complete the either the MSQ or GQ survey were eliminated for statistical purposes. As school mission may influence specialty preferences, and the MSQ response rate varies greatly between medical schools, how students are encouraged to complete the survey is important for this study (AAMC, 2013a). Although the censored respondents appear to have similar characteristics as the identified study sample, such a large "attrition" should be acknowledged as a limitation. Finally, because the motivational battery was not included in the 2007 version of the MSQ, these respondents, the most current cohort in the TrEMUR dataset, were eliminated. This necessary elimination highlights the importance of the period in which this data was collected, and the reader should consider how and to what extent the current cohort of medical students may be different from medical students who matriculated from 2001 to 2006, as discussed above.

Secondly, we should consider the data collection tools themselves. The MSQ, which students can voluntarily complete online from June to September of their first year of medical school, represents a large battery of items (AAMC, 2014b). As with any large survey or questionnaire that stresses one's cognitive load, we must consider that some students may have experienced "survey fatigue" and their responses may not have been an accurate indication of their true perceptions or history of experiences. This is true of any survey, however, I acknowledge it because much of this study dealt with the 22-item

battery that was a self-inspection of reasons for pursuing medicine as a career. Another concern is that the motivational factors were only measured at matriculation, yet I used them to find independent fitted odds of one's *graduating* specialty preference. Their role in explaining these graduating preferences was diminished, and this suggests that one's values at matriculation do in fact change over time.

We must also ask how these 31,425 respondents compare to today's population of medical students. Looking specifically at *Table 4.1*, we see that 51.6% of the final sample self-identified as female, compared with 50.2% of the 2014 national matriculating class, the most recent data available (AAMC, 2014b, p. 23). Similarly, we see that White students represented 67.8% of the final sample versus 63.8% in the 2014 national matriculating class. Additionally, Asian, URM, and Other/Unknown students represented 16.3%, 12.9%, and 3.0% of the sample, respectively compared with the 2014 national matriculating class which was 22.4% Asian, 16.3% URM, and 3.9% Other. These demographic comparisons demonstrate that my sample has slightly more White students than the most recent matriculating class, while the proportion of Asian and URM students is lower.

For the experiential variables, I found that 70.1% of students in my sample indicated that they had "paid or volunteer work in hospitals, medical clinics, or labs," however the matriculating class of 2014 did not see that same prompt. Rather, 91.5% responded that they had "volunteered in the healthcare field" and 93.5% had "shadowed a physician or other healthcare professional" suggesting that nearly all students some kind

of previous exposure to the healthcare field before entering medical school (AAMC, 2014b, p. 6). Secondly, 52.5% of the sample indicated that they had undertaken a "laboratory research apprenticeship for college students" and 58.6% of the 2014 matriculating class responded affirmatively to this same prompt. This indicates an increase in hands-on laboratory experiences, and further suggests that important experiential factors in this dissertation are already by leveraged with today's population pf students. (Unfortunately, there was no data for parental occupation in the *MSQ 2014 All Schools Summary Report.*)

There were vast differences between the sample and the 2014 matriculating class in terms of their specialty preferences at matriculation. In this study's sample, 22.6%, 39.6%, and 37.7% of the sample preferred primary care, non-primary care, or were undecided at matriculation, respectively. In contrast, 35.2% of the 2014 matriculating class preferred a primary care specialty, though they included Obstetrics and Gynecology which accounted for 4.3% of the total class. It is also surprising that only 15.8% of these students indicated that they were undecided about their specialty at matriculation, suggesting a difference of 58.2% from this study's sample; however, this can be explained by the different item prompts used in the survey tools and these proportions should not be directly compared.

Taken together, this comparison with the study's sample shows that today's students, as a whole: a) are more racially/ethnically diverse, b) have more healthcare and

laboratory experiences before matriculating, and c) may have stronger primary care career inclinations and are less likely to be undecided.

I will again cite Lawson, Hoban, and Mazmanian (2004, p. S38) who claim that researchers should "design studies with a specific field as a dependent variable" and that "research has shown that physicians within these three fields are not a homogenous group" (Lawson & Hoban, 2003, p. 70). In this study, I categorically rejected this sensible piece of advice because I wanted to explore population trends at the highest possible level, and because the number of results with separate internal medicine, pediatrics, and family medicine models would have been overwhelming. Nonetheless, this study can serve as a benchmark for future specialty-specific studies.

Finally, asking students about their specialty preference at matriculation and graduation is problematic, but for different reasons. For matriculation, I question whether students have truly settled on a preference by this stage of their career. For the 22.6% (7,103/31,425) of students who indicated a primary care preference at matriculation, I suspect that a proportion *truly* had this preference, another proportion had another preference but acquiesced to present a positive image, and another proportion didn't have any preference, but recognized the presented primary care specialties and chose them anyway. For this last group, we should pause and reflect on what bodies of knowledge students had when responding to this questionnaire, or if they even were aware that certification is an important professional step in medicine. Did popular culture inform their response? What about personal interactions with doctors?

170

Turning the graduation and the GQ, both my committee and I found it strange that 15.3% (4,821/31,425)—nearly one in seven students—were still undecided by graduation. However, to interpret the GQ as an "exit survey" is incorrect, as students may complete the survey from February to June (AAMC, 2014a). Therefore, some students may have completed the survey before "the match" (National Resident Matching Program [NRMP]), which has undeniable consequences for this study. That is, if students completed the GQ before the match, we could understand how their preferences might still be undecided. However, if they completed the GQ *after* the match, we really must wonder why (or even *how*) most graduates of LCME-accredited institutions could be undecided. This uncertainty cloud represents, therefore, a limitation to the interpretations of the results.

Lastly, and most critically, while these findings can be considered representative of the population of medical graduates who took the PMQ between 2001 and 2006, the findings should not be considered as causal in any way. For example, those with higher Doctor-Patient Relationship factor scores were more likely to prefer primary care, but this does not suggest that this factor predicts or causes such a preference. Instead, the data and tools discussed in this study can only indicate which population characteristics are more prevalent or more strongly expressed between predetermined groups of individuals.

Implications
In this, the final section, I discuss how my interpretation of the results within the integrative conceptual model has implications for the field, and directly responds to the problem of interest described in *Chapter 1*. Three potential kinds of implications are explored: implications for research, policy, and practice.

Implications for Research

This study was based on the assumption that specialty preferences are influenced by one's social experience. The socialization theories applied in the conceptual model have a number of assumptions themselves. For one, we suspected that each of the experiential variables (parent occupation, hospital experience, and college research) would impact individuals in unique ways (Weidman, Twale, & Stein, 2001; Hafler et al., 2011), but that in general, they afforded an opportunity to develop a professional identity. It was also assumed that the development of one's pre-medical identity was non-linear (Bennett & Philips, 2010). That is, one could have a strong orientation towards a primary care specialty, but through various experiences begin to abandon this career interest. To state it differently, medical students' career interests can be unstructured and are thought to incorporate new experiences constantly.

The reason that the pre-medical identity use integrated into the conceptual model was that it aligned with the "pinnacle" of the process of professional socialization: identity formation (Austin, 2002; Weidman, Twale, & Stein, 2001). Rather than just simple knowledge transfers (Lentz & Laband, 1989), having a situated experience allows the novice participant to see themselves in a role (Burack et al., 1997) and practice the professional and social norms associated with that role. Because location plays such an important role in forming an identity (Lave & Wenger, 1991), it was important that the experiential variables were situated. Parental occupation was chosen because we wanted to see if children of physicians were exposed to medical communities of practice in the home (Nowacek and Sachs, 1990; Kleshinski et al., 2009), and whether these children are exposed to knowledge or perceptions of the specialties (Colquitt et al., 1996). The results did show that children of physician parents indicated slightly different specialty preferences than children of parents with other occupations at matriculation. However, it remains unclear what kinds of advanced knowledge led them to be more "undecided" at matriculation. Was it that they had a greater knowledge of the variety of specialties and were truly undecided? Or, were they undecided because their physician-parents had encouraged them to not decide on a specialty so early?

In contrast to one's home-based parental influence, the Hospital Work and College Research variables clearly represent socialization experiences that are situated in a professional context, where students were allowed to try on professional roles. That having a prior hospital experience was so strongly associated with expressing a career preference suggests that it was an important experience for students' identity development. Having a college research experience played a less influential role in the model, and this may be explained by the fact that the laboratory context could have been of a fundamentally different nature than the hospital setting. That is, students with research experience may have developed a "researcher identity" but not conceptually connected that identity to their pre-medical "physician identity." This should encourage future researchers to identify the nature, location, and intensity of professional socialization experiences so that they can be linked more closely to outcomes of interest.

Turning to the research implications of these interpretations, this study was designed as a response to the gap in the literature created by the primary care versus nonprimary care duality (e.g. Colquitt et al., 1996), which limits the outcomes of interest to two categorically different groups of individuals. Although this outcome structure is common, the reviewed studies that *did* use undecided as an outcome category (e.g. Compton et al., 2008; Jeffe, Whelan, & Andriole, 2010) failed to explore why or how these students differed from students with defined career plans and did not provide a sound rationale for their inclusion. Responding to this gap in the literature, my conclusions indicated that there are important ways that undecided students differ from those with a primary care and non-primary care preference. Moreover, interpreting these findings through socialization theory has revealed new knowledge and potential interventions, which I discuss below. As such, this study supports the use of undecided in the outcome structure as it provides a sensible alternative to the traditional binary outcome structure. It also provides future researches with probable hypotheses for most common predictors.

With that in mind, the proportion of undecided respondents at graduation suggests that how future researchers structure the outcome category undecided may impact the findings. For a survey analysis, they may consider using the same prompt as the AAMC's MSQ and GQ (see *Figure 3.2* and *3.3*), or may consider delineating those who are truly undecided from those who have not yet considered their career preferences. That 15.3% (4,821/31,425) of students were still undecided at graduation raises questions, and highlights the importance of accurately capturing undecided specialty preferences. It also brings into question the interpretation of a specialty indication as a commitment.

Second, while this study was not truly longitudinal in nature, the results nonetheless support a longitudinal framework. Career preference "stability" was indirectly explored in this study, particularly in the third research question, where the predictor "Prior Specialty Preference" was considered. In fact, the fitted odds for these predictors were extremely statistically significant (see *Table 4.11 A* and *B*) and suggest that prior specialty preference plays a non-trivial role in graduates career decisionmaking. However, one could also argue that the stability from matriculation to graduation was relatively low (37.2% for primary care, 18.4% for undecided, and 74.7% for nonprimary care). While one's expectation for career preference stability determines whether this is a "good" or "bad" level, an implication of these findings is that future researchers would be judicious in capturing specialty preferences at key moments in medical students' lives. Borrowing from the literature, Scott and colleagues surveyed students during the "first 2 weeks of their programs, at the end of their preclinical training, and again at the end of their clinical training" (2007, p. 1956), though other data capture strategies such as an end-of-year survey may also be sensible, and more practical.

The findings also imply that future research should focus on how motivations change throughout one's medical school experience. As I discussed above, the motivational factors had stronger relationships with outcome in the matriculating models than in the graduation models. Thus, an implication of this study is that researchers should apply the same battery of motivational factors at each data capture to understand how changes in motivations are associated with changes in specialty preferences. In terms of developing this motivational battery, nearly all studies that were reviewed developed a motivational structure through an exploratory factor analysis (e.g. Jeffe, Whelan, & Andriole, 2010; Hauer et al., 2010; Scott et al., 2011), which can be used to develop one *possible* factor structure. This study is, to my knowledge, the first study to apply a split-sample exploratory-confirmatory analysis of motivational factors, and future researchers may consider using this factor structure develop future motivational batteries for use with medical students. With that in mind, this analysis was a *post hoc* use of a pre-existing tool developed by the AAMC, and the Lifestyle Management and Command & Control factors, while a novel contribution to the field, were comprised of just two and three items, respectively. As there is growing evidence that lifestyle choices are key to current medical students' career interests (Dorsey et al., 2005; Newton, Grayson, & Thompson, 2005), these findings suggest that focusing a motivational battery around lifestyle characteristics is essential.

Research and theory recommendations:

- Future investigations of specialty interest during medical school should, if possible, focus on undecided students and be mindful of the reasons that students would indicate that they are undecided in interpreting the results.
- The results of this study confirm that specialty preferences and career motivations shift dramatically during one's medical school experience. Researchers should consider longitudinal approaches, with data collection corresponding with influential experiences. Additionally, researchers should focus on developing a motivational battery that is reliable and valid for use at multiple time periods, and reflects the motivations of *current* medical students.
- A multi-institution, qualitative research approach is opportune. This study, and others like it, have presented important population trends, but have yet to thoroughly unpack the mechanisms by which specialty preferences are informed or changed. A more precise understanding how students incorporate experiences and career advice into their body of knowledge would provide policy makers and practioners with specific areas to target interventions.

Implications for Policy

The results of this dissertation may have implications for groups concerned about primary care pipeline policies, such as the Council on Graduate Medical Education (COGME), the HRSA, the AAMC's Center for Workforce Studies, the AMA, the Urban Institute, the Agency for Healthcare Research and Quality (AHRQ), and the Robert Graham Center to name just a few. Truly, there are many groups—governmental and non-governmental—eager for data that could influence institutional, state, and national policies. In this section, I discuss how these results might be used to inform or influence the policies and policy recommendations of these groups.

At the core of the discussion circulating amongst researchers and policymakers in this arena is whether our medical education system is producing a sufficient number of physicians interested and prepared for a primary care career. To succinctly restate the issue, the scarcity of primary care physicians: a) is impacted by the way that medical students are introduced to specialties, b) is further constrained by a limited number of residencies, c) is expected to worsen by the year 2025, and d) has implications for the state of healthcare delivery.

The descriptive statistics from this study have shown that a substantial proportion of matriculating medical students (22.6%; 7,103/31,425) and graduating medical students (23.5%; 7,400/31,425) prefer a primary care specialty. While *Figures 5.1* to *5.3* highlighted the change in preferences between these two points in time, it may be reasonable to assume that undergraduate medical schools—as a whole—could be expected to sustain about one-quarter of all medical students' interest in a primary care preference. (Given *my* definition of primary care, which is one of many tenable categorizations.) As econometric analyses of physician shortages become more sophisticated, these figures may help researchers create more robust models of the primary care physician supply chain. The shortage estimates from the 2006, 2010, and 2015 workforce studies commissioned by the AAMC varied greatly, and it is clear that

having a complete picture of the environment in which students plan their careers is vital. This study shows the relationships between key variables and career decision-making at critical moments in their career training.

Additionally, this study may also have implications for third party organizations that either independently research primary care pipeline issues or evaluate governmental and institutional policies. For example, the Robert Wood Johnson Foundation recently funded a policy analysis by Pauly, Naylor, and Weiner that synthesized the literature of primary care shortages and provided a list of possible strategies to alleviate it (2014.) One of these strategies was to increase "the supply of primary care physicians via financial and non-financial incentives" (p. 7), and this study shows very intentionally that there are five such motivational factors that could be leveraged to increase this supply through changes to such incentives. Or, consider a recent study supported by the Robert Graham Center that inspected school mission in terms of workforce outcomes, in particular primary care physician output (Morley et al., 2015). This study could help explain how institutional goals and strategies support or inhibit the preparation of primary care physicians. Finally, as many of these organizations have developed and currently maintain databases of primary and secondary research on medical education trends, they may consider using this study to formulate a policy and advocacy agenda around professional socialization in medical schools.

The development of a policy research agenda aside, the results of this study also show that, on average, career intentions significantly shift during one's medical school experience. For policymakers, this can be at once troublesome and exciting, depending on their point of view. Taking the second perspective, this study demonstrates tangible ways of identifying students who are undecided versus having a primary care or nonprimary specialty preference, and policies could be developed to support students who are undecided. I will take one example from the findings to exemplify such a strategy:

In *Table 5.1* we see that, at graduation, self-identified females are more likely to be undecided than to indicate a non-primary care preference, but are *also* more likely to be undecided than to have primary care preference. As discussed in the literature review, women are more likely to prefer primary care specialties than non-primary care specialties (e.g. Jeffe, Whelan, & Andriole, 2011), and these findings suggest that a national policy to support undecided female students to consider a primary care specialty would be wise. As the impact of role models may have a stronger impact on women than on men (Wendel, Godellas, & Prinz, 2003), national medical education leaders should promote policies that help identify such undecided female students early on and match them with faculty members and experiences that are shown to promote an interest in primary care. Burack and colleagues noted in a study on specialty preferences that choosing a specialty is a matter of "trying on possible selves" (1997, p. 534.), and this example shows how intentionally addressing the information gaps of undecided female medical students could lead to a greater proportion of medical students leaving with a proclivity for primary care careers. (I keep this paragraph brief as there are many such examples that could be drawn from the data and because many of the strategies one could formulate as national policies are treated as institutional practices below.)

As stated in the introduction, even if a greater proportion of medical students do exit medical school with such a proclivity, it is also reasonable to question whether the subsequent step in the pathway, GME residencies, would be equipped to handle such an influx. The most recent match results from the NRMP categorically suggest 'no.' The 2015 match saw 96.2% of all PGY-1 positions filled, with family medicine at 95.1% filled, pediatrics at 99.5% filled, and internal medicine at 98.9% filled (NRMP, 2015, p.4). These figures certainly confirm the presence of the bottleneck that many researchers have identified as the main barrier to systemic change. We should also consider the "success" of the match, that is, whether students *stated* specialty preference actually matched with the program they were accepted into. Based on 2014 match data of US seniors, 96.4% (1,327/1,367) students who preferred family medicine matched with that specialty. This compares with a 96.1% (1,818/1,890) success rate for pediatrics, and 96.9% (3,300/3,405) for internal medicine (NRMP, 2014b, p. 3). Taken together, we see that not only are these residencies at capacity, but those interested in these specialties are also generally matching to these residencies.

Based on the findings of this study and these NRMP trends, it is sensible to conclude that substantial increases in students graduating from medical school with an interest in primary care would most only result in incremental changes, at best, of residents leaving with the certification to do so. Additionally, with the number of PGY-1 positions having increased only incrementally themselves since 2002, this is likely to be the case (NRMP, 2014c). Therefore, the most obvious policy response to this bottleneck

is simply that the number of residency positions should be increased. This stance has been taken by the AAMC, who state that "to ensure an adequate physician workforce, the Medicare freeze on residency training must end," essentially calling on the largest funder of residency positions, the Department of Health and Human Services (HHS), to increase their support (AAMC, 2014c, p.1). The American Academy of Family Physicians (AAFP) position that "Medicare should also consider giving an incentive back to hospitals for every primary care physician they produce" represents a more aggressive financial strategy towards rectifying the future primary care shortage (2013, p.2). Other primary care policy groups echo this request, such as the *GME Initiative* which calls for 40% of all residencies to produce primary care physicians (Voorhees et al., 2013).

Two even more radical policies should be considered. The *Veterans Access*, *Choice and Accountability Act of 2014*, or the *Choice Act*, is intended to "provide timely, high-quality health care for Veterans" (Veterans Affairs [VA], 2014a, p.1). One provision mandated the VA to increase its number of residency positions by up to 1,500 over a five year period, and place particular emphasis on "residency positions that improve Veterans" access to primary care, mental health, and other specialties the Secretary deems appropriate" (VA, 2014b, p.3). Given that between 2014 and 2015, the number of PGY-1 positions increase by 615 (2.3%; 26,678 in 2014, and 27,293 in 2015), such a policy would play a significant role in reducing the primary care physician shortage (NRMP, 2014a; NRMP, 2015). This expenditure, targeted at improving the healthcare service of a high-needs population, may provide a policy guide for larger national changes to the GME residency if successful.

Another radical policy provides an altogether different approach to alleviating the primary physician shortage, in particular for primary care health professional shortage areas (HPSA), typically rural and lower-income areas (HRSA, 2015). Missouri, fraught with HSPAs, recently passed Senate Bill 716, which would allow medical school graduates the opportunity to practice primary care medicine in rural areas without completing a residency under the title of "assistant physician" (Office of Missouri Governor Jay Nixon, 2015). Although they would be practicing medicine under the supervision and responsibility of a "collaborating physician," there is fierce opposition from the AAFP and the American Association of Physicians Assistants (AAPA), who question the quality of care that these practitioners would deliver as these assistant physicians would not be fully trained or prepared for the many challenges of a family physician (AAPA, 2015). Certainly, unanswered questions remain and oversight systems would be required under this scenario, and the extreme shortage of primary care physicians has resulted in such a radical policy. And, as the results of this study have shown, there are tangible ways to prepare medical students for primary care careers even non-traditional ones—provided there are opportunities awaiting them.

Importantly, both the *Choice Act* and Missouri's *Senate Bill 716* highlight the need to align medical students' professional goals, societal needs, and public policy into a single policy framework. While other developed nations have aligned physician training policies to the needs of its people and capacities of its training system (for example, the National Health Service's *Health Education England*), the US lacks such a level of

coordination. Noting the implications, John Iglehart recently wrote that "the absence of health-workforce planning, a hallmark of the freewheeling U.S. market economy, may come back to haunt policymakers, particularly when physician shortages become more apparent as the ACA's coverage expansion takes hold" (2013, p. 299). This study could be used to develop an "undergraduate to practicing physician" pipeline policy.

Policy recommendations:

- As the production of medical students at LCME-accredited institutions increases, policy groups should help medical school admission and culture teams identify best practices for aligning undecided students with potential primary care careers through experiences and mentorships.
- To that end, internal medicine, pediatrics, and family medicine residencies are filled, students who preferred that specialty matched at high rates to those residencies. Expansion of federally-supported residency positions is the most straightforward approach; however, this should be considered in light of complimentary policies such as Missouri's assistant physician proposal, redefining the role of nurse practioners, and a rethinking of primary care service staffing models.
- This work should inform the advocacy efforts of groups interested in increasing the production of primary care physicians. Specifically, they can point to the leverage points for increasing interest in a primary care career during medical school, and the fact the GME residency bottleneck would prevent additional

medical school graduates from achieving the desired career outcome. Aligning social needs and training efforts must take a systems-level approach.

Implications for Practice

The new knowledge generated in this study can also be applied to the existing practices, particularly within medical schools. I recognize two immediate applications: assessing fit during applications, and developing an intentional response to the so-called "hidden curriculum."

Assessing fit. In regards to the admissions, it is clear that the results of this study could be used to inform and improve admission committees' bodies of evidence. Certainly, which metrics should be used in determining admission has implications for the composition of the physician workforce, and the debate around which metrics to use is both contentious and well documented (e.g. Ferguson, James, & Madeley, 2002; Siu, & Reiter, 2009).

Within the field, most researchers and professionals recognize the MCAT as *the* critical tool of assessment for determining admission, with predictive models supporting its utility for this purpose (Julian, 2005; Dunleavy, Kroopnick, Dowd, Searcy, & Zhao, 2013). Despite this, there have been calls—based on sound justification – for medical schools to broaden the criteria. Dr. Jordan Cohen stated that "admission committees might well find many instances in which truly compelling personal characteristics would trump one or two isolated blemishes in the academic record" (2001, np), while William

McGaghie more forcefully noted that "admissions committees give lip-service to the importance of personal and social traits...one reason why candidates' personal qualities receive little attention is the difficulty of measuring these attributes quantitatively" (1990, p. 137). The results of this study speak to such person qualities.

Personal qualities are known to distinguish those with a preference for primary care from others (Jeffe et al, 2010), and as Rabinowitz succinctly put it, "identifying and giving weight to these factors in the medical school admission process is likely to increase the number of graduates who choose generalist specialties" (2001, p. S39). The research supports this supposition: Martini, Veloski, Barzansky, Xu, and Fields found that the criteria for admission played the strongest role in predicting students' choice of generalist careers (1994).

Herein lays the first major implication for practice: the pre-medical experiences and motivational factors explored in this study may be used to assess the individual qualities that are associated with both short and long-term interest in a primary care career. In sum, the predictors in this study can be used to determine a student's "fit" for a medical school. Specifically looking at the students who indicated an undecided specialty preference, we can see that particular motivational factors play a role in preferences.

For example, Intellectual Intrigue, Command & Control, and Doctor-Patient Relationship variables all statistically distinguished undecided students from those with a non-primary care preference at matriculation and graduation. Thus, it may be in the best interest of medical schools with a specific mission to produce primary care generalists to inspect these qualities during the admissions process, and then make strategic decisions. However, because these motivational factors did not as strongly distinguish undecided students from those with a primary care preference (signaled in *Table 5.1* by the many [?] symbols), it may prove more difficult to differentiate students who are truly undecided from those with a primary care preference during the admissions process. Additionally, that the two pre-medical experiential variables (Hospital Experience, College Research) had statistically significant relationships with the outcomes of interest suggests that these could be used as part of the admissions criteria. Finally, and perhaps most importantly, students with a defined career preference at matriculation were more likely to prefer primary care than to be undecided at graduation. Similar to previous arguments made by Colquitt and colleagues (1996), admissions committees should inspect potential students' career plans as well as the basis for these plans, as they have long term implications for students' specialty preferences.

Recommendations for practice:

- At schools with a strong focus on preparing students for primary care careers, admissions should look for students with a strong interest in the Doctor-Patient Relationship and students with experience working in hospitals. Naturally, they should look for students' with a strong interest in primary care, as this was found to be the strongest indicator of long-term interest in primary care.
- These medical schools can also use results of this dissertation to develop a market strategy and recruitment materials to better inform potential students of the

institutional mission and training opportunities that would match their personal and career interests. Additionally, these marketing materials could be used to dispel misrepresentations of primary care careers to broaden the potential candidate pool.

(As an aside, although it may be tempting to discuss the role that these themes could play at the undergraduate or even high school level, I urge the reader to note that this dissertation focuses on the population of medical graduates, and does not compare them to non-admitted students. Thus, while the pre-medical familial and professional socialization experiences discussed are germane to developing future medical students' career plans, these results do not indicate what influences or experience impact *admission into* or *graduation from* medical school, and should not be interpreted as such. Moreover, unless an undergraduate program has an explicit responsibility to prepare primary care physicians, even the results regarding college-level experiences are not germane to undergraduate programs.)

Hidden Curriculum. The results of this study also respond directly to the concerned about the impact that the "hidden curriculum" of medical schools has on their students career decisions. To provide a working definition, the hidden curriculum encapsulates "the set of influences that function at the level of organisational structure and culture including, for example, implicit rules to survive the institution such as customs, rituals, and taken for granted aspects" (Lempp & Seales, 2004, p. 770).

Although the concept of a hidden curriculum is by no means found only in medical education, it is not unreasonable to assume that it is more pervasive in medical schools than in other professional or graduate institutions. To that point, a leading thinker on this issue, Frederick Hafferty succinctly describes the hidden curriculum as the "fundamental distinction between what students are taught, and what they learn" (1998, p. 404), and notes the ubiquity of such informal learning; from the construction of a new building to faculty composition, students receive, interpret, and internalize clues about "what matters." To that end, how resources are distributed, who comprises the leadership team, and what the curriculum emphasizes all provide medical students with implicit messages about what is valued in their professional community.

Given the importance of developing a professional identity during one's undergraduate medical education (Sinclair, 1997), it should not be surprising that students are searching broadly for clues and indicators of career success, developing conceptual hierarchies, and seeking out experiences that will help them achieve their goals. Without a doubt, these clues and indicators provided by the hidden curriculum play a role in determining one's medical specialty preference and ultimately, selection. As my conceptual model indicated, the pre-medical identity is fluid, and various socialization experiences during one's undergraduate medical training surely play a key role solidifying one's *professional* identity. I posit that regardless of whether these experiences are formalized, where the hidden curriculum expresses itself through intentional practices within a community of practice (Wenger, 1998), or completely unintentional and informal, the hidden curriculum has a non-trivial impact on career preferences.

In light of that, the results of this study support the generally support the hypothesis that students are exposed to a hidden curriculum that either limits students ability to pursue experiences that would support their interest in a primary care career (Scott et al, 2011; Zinn et al, 2001) or directly and intentionally inhibits it, for example, through "badmouthing" (Erickson et al., 2013). As the descriptive statistics in *Chapter 4* and the trends in *Figure 5.1* to *5.3* demonstrate, overall student interest in pediatrics and family medicine diminishes during their medical school experience, and the unexpected rise in the number of students interested in general internal medicine at graduation may be explained future sub-specialization or hospitalist plans.

Thus, these results may have profound implications for how medical schools develop structures and supports for students to first *identify* and then *pursue* their preferred specialty. Beginning with identification, the results of this study demonstrated the presence of five motivational factors that could be intentionally integrated into the curriculum or other support structures to disabuse medical students of entrenched, and perhaps fundamentally incorrect, perspectives of medical specialties. As a hypothetical example for how this research could be used, medical school faculty could use items from the "Doctor-Patient Relationship" motivational factor to show undecided students how a non-primary care physician *can* in fact develop these relationships and must develop good bedside manners. The socialization lens used in this dissertation also

suggests that location is important, and is consistent with Osler's main tenet that the knowledge gained by a novice should be at the bedside (1893).

While only some of the motivational factors had significant relationships with specialty preferences, all of them represent conceptually unique motivations that medical students must weigh as they begin identifying specialties suitable to their personalities and career goals. As previously stated, it is uncertain how and from where medical students build their *early* bodies of career knowledge around potential careers, and providing them with both adequate and accurate information would help them align their future choices to their goals. Such a strategy could change an environment where hierarchies of specialties are presented to medical students (Erickson et al, 2013), even within schools where producing primary care generalists is engrained in the institution's mission (Block, Clark-Chiarelli, & Singer, 1998; Zinn et al, 2001; Duerson, Crandall, & Dwyer, 1989).

Acknowledging that the development of these *early* bodies of career knowledge is ambiguous, the results of this study also show that those with a career preference at matriculation—even those with a non-primary care preference—were more likely than those who were undecided to prefer primary care at graduation. This suggests that the hidden curriculum express itself more heavily on those without strong career preferences, and that these students' preferences are more malleable. With this in mind, identifying and articulating one's specialty preference is a necessary but insufficient condition for sparking students' preference for primary care careers. To be certain, providing support structures will help students *pursue* their preference. Thus, if we accept the premise of Hafferty and Franks, that "medical training at root is a process of moral enculturation", then we must also accept that social structures can be developed to combat the impact of the hidden curriculum on students' interest in primary care (1994, p. 861). In my integrative conceptual model, I incorporated the language of Lave and Wenger, because it was clear that learning was truly "situated" within the physical and social ecology of a medical school, just as the hidden curriculum hypothesis recognizes.

The results of this study certainly suggest that students not only modify their career preferences (*Figures 5.1* to *5.3*), but also their motivations for being a physician in the first place. As the final logistic regression models showed, the motivational factors played a *much* smaller role in distinguishing those who were undecided from those who preferred primary care at graduation than at matriculation. Thus, if underlying career motivations shift because of medical school experiences, then the creation of "communities of practice" that support those with primary care interests.

Examples of communities of practice abound. The creation, continuance, or development of a required family practice rotation, which is already in place in medical schools (American Academy of Family Physicians [AAFP], 2015), is supported by the results of this study. The research also suggests that student-faculty interactions play a strong role in determining interest in a primary care career (Senf, Campos-Outcalt, & Kutob, 2003), as faculty are role models and ideally placed to deliver career advice. To that end, the faculty composition of medical schools should be reviewed to ensure that students have multiple points of contact with primary care faculty to better understand their work and receive career advice.

Finally, and perhaps most obviously, the results of this study reinforce the need for explicit primary care community of practice, which appear to be the most sensible and economical social structure to support student interest in primary care careers. The *Center for Primary Care* at Harvard Medical School is one such example of a community of practice that provides information and mentorship opportunities for medical students. (Harvard Medical School, 2015). One wonders if this kind of community of practice would help alleviate doubts of those truly committed to primary care, such as Mara Gordon, who recently wrote in *the Atlantic* that "every time I tell someone I plan to go into primary care, I wonder if they think I wasn't smart enough to do something else" (2014, n.p.) Such perceptions can be powerful, and a community of practice would provide students with confidence as they prepare to enter their preferences into the NRMP's match algorithm. This model could be expanded to other institutions whose goal is increasing the production of primary care physicians.

Recommendations for practice:

Faculty should work particularly with "undecided" students early in their education to determine their motivations, and help them align these motivations with specialties. This is important, as many students do not enter medical school with defined career preferences. Additionally, medical schools with institutional missions related to primary care should highlight for undecided students how the motivational factors Intellectual Intrigue, Common & Control, and Lifestyle Management can be incorporated into primary care careers. The AAMC's *Careers in Medicine* page provides useful tools for this purpose (2015e).

This study, based on socialization theory, supports the implementation of primary care focused communities of practice to resist the erosion of in primary care careers.

Personal Reflection

"Sound decisions are made in the presence of sound data." I want to briefly conclude this dissertation by expanding that statement as a final, personal reflection.

For me, the process of developing this dissertation has felt very much like a writing a technical report, as I hoped to provide the field with a body of evidence not yet scene in the literature, do it with robust methods and a compelling dataset, and provide tools for interpretation. This much, I think I have achieved. My second hope is that this study will also be used to spark efforts to create a stronger undergraduate-to-physician pipeline research and policy agenda. This will, I think, allow us to develop a system where students are informed and make the good life decisions, with the needs of society equally prioritized.

Table 5.1

	Primary Care versus Undecided		Non-Primary Care versus Undecided		
	Matriculation	Graduation	Matriculation	Graduation	
Gender					
Male ^R	R	R	R	R	
Female	+	-	-	-	
Race/Ethnicity					
White ^R	R	R	R	R	
Asian	?	-	-	-	
URM	+	?	+	?	
Other/ Unknown	-	-	?	?	
Parental Occupation					
Other-Parent ^R	R	R	R	R	
Professional-Parent	?	+	-	?	
Physician-Parent	-	?	-	+	
Hospital Work					
No ^R	R	R	R	R	
Yes	+	+	+	+	
College Research					
No ^R	R	R	R	R	
Yes	-	+	?	+	
Motivational Factors					
Professional Benefits	?	?	-	?	
Intellectual Intrigue	-	+	+	+	
Command & Control	-	?	+	+	
Doctor-Patient Relationship	+	?	-	-	
Lifestyle Management	-	?	?	?	
Prior Specialty Pref.					
Undecided ^R	N/A	R	N/A	R	
Primary Care	N/A	+	N/A	-	
Non-Primary Care	N/A	+	N/A	+	

Summary	of the	Multinomial	Logistic	Reg	ression	Results
-			р.		a	

Note. [+] indicates that the variable positively predicts primary or non-primary care preference; [-] indicates that the variable positively predicts an undecided preference; [?] indicates that no relationship could be identified; [R] the variable was either not included in the analysis or was the reference category; [N/A] signifies that the predictor was not included in the model.



Figure 5.1. Flowchart of medical students with a primary care specialty preference at matriculation. NB. Lines show the percentage indicating a preference for primary care, undecided or non-primary care specialties at graduation. (N=7,103)



Figure 5.2. Flowchart of medical students with undecided specialty preferences at matriculation. NB. Lines show the percentage indicating a preference for primary care, undecided or non-primary care specialties at graduation. (N=12,455).



Figure 5.3. Flowchart of medical students with a non-primary care specialty preference at matriculation. NB. Lines show the percentage indicating a preference for primary care, undecided or non-primary care specialties at graduation. (N=11,867)



Figure 5.4. Conceptual chart showing typical outcome specification shown with the solid line, and the outcome specification of this analysis shown with the dashed lines.

REFERENCES

- Abdi, H. (2003). Factor rotations in factor analyses. In Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (Eds), *The Encyclopedia for Research Methods for the Social Sciences* (792-795). Sage: Thousand Oaks, CA.
- Albanese, M. A., Snow, M. H., Skochelak, S. E., Huggett, K. N., & Farrell, P. M. (2003). Assessing personal qualities in medical school admissions. *Academic Medicine*, 78(3), 313-321
- American Academy of Family Physicians [AAFP]. (2015). *Rotations & Electives*. Retrieved from: <u>http://www.aafp.org/medical-school-residency/medical-school/rotations.html</u>
- American Association of Physicians Assistants [AAPA]. (2015). *Re: SB 716 and SB 754, Modifying Provisions Relating to Public Health.* Retrieved from: https://www.aapa.org/WorkArea/DownloadAsset.aspx?id=2876
- American Board of Medicine Specialties [ABMS]. (2014). *Family Medicine*. Retrieved from: <u>http://www.abms.org/member-boards/contact-an-abms-member-board/american-board-of-family-medicine/</u>
- American College of Physicians [ACP]. (2014). Internal medicine subspecialties. Retrieved from: <u>http://www.acponline.org/patients_families/about_internal_medicine/subspecialties/about_internal_medicine/subspec</u>
- American Academy of Family Physicians [AAFP]. (2013). *Transforming Graduate Medical Education*. Retrieved from:

http://www.aafp.org/dam/AAFP/documents/advocacy/workforce/gme/ES-Transforming-GME-100309.pdf

American Medical Association [AMA]. (2014a). *Bill would fund residency slots for underserved areas*. Retrieved from: <u>http://www.ama-assn.org/ama/pub/ama-</u> wire/ama-wire/post/bill-would-fund-residency-slots-underserved-areas

American Medical Association [AMA]. (2014b). *Graduates' career plan. FRIEDA Online*. Retrieved from: <u>https://freida.ama-assn.org/Freida/</u>

Association of American Medical Colleges [AAMC]. (2004). *The status of the new AAMC definition of "underrepresented in medicine" following the Supreme Court's decision in Grutter*. Retrieved from https://www.aamc.org/download/54278/data/urm.pdf

Association of American Medical Colleges [AAMC]. (2006). AAMC Statement on the Physician Workforce, June 2006. Retrieved from https://www.aamc.org/download/55458/data/workforceposition.pdf

Association of American Medical Colleges [AAMC]. (2008). *Diversity in Medical Education Facts and Figures 2008*.. Retrieved from: <u>https://www.aamc.org/download/386172/data/diversityinmedicaleducation-</u> factsandfigures2008.pdf

Association of American Medical Colleges [AAMC]. (2012a). *Diversity in Medical Education: Facts & Figures 2012*. Retrieved from <u>https://members.aamc.org/eweb/upload/Diversity%20in%20Medical%20Educatio</u> <u>n%20Facts%20and%20Figures%202012.pdf</u>

Association of American Medical Colleges [AAMC]. (2012b). Table 8: Applicants to U.S. Medical Schools by Race, Selected Combinations within Hispanic or Latino *Ethnicity, and Sex, 2009-2012.* Retrieved from: https://www.aamc.org/download/321472/data/2012factstable8.pdf

Association of American Medical Colleges [AAMC]. (2013a). *Matriculating Student Questionnaire: 2013 All Schools Summary Report*. Retrieved from: https://www.aamc.org/download/363478/data/msq2013report.pdf

Association of American Medical Colleges [AAMC]. (2013b). *Medical School Graduation Questionnaire, 2013 All School Summary Report.* Retrieved from: <u>https://www.aamc.org/download/350998/data/2013gqallschoolssummaryreport.pd</u> <u>f</u>

Association of American Medical Colleges [AAMC]. (2014a). *Medical School Gradation Questionnaire: 2014 All Schools Summary Report*. Retrieved from: <u>https://www.aamc.org/download/397432/data/2014gqallschoolssummaryreport.pdf</u>

Association of American Medical Colleges [AAMC]. (2014b). *Matriculating Student Questionnaire: 2014 All Schools Summary Report*. Retrieved from: <u>https://www.aamc.org/download/419782/data/msq2014report.pdf</u>

Association of American Medical Colleges [AAMC]. (2014c). *Physician Shortages to Worsen Without Increases in Residency Training*. Retrieved from: https://kaiserhealthnews.files.wordpress.com/2011/01/md-shortage.pdf

Association of American Medical Colleges [AAMC]. (2015a). *Graduation Questionnaire*. Retrieved from: <u>https://www.aamc.org/data/gq/</u>

Association of American Medical Colleges [AAMC]. (2015b). *Post-MCAT Questionnaire (PMQ)*. Retrieved from: <u>https://www.aamc.org/data/pmq</u> Association of American Medical Colleges [AAMC]. (2015c). *Matriculating Student Questionnaire (MSQ)*. Retrieved from: <u>https://www.aamc.org/data/msq</u>

Association of American Medical Colleges [AAMC]. (2015d). *Student Records System* (*SRS*). Retrieved from: https://www.aamc.org/services/srs/

Association of American Medical Colleges [AAMC]. (2015e). *Careers in Medicine*. Retrieved from: <u>https://www.aamc.org/cim/</u>

- Andriole, D. A., Whelan, A. J., & Jeffe, D. B. (2008). Characteristics and career intentions of the emerging MD/PhD workforce. JAMA, 300(10), 1165-1173.
- Austin, A. E. (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. *The Journal of Higher Education*, 73(1), 94-122.
- Austin, A. E., Campa, H., Pfund, C., Gillian-Daniel, D. L., Mathieu, R., & Stoddart, J. (2009). Preparing STEM doctoral students for future faculty careers. *New Directions for Teaching and Learning*, *117*, 83-95.
- Babbott, D., Baldwin, D. C., Jolly, P., & Williams, D. J. (1988). The stability of early specialty preferences among US medical school graduates in 1983. *JAMA*, 259(13), 1970-1975.
- Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall.

- Babaria, P., Bernheim, S., & Nunez-Smith, M. (2011). Gender and the pre-clinical experiences of female medical students: a taxonomy. *Medical education*, 45(3), 249-260.
- Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of statistical psychology*, *3*(2), 77-85.
- Bennett, K. L., & Phillips, J. P. (2010). Finding, recruiting, and sustaining the future primary care physician workforce: a new theoretical model of specialty choice process. *Academic Medicine*, 85(10), S81-S88.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, 88(3), 588.
- Bland, C. J., Meurer, L. N., & Maldonado, G. (1995). Determinants of primary care specialty choice: a non-statistical meta-analysis of the literature. *Academic Medicine*, 70(7), 620-41.
- Block, S. D., Clark-Chiarelli, N., & Singer, J. D. (1998). Mixed messages about primary care in the culture of US medical schools. *Academic Medicine*, 73(10), 1087-1094.
- Boden, D., Borrego, M., & Newswander, L. K. (2011). Student socialization in interdisciplinary doctoral education. *Higher Education*, 62(6), 741-755.
- Bodenheimer, T., Chen, E., & Bennett, H. D. (2009). Confronting the growing burden of chronic disease: can the US health care workforce do the job? *Health Affairs*, 28(1), 64-74.
- Bodenheimer, T., & Pham, H. H. (2010). Primary care: current problems and proposed solutions. *Health Affairs*, 29(5), 799-805.

- Bollen, K. A. (1990). Overall fit in covariance structure models: Two types of sample size effects. *Psychological Bulletin*, *107*(2), 256.
- Brass, L. F., Akabas, M. H., Burnley, L. D., Engman, D. M., Wiley, C. A., & Andersen, O. S. (2010). Are MD-PhD programs meeting their goals? An analysis of career choices made by graduates of 24 MD-PhD programs. *Academic Medicine*, 85(4), 692-701.
- Brim, O. (1966). Socialization through the life-cycle. In O. Brim, and S. Wheeler (Eds.), *Socialization after childhood* (pp. 3-49). NY, NY: Wiley.
- Brotherton, S. E., Rockey, P. H., & Etzel, S. I. (2005). US graduate medical education, 2004-2005: trends in primary care specialties. *JAMA*, *294*(9), 1075-1082.
- Brown, S. D., & Lent, R. W. (1996). A social cognitive framework for career choice counseling. *The Career Development Quarterly*, 44(4), 354-366.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, 154, 136-136.
- Buchmueller, T. C., Grumbach, K., Kronick, R., & Kahn, J. G. (2005). Book review: The effect of health insurance on medical care utilization and implications for insurance expansion: A review of the literature. *Medical Care Research and Review*, 62(1), 3-30.
- Burack, J. H., Irby, D. M., Carline, J. D., Ambrozy, D. M., Ellsbury, K. E., & Stritter, F. T. (1997). A study of medical students' specialty-choice pathways: trying on possible selves. *Academic Medicine*, 72(6), 534-41.

- Campos-Outcalt, D., Senf, J., Watkins, A. J., & Bastacky, S. (1995). The effects of medical school curricula, faculty role models, and biomedical research support on choice of generalist physician careers: a review and quality assessment of the literature. *Academic Medicine*, 70(7), 611-619.
- Carrier, E., Yee T., & Stark, L. (2011). Matching Supply to Demand: Addressing the U.S. Primary Care Workforce Shortage, Policy Analysis No. 7, National Institute for Health Care Reform, Washington, D.C.
- Cassel, C. K., & Reuben, D. B. (2011). Specialization, subspecialization, and subsubspecialization in internal medicine. *New England Journal of Medicine*, 364(12), 1169-1173.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245-276.
- Center for Workforce Studies. (2010). *The Impact of Health Care Reform on the Future Supply and Demand for Physicians Updated Projections through 2025*. Retrieved from <u>https://www.aamc.org/download/158076/data/updated_projections_through_2025</u> <u>.pdf</u>

Center for Workforce Studies. (2012). *Recent Studies and Reports on Physician Shortages in the US*. Retrieved from <u>https://www.aamc.org/download/100598/data/recentworkforcestudies.pdf</u>

Center for Workforce Studies. (2014). *Results of the 2013 Medical School Enrolment Survey*. Retrieved from: <u>https://www.aamc.org/download/431876/data/resultsofthe2013medicalschoolenro</u> <u>llmentreport.pdf</u>

- Chen, C., Petterson, S., Phillips, R. L., Mullan, F., Bazemore, A., & O'Donnell, S. D. (2013). Toward graduate medical education (GME) accountability: Measuring the outcomes of GME institutions. *Academic Medicine*, 88(9), 1267-1280.
- Chen, R. M., Petterson, S. M., & Bazemore, A. W. (2014). Projected impact of the primary care residency expansion program using historical trends in graduate placement. *American family physician*, 89(7), 518-518.
- Click, I. A. (2013). Practice Characteristics of Graduates of East Tennessee State University Quillen College of Medicine: Factors Related to Career Choices in Primary Care (Doctoral Dissertation). Retrieved from: <u>http://dc.etsu.edu/cgi/viewcontent.cgi?article=2275&context=etd</u>
- Clinite, K. L., Reddy, S. T., Kazantsev, S. M., Kogan, J. R., Durning, S. J., Blevins, T., ... & DeZee, K. J. (2013). Primary care, the ROAD less traveled: what first-year medical students want in a specialty. *Academic Medicine*, 88(10), 1522-1528.
- Clinite, K. L., DeZee, K. J., Durning, S. J., Kogan, J. R., Blevins, T., Chou, C. L., ... & Reddy, S. T. (2014). Lifestyle factors and primary care specialty selection:
 Comparing 2012–2013 graduating and matriculating medical students' thoughts on specialty lifestyle. *Academic Medicine*, *89*(11), 1483-1489.
- Cohen J. *Facing the future*. *President's Address*. 112th Annual Meeting of the Association of American Medical Colleges, November 4, 2001, Washington, DC.
- Cole, D., & Espinoza, A. (2011). The postbaccalaureate goals of college women in STEM. *New Directions for Institutional Research*, *152*, 51-58.
- Colquitt, W. L., Zeh, M. C., Killian, C. D., & Cultice, J. M. (1996). Effect of debt on US medical school graduates' preferences for family medicine, general internal medicine, and general pediatrics. *Academic Medicine*, 71(4), 399-411.
- Colwill, J. M., Cultice, J. M., & Kruse, R. L. (2008). Will generalist physician supply meet demands of an increasing and aging population? *Health Affairs*, 27(3), 232-241.
- Compton, M. T., Frank, E., Elon, L., & Carrera, J. (2008). Changes in US medical students' specialty interests over the course of medical school. *Journal of general internal medicine*, 23(7), 1095-1100.
- Cooper, R. A., Getzen, T. E., McKee, H. J., & Laud, P. (2002). Economic and demographic trends signal an impending physician shortage. *Health Affairs*, 21(1), 140-154.
- Council on Graduate Medical Education [COGME]. (1992). *Third Report: Improving Access to Health Care Through Physician Workforce Reform: Directions for the 21st Century*. Retrieved from: <u>http://www.hrsa.gov/advisorycommittees/bhpradvisory/cogme/Reports/thirdreport</u> <u>.pdf</u>
- Council on Graduate Medical Education [COGME]. (2010). Advancing Primary Care. Retrieved from: <u>http://www.hrsa.gov/advisorycommittees/bhpradvisory/cogme/Reports/twentiethr</u> eport.pdf
- Council of Pediatric Subspecialties [CoPS]. (2014). *Descriptions of pediatric subspecialties*. Retrieved from: <u>http://www.pedsubs.org/SubDes/index.cfm</u>
- DeZee, K. J., Maurer, D., Colt, R., Shimeall, W., Mallory, R., Powers, J., & Durning, S. J. (2011). Effect of financial remuneration on specialty choice of fourth-year US medical students. *Academic Medicine*, 86(2), 187-193.

- Dill, M. J., & Salsberg, E. S. (2008). The complexities of physician supply and demand: projections through 2025. Association of American Medical Colleges.
- Dinno, A. (2009). Implementing Horn's parallel analysis for principal component analysis and factor analysis. *Stata Journal*, *9*(2), 291.
- Dohn, H. (1996). Choices of careers in medicine: some theoretical and methodological issues. *Medical Education*, *30*(3), 157-160.
- Dornan, T. (2005). Osler, Flexner, apprenticeship and the new medical education'. *Journal of the Royal Society of Medicine*, 98(3), 91-95.
- Dorsey, E. R., Jarjoura, D., & Rutecki, G. W. (2005). The influence of controllable lifestyle and sex on the specialty choices of graduating US medical students, 1996-2003. Academic Medicine, 80(9), 791-796.
- Duerson, M. C., Crandall, L. A., & Dwyer, J. W. (1989). Impact of a required family medicine clerkship on medical students' attitudes about primary care. *Academic Medicine*, 64(9), 546-548.
- Dunleavy, D. M., Kroopnick, M. H., Dowd, K. W., Searcy, C. A., & Zhao, X. (2013). The predictive validity of the MCAT exam in relation to academic performance through medical school: a national cohort study of 2001–2004 matriculants. Academic Medicine, 88(5), 666-671.
- Eckleberry-Hunt, J., & Tucciarone, J. (2011). The challenges and opportunities of teaching "Generation Y". *Journal of graduate medical education*, *3*(4), 458-461.
- Erikson, C. E., Danish, S., Jones, K. C., Sandberg, S. F., & Carle, A. C. (2013). The role of medical school culture in primary care career choice. *Academic Medicine*, 88(12), 1919-1926.

- Ferguson, E., James, D., & Madeley, L. (2002). Factors associated with success in medical school: systematic review of the literature. Bmj, 324(7343), 952-957.
- Freeman, B. S. (2004). *The ultimate guide to choosing a medical specialty*. Lange Medical Books/McGraw Hill.
- Freeman, J. D., Kadiyala, S., Bell, J. F., & Martin, D. P. (2008). The causal effect of health insurance on utilization and outcomes in adults: a systematic review of US studies. *Medical care*, 46(10), 1023-1032.
- Frintner, M. P., Mulvey, H. J., Pletcher, B. A., & Olson, L. M. (2013). Pediatric resident debt and career intentions. *Pediatrics*, 131(2), 312-318.

Frisch, S. (2013). The primary care physician shortage. BMJ, 347(f6559).

- Funkenstein, D. H. (1978). Medical Students, Medical Schools and Society during Five Eras: Factors Affecting the Career Choices of Physicians 1958-1976. Cambridge, MA: Ballinger Publishing
- Gardner, S. K. (2010). Keeping up with the Joneses: Socialization and culture in doctoral education at one striving institution. *The Journal of Higher Education*, 81(6), 728-749.
- George, D., & Mallery, M. (2010). SPSS for Windows Step by Step: A Simple Guide and *Reference*. Boston, MA: Pearson.
- Gold, J. A., Barg, F. K., & Margo, K. (2014). Undergraduate Students' Perspectives on Primary Care. *Journal of primary care & community health*.

- Goodson, J. D. (2010). Patient Protection and Affordable Care Act: promise and peril for primary care. *Annals of internal medicine*, *152*(11), 742-744.
- Gordon, M. (September 18, 2014). Why I'm Becoming a Primary-Care Doctor. *The Atlantic*. Retrieved from:
 <u>http://www.theatlantic.com/health/archive/2014/09/why-im-becoming-a-primary-care-doctor/379231/</u>
- Gorenflo DW, Ruffin MT, Sheets KJ. (1994). A multivariate model for specialty preference by medical students. *Journal of Family Practice* 9(6):570-577.
- Greenbaum, P. E., & Dedrick, R. F. (1998). Hierarchical confirmatory factor analysis of the Child Behavior Checklist/4–18. Psychological Assessment, 10(2), 149.
- Greenberg, L., & Cultice, J. M. (1997). Forecasting the need for physicians in the United States: the Health Resources and Services Administration's physician requirements model. *Health Services Research*, 31(6), 723.
- Grbic, D., Garrison, G., & Jolly, P. (August, 2010). Analysis in brief: diversity of U.S. medical school students by parental education. Retrieved from <u>https://www.aamc.org/download/142770/data/aibvol9_no10.pdf</u>
- Hafferty F. (2008). Professionalism and the socialization of medical students. In: Creuss
 R., Creuss S., Steinert Y. (Eds.), *Teaching Medical Professionalism*. Cambridge,
 UK: Cambridge University Press.
- Hafferty, F. W. (1998). Beyond curriculum reform: confronting medicine's hidden curriculum. *Academic Medicine*, *73*(4), 403-7.
- Hafferty, F. W., & Franks, R. (1994). The hidden curriculum, ethics teaching, and the structure of medical education. *Academic Medicine*, *69*(11), 861-71.

- Hafler, J. P., Ownby, A. R., Thompson, B. M., Fasser, C. E., Grigsby, K., Haidet, P., ... & Hafferty, F. W. (2011). Decoding the learning environment of medical education:
 A hidden curriculum perspective for faculty development. *Academic Medicine*, 86(4), 440-444.
- Harvard Medical School. (2015). *Center for Primary Care: What We Do*. Retrieved from: <u>https://primarycare.hms.harvard.edu/what-we-do</u>
- Hauer, K. E., Durning, S. J., Kernan, W. N., Fagan, M. J., Mintz, M., O'Sullivan, P. S., ...
 & Schwartz, M. D. (2008). Factors associated with medical students' career choices regarding internal medicine. *JAMA*, *300*(10), 1154-1164.
- Health Resources and Services Administration [HRSA]. (2006). Physician supply and demand: Projections to 2020. HRSA, Bureau of Health Professionals. Retrieved from http://bhpr.hrsa.gov/healthworkforce/supplydemand/medicine/physician2020proje ctions.pdf

Health Resources and Services Administration [HRSA]. (2008). The Physician Workforce: Projections and Research into Current Issues Affecting Supply and Demand. Retrieved from <u>http://bhpr.hrsa.gov/healthworkforce/reports/physwfissues.pdf</u>

Health Resources and Services Administration [HRSA]. (2013). *Projecting the supply and demand for primary care practitioners through 2020*. Retrieved from <u>http://bhpr.hrsa.gov/healthworkforce/supplydemand/usworkforce/primarycare/projectingprimarycare.pdf</u>

Health Resources and Services Administration [HRSA]. (2014). Primary Medical Care HPSA Designation Criteria. Retrieved from http://bhpr.hrsa.gov/shortage/hpsas/designationcriteria/primarycarehpsacriteria.ht ml

- Health Resources and Services Administration [HRSA]. (2015). *HPSA Find*. Retrieved from: http://datawarehouse.hrsa.gov/tools/analyzers/hpsafind.aspx
- Hofer, A. N., Abraham, J., & Moscovice, I. (2011). Expansion of coverage under the Patient Protection and Affordable Care Act and primary care utilization. *Milbank Quarterly*, 89(1), 69-89.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179-185.
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Humes, K., Jones, N. A., & Ramirez, R. R. (2011). Overview of race and Hispanic origin, 2010. US Department of Commerce, Economics and Statistics Administration, US Census Bureau.
- Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models.* Sage.
- Iglehart, J. K. (2013). The residency mismatch. *New England Journal of Medicine*, 369(4), 297-299.

- Institute of Medicine [IOM]. (2008). *Retooling for an Aging America*. Washington, DC: The National Academies Press.
- Jeffe, D. B., Whelan, A. J., & Andriole, D. A. (2010). Primary care specialty choices of United States medical graduates, 1997–2006. Academic Medicine, 85(6), 947-958.
- Jeffe, D. B., Yan, Y., & Andriole, D. A. (2012). Do research activities during college, medical school, and residency mediate racial/ethnic disparities in full-time faculty appointments at US Medical schools? *Academic Medicine:* 87(11), 1582.
- Jeffe, D. B., Andriole, D. A., Wathington, H. D., & Tai, R. H. (2014). The Emerging Physician–Scientist Workforce: Demographic, Experiential, and Attitudinal Predictors of MD–PhD Program Enrollment. *Academic Medicine*, 89(10), 1398-1407.
- Jolly, P. (January, 2008). Analysis in brief: diversity of U.S. medical students by parental income. Retrieved from https://www.aamc.org/download/102338/data/aibvol8no1.pdf
- Jolly, P., Erikson, C., & Garrison, G. (2013). US graduate medical education and physician specialty choice. *Academic Medicine*, 88(4), 468-474.
- Josiah Macy Jr. Foundation. (2011). Ensuring an Effective Physician Workforce for America: Recommendations for an Accountable Graduate Medical Education System. Retrieved from: <u>http://www.macyfoundation.org/docs/macy_pubs/Effective_Physician_Workforce_Conf_Book.pdf</u>
- Julian, E. R. (2005). Validity of the Medical College Admission Test for predicting medical school performance. *Academic Medicine*, 80(10), 910-917.

- Kitsis, E. A., & Goldsammler, M. (2013). Physician shadowing: a review of the literature and proposal for guidelines. *Academic Medicine*, 88(1), 102-110.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3), 187-200.
- Kanter, R. M. (1968). Men and women of the corporation. New York: Basic Books.
- Kassebaum, D. G., & Szenas, P. L. (1994). Factors influencing the specialty choices of 1993 medical school graduates. *Academic Medicine*, 69(2), 163-70.
- Kassebaum, D. G., Szenas, P. L., Ruffin, A. L., & Masters, D. R. (1995). The research career interests of graduating medical students. *Academic Medicine*, 70(9), 847-852.
- Kassler, W. J., Wartman, S. A., & Silliman, R. A. (1991). Why medical students choose primary care careers. *Academic Medicine*, 66(1), 41-3.
- Kleshinski, J., Khuder, S. A., Shapiro, J. I., & Gold, J. P. (2009). Impact of preadmission variables on USMLE step 1 and step 2 performance. *Advances in Health Sciences Education*, 14(1), 69-78.
- Lambert, E. M., & Holmboe, E. S. (2005). The relationship between specialty choice and gender of US medical students, 1990–2003. *Academic Medicine*, *80*(9), 797-802.
- Laurant, M., Reeves, D., Hermens, R., Braspenning, J., Grol, R., & Sibbald, B. (2014). *Substitution of doctors by nurses in primary care*. Retrieved from: <u>http://www.bibliotecacochrane.com/pdf/CD001271.pdf</u>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

- Lawson, S. R., & Hoban, J. D. (2003). Predicting career decisions in primary care medicine: a theoretical analysis. *Journal of Continuing Education in the Health Professions*, 23(2), 68-80.
- Lawson, S. R., Hoban, J. D., & Mazmanian, P. E. (2004). Understanding primary care residency choices: a test of selected variables in the Bland-Meurer model. *Academic Medicine*, 79(10), S36-S39.
- Ledesma, R. D., & Valero-Mora, P. (2007). Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis. *Practical Assessment, Research & Evaluation, 12*(2), 1-11.
- Lempp, H., & Seale, C. (2004). The hidden curriculum in undergraduate medical education: qualitative study of medical students' perceptions of teaching. *BMJ*, 329(7469), 770-773.
- Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of vocational behavior*, 45(1), 79-122.
- Lentz, B. F., & Laband, D. N. (1989). Why so many children of doctors become doctors: Nepotism vs. human capital transfers. *Journal of Human Resources*, 396-413.
- Lewis, J. L., Menzies, H., Najera, E. I., & Page, R. N. (2009). Rethinking trends in minority participation in the sciences. *Science Education*, 93, 961 – 977.
- Linde, C. (1993). *Life stories: The creation of coherence*. New York: Oxford University Press.

- Martini, C. J., Veloski, J. J., Barzansky, B., Xu, G., & Fields, S. K. (1994). Medical school and student characteristics that influence choosing a generalist career. *JAMA*, 272(9), 661-668
- Massachusetts Medical Society [MMS]. (2013). 2013 MMS Patient Access to Care Study. Retrieved from: <u>http://www.massmed.org/News-and-Publications/Research-and-Studies/2013-MMS-Patient-Access-to-Care-Study-%28pdf%29/</u>
- McAdams, D. P. (1993). *The stories we live by: Personal myths and the making of the self.* New York: William Morrow.
- McGaghie, W. C. (1990). Perspectives on medical school admission. Academic Medicine, 65(3), 136-9.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, 1(2), 130.
- Markert, R. J. (1991). Why medical students change to and from primary care as career choice. *Family medicine*, *23*(5), 347-350.
- Marrast, L. M., Zallman, L., Woolhandler, S., Bor, D. H., & McCormick, D. (2014). Minority physicians' role in the care of underserved patients: diversifying the physician workforce may be key in addressing health disparities. *JAMA internal medicine*, 174(2), 289-291.
- Merton, R. K., Reader, G., and Kendall, P. L. (1957). *The student physician*. Cambridge, MA:Harvard University Press.
- Mitchell, W. D. (1975). Medical student career choice: a conceptualization. *Social Science & Medicine*, *9*(11), 641-653.

- Morley, C. P., Mader, E. M., Smilnak, T., Bazemore, A., Petterson, S., Rodríguez, J. E., & Campbell, K. M. (2015). The social mission in medical school mission statements: associations with graduate outcomes. *Family medicine*, 47(6), 427-434.
- Muthén, B. (2004). Latent variable analysis. The Sage handbook of quantitative methodology for the social sciences. Thousand Oaks, CA: Sage Publications, 345-68.
- Muthén, L. & Muthén, B. 2007). *Mplus User's Guide. Fifth Edition*. Los Angeles, CA: Muthén & Muthén.
- Muthén, L. & Muthén, B. (August, 2009). Exploratory Factor Analysis, Confirmatory Factor Analysis, and Structural Equation Modeling For Continuous Outcomes. Lecture conducted at Johns Hopkins University. Baltimore, MD.
- National Resident Matching Program [NRMP]. (2014a). Results and data: 2014 main residency
- Match. Retrieved from <u>http://www.nrmp.org/wp-content/uploads/2014/04/Main-Match-Results-and-Data-2014.pdf</u>

National Resident Matching Program [NRMP]. (2014b). *Charting Outcomes in the Match*. Retrieved from: <u>http://www.nrmp.org/wp-</u> <u>content/uploads/2014/09/Charting-Outcomes-2014-Final.pdf</u>

National Resident Matching Program [NRMP]. (2014c). *First-Year Graduate Medical Education in the United States: 2002-2014*. Retrieved from: <u>http://www.nrmp.org/wp-content/uploads/2014/04/First-Year-GME-in-the-US-rev-for-WWW_4.8.14-FINAL.pdf</u>

- National Resident Matching Program [NRMP]. (2015). *Results and Data: 2015 Main Residency Match*. Retrieved from: <u>http://www.nrmp.org/wp-content/uploads/2015/05/Main-Match-Results-and-Data-2015_final.pdf</u>
- Newton, D. A., Grayson, M. S., & Thompson, L. F. (2005). The variable influence of lifestyle and income on medical students' career specialty choices: data from two US medical schools, 1998–2004. *Academic Medicine*, 80(9), 809-814.
- Nivet, M. A. (2011). Commentary: Diversity 3.0: A necessary systems upgrade. *Academic Medicine*, 86(12), 1487-1489.
- Nowacek, G., & Sachs, L. (1990). Demographic variables in medical school admission. *Academic Medicine*, 65(3), 140-144.
- Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric Theory (3rd Edition)*. New York, New York: McGraw-Hill.
- Office of Missouri Governor Jay Nixon. (2015). *Gov. Nixon signs bill to reauthorize Missouri Rx program.* Retrieved from: <u>http://governor.mo.gov/news/archive/gov-</u> <u>nixon-signs-bill-reauthorize-missouri-rx-program</u>
- O'Neil, E., & Dower, C. (2011). Primary care health workforce in the United States. *The* Synthesis project. Research synthesis report, (22). Retrieved from: <u>http://www.rwjf.org/content/dam/farm/reports/issue_briefs/2011/rwjf402104/suba</u> <u>ssets/rwjf402104_1</u>
- Osipow, S. H. (1990). Convergence in theories of career choice and development: Review and prospect. *Journal of vocational behavior*, *36*(2), 122-131.
- Osler, W. (1893). The principles and practice of medicine. D. Appleton & Company.

- Owen, J. A., Hayden, G. F., & Connors Jr, A. F. (2002). Can medical school admission committee members predict which applicants will choose primary care careers? *Academic Medicine*, 77(4), 344-349.
- Page, K. R., Castillo-Page, L., Poll-Hunter, N., Garrison, G., & Wright, S. M. (2013). Assessing the evolving definition of underrepresented minority and its application in academic medicine. *Academic Medicine*, 88(1), 67-72.

Patient Protection and Affordable Care Act, 42 U.S.C. § 18001 (2010).

Pauly, M., Naylor, M., & Weiner, J. (2014). "Primary Care Shortages: It's More Than Just a Head Count." Retrieved from:

http://www.inqri.org/sites/default/files/INQRI%20Brief%20FINAL1.pdf .

- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. Thompson Learning Inc.
- Petterson, S. M., Liaw, W. R., Phillips, R. L., Rabin, D. L., Meyers, D. S., & Bazemore,
 A. W. (2012). Projecting US primary care physician workforce needs: 2010-2025. *The Annals of Family Medicine*, 10(6), 503-509.
- Petterson, S. M., Phillips Jr, R. L., Bazemore, A. W., & Koinis, G. T. (2013). Unequal distribution of the US primary care workforce. *American family physician*, 87(11).
- Phillips, J., Weismantel, D., Gold, K., & Schwenk, T. (2012). How do medical students view the work life of primary care and specialty physicians? *Family Medicine-Kansas City*, 44(1), 7.

- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta psychologica*, 104(1), 1-15.
- Rabinowitz, H. K. (1999). The role of the medical school admission process in the production of generalist physicians. *Academic Medicine*, 74(Supplement), S39-44.
- Reed, V. A., Jernstedt, G. C., & Reber, E. S. (2001). Understanding and improving medical student specialty choice: a synthesis of the literature using decision theory as a referent. *Teaching and Learning in Medicine*, 13(2), 117-129.
- Reis, H. T., & Judd, C. M. (2000). *Handbook of research methods in social and personality psychology*. Cambridge University Press.
- Reybold, L. E. (2003). Pathways to the professorate: The development of faculty identity in education. *Innovative Higher Education*, 27(4), 235-252.
- Ricketts, T. C. (2011). The health care workforce: will it be ready as the boomers age? A review of how we can know (or not know) the answer. *Annual review of public health*, 32, 417-430.
- Robert Wood Johnson Foundation. (2003). *Generalist Physician Initiative: An RWJF National Program*. Retrieved from <u>http://www.rwjf.org/content/dam/farm/reports/program_results_reports/2003/rwjf</u> <u>69797</u>
- Rogers, M.E., Creed, P. A., & Searle, J. (2009). The Development and Initial Validation of Social Cognitive Career Theory Instruments to Measure Choice of Medical Specialty and Practice Location. *Journal of Career Assessment*, 17, 324-337.

- Rosenblatt, A., & Andrilla, C. (2005). The impact of US medical students' debt on their choice of primary care careers: an analysis of data from the 2002 medical school graduation questionnaire. *Academic Medicine*, 80(9), 815-819.
- Salsberg, E., Rockey, P. H., Rivers, K. L., Brotherton, S. E., & Jackson, G. R. (2008). US residency training before and after the 1997 Balanced Budget Act. *JAMA*, 300(10), 1174-1180.
- Siegel, R. (Host). (2014, November, 10). Secretary of Embattled VA: 'I Can Reset The Tone'. [Radio broadcast episode]. In Sanders, S. (Producer), *All Things Considered*. Washington, DC, National Public Radio.
- Sandmann, L., Saltmarsh, J., & O'Meara, K. (2008). An integrated model for advancing the scholarship of engagement: Creating academic homes for the engaged scholar. *Journal of Higher Education Outreach and Engagement*, 12(1), 47-64.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99(6), 323-338.
- Schwartz, R. W., Simpson, W. G., Strodel, W. E., Jarecky, R. K., Griffen Jr, W. O., & Byron Young, A. (1989). Career change: in quest of a controllable lifestyle. *Journal of Surgical Research*, 47(3), 189-192.
- Schwartz, R. W., Haley, J. V., Williams, C., Jarecky, R. K., Strodel, W. E., Young, B., & Griffen Jr, W. O. (1990). The controllable lifestyle factor and students' attitudes about specialty selection. *Academic Medicine*, 65(3), 207-10.
- Schwartz, M. D., Basco, W. T., Grey, M. R., Elmore, J. G., & Rubenstein, A. (2005). Rekindling student interest in generalist careers. *Annals of internal medicine*, 142(8), 715-724.

- Schwartz, M. D. (2012). Health care reform and the primary care workforce bottleneck. *Journal of general internal medicine*, 27(4), 469-472.
- Scott, I., Gowans, M., Wright, B., Brenneis, F., Banner, S., & Boone, J. (2011). Determinants of choosing a career in family medicine. *Canadian Medical Association Journal*, 183(1), E1-E8.
- Scott, I., Wright, B., Brenneis, F., Brett-MacLean, P., & McCaffrey, L. (2007). Why would I choose a career in family medicine? Reflections of medical students at 3 universities. *Canadian Family Physician*, 53(11), 1956-1957.
- Senf, J. H., Campos-Outcalt, D., Watkins, A. J., Bastacky, S., & Killian, C. (1997). A systematic analysis of how medical school characteristics relate to graduates' choices of primary care specialties. *Academic Medicine*, 72(6), 524-33.
- Senf, J. H., Campos-Outcalt, D., & Kutob, R. (2003). Factors related to the choice of family medicine: a reassessment and literature review. *The Journal of the American Board of Family Practice*, 16(6), 502-512.
- Sharp, L. M., and Frankel, J. (1983). Respondent Burden: A Test of Some Common Assumptions. *Public Opinion Quarterly*, 47(1), 36–53.
- Sinclair, S. (1997). Making doctors: an institutional apprenticeship. Berg Publishers.
- Siu, E., & Reiter, H. I. (2009). Overview: what's worked and what hasn't as a guide towards predictive admissions tool development. Advances in Health Sciences Education, 14(5), 759-775.

- Stage, F. K. (1990). LISREL: An introduction and applications in higher education. In J.
 C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 427–466). New York: Agathon Press.
- Starfield, B., Shi, L., & Macinko, J. (2005). Contribution of primary care to health systems and health. *Milbank quarterly*, *83*(3), 457-502.
- Stevens, J. P. (2012). Applied multivariate statistics for the social sciences. Routledge.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.)*. Upper Saddle River, NJ: Pearson Allyn & Bacon.
- Tanaka, J. (1993). Multifaceted conceptions of fit in structural equation models. In Bollen & Long (Eds.), *Testing structural equation models*. (pp. 10-39) Newbury Park, CA: Sage.
- Thomson, G. H. (1951). *The Factorial Analysis of Human Ability*. London: University of London Press.
- Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association.
- Ullman, J. B. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (Eds.), Using multivariate statistics (4th ed.). Needham Heights, MA: Allyn & Bacon.
- U.S. Census Bureau. (2012). Table 2. Projections of the Population by Selected Age Groups and Sex for the United States: 2015 to 2060 (NP2012-T2). Washington, DC: U.S. Census Bureau.

- U.S. Congress, Office of Technology Assessment. (1988). *Forecasts of Physician Supply and Requirements*. Washington, DC: Government Printing Office.
- Van Prooijen, J. W., & Van Der Kloot, W. A. (2001). Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement*, 61(5), 777-792.
- Veterans Affairs [VA]. (2014a). Veterans Access, Choice and Accountability Act of 2014. Retrieved from: <u>http://www.va.gov/opa/choiceact/documents/Fact-Sheet-</u> <u>Residency-Program.pdf</u>
- Veterans Affairs [VA]. (2014b). Summary, Veterans Access, Choice and Accountability Act of 2014 ("Choice Act"). Retrieved from: <u>http://www.va.gov/opa/choiceact/documents/Choice-Act-Summary.pdf</u>.
- Voorhees, K. I., Prado-Gutierrez, A., Epperly, T., & Derksen, D. (2013). A proposal for reform of the structure and financing of primary care graduate medical education. *Fam Med*, 45(3), 164-170.
- Weidman, J. C., Twale, D. J., & Stein, E. L. (2001). Socialization of Graduate and Professional Students in Higher Education: A Perilous Passage? ASHE-ERIC
 Higher Education Report, Volume 28, Number 3. Jossey-Bass Higher and Adult
 Education Series. San Francisco, CA: Jossey-Bass, Publishers, Inc.
- Weiner, J. P. (1994). Forecasting the effects of health reform on US physician workforce requirement: evidence from HMO staffing patterns. *Jama*, 272(3), 222-230.
- Wendel, T. M., Godellas, C. V., & Prinz, R. A. (2003). Are there gender differences in choosing a surgical career?. Surgery, 134(4), 591-596.
- Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. Cambridge, UK: Cambridge University Press.

- Wetzel, A. (2012). Factor analysis methods and validity evidence: a review of instrument development across the medical education continuum. *Academic Medicine*, 87(8), 1060-1069
- Xierali, I., Nivet, M., & Fair, M. (2014). Analysis in brief: Analyzing Physician Workforce Racial and Ethnic Composition Associations: Physician Specialties (Part I). Retrieved from: <u>https://www.aamc.org/download/401798/data/aug2014aibpart1.pdf</u>
- Yuan, K. H., & Bentler, P. M. (1998). Normal theory based test statistics in structural equation modelling. *British Journal of Mathematical and Statistical Psychology*, 51(2), 289-309.
- Zinn, W. M., Sullivan, A. M., Zotov, N., Peters, A. S., Connelly, M. T., Singer, J. D., & Block, S. D. (2001). The effect of medical education on primary care orientation: results of two national surveys of students' and residents' perspectives. *Academic Medicine*, 76(4), 355-365.

APPENDIX A:

STATISTICAL SYNTAX

```
*
*
*
/*
Dissertation Syntax (Revised and Annotated Version)
Started: 01, September 2014
Last Updated: 03, July 2015
Author: Daniel M Read
Affiliation: University of Virginia,
               Curry School of Education
                  CISE Department
                     Robert H. Tai Laboratory
Summary of research questions: How can primary-care-committed, primary-care-positive,
undecided, and non-primary-care-committed students be identified in the
admissions process or early in medical school?
*/
*
*
*
*
*
*Loading the dataset; inclusion and exclusion criteria*
*Opens the dataset:
      use
"C:\Users\daniel_read\Desktop\PERSONAL\TREMUR_DIS\Data_DoFiles\tremur_data_working_file_2
0130128.dta", clear
      *Begins a log file:
      log using
"C:\Users\daniel_read\Desktop\PERSONAL\TREMUR_DIS\Data_DoFiles\Dissertation_Exploration_v
_17_May", replace
      *Data Screening:
      *Confirm N:
      count
      *Criteria #1: No Duplicates:
      drop if R_ID=="262D84E99F25D09E" | R_ID=="CEEC1B3755EDE3D7"
             count
             *2 deletions
```

```
*Criteria #2: Must have matriculated into a medical school
drop if MAT AYEAR==" "
        count.
        *162,700 deletions
*Criteria #3: Must have taken MCAT before matriculating:
drop if MAT AYEAR=="1995-1996"
drop if MAT_AYEAR=="1996-1997"
drop if MAT_AYEAR=="1997-1998"
drop if MAT_AYEAR=="1998-1999"
drop if MAT_AYEAR=="1999-2000"
drop if MAT AYEAR=="2000-2001"
drop if MAT AYEAR=="2001-2002" & MCAT ADMIN YEAR=="2002"
drop if MAT_AYEAR=="2001-2002" & MCAT_ADMIN_YEAR=="2003"
drop if MAT_AYEAR=="2002-2003" & MCAT_ADMIN_YEAR=="2003"
drop if MAT_AYEAR=="2001-2002" & MCAT_ADMIN_YEAR=="2004"
drop if MAT AYEAR=="2002-2003" & MCAT ADMIN YEAR=="2004"
drop if MAT AYEAR=="2001-2002" & MCAT ADMIN YEAR=="2006"
        count
        *27 deletions
*And, did not matriculate into an MDPHD program:
drop if MD PHD IND=="Y"
        count
        *3948 deletions
*Criteria #4: Must have graduated from medical school
drop if GRAD_AYEAR==" "
        count
        *10,282 deletions
                *... and in the appropriate time range (2005-2006 to 2011-2012)
                tab GRAD AYEAR
                drop if GRAD_AYEAR=="2003-2004"
                drop if GRAD AYEAR=="2004-2005"
                tab GRAD_AYEAR
                count
                *59 deletions
        *
         *
         *
        *
         *At this point we have the graduating population of interest; N=85,797
        *
         *
        *
         *
         *
         *
        *
*Additionaly Exclusion: Must have completed the GQ & the MSQ:
generate GQ_completer=0
replace GQ_completer=1 if GQ_year_r>=1
```

```
tab GQ completer
```

```
generate MSQ_completer=0
replace MSQ completer=1 if MSQ YEAR!=.
tab MSQ_completer
tab GQ completer MSQ completer, cell
       drop if GQ_completer==0
       *32400 deletions
       drop if MSQ\_completer==0
       *14,020 deletions
       count
       *
        *
       *
        *
        *
        *
       *
        * Found that 33,120 had not completed the GQ in the timeframe;
       *
        * Found that 26,636 had not completed the MSQ in the timeframe;
       *
        *
       *
        *
        *
```

*Additionaly Exclusion: Must have all attitudinal factors

egen MISSING_FACTORS=rowmiss(FAC_AUTHORTY FAC_AUTONOMY FAC_CONTACT FAC_CONTROL FAC_DECISION FAC_DEXTER FAC_EVALUATE FAC_EXPERT FAC_INCOME FAC_INNOVATE FAC_INTELL_CHALL FAC_INTERACT FAC_LEADER FAC_MAKE_DIFF FAC_PROMOTE FAC_PSY_CARE FAC_RESEARCH FAC_SECURE FAC_SOCIAL FAC_STATUS FAC_STRESS FAC_TECHNOLOGY)

tab MISSING FACTORS

gen MISSING_FACTORS_CAT=0
replace MISSING_FACTORS_CAT=1 if MISSING_FACTORS>=1
tab MISSING_FACTORS_CAT MSQ_YEAR, column

drop if MSQ_YEAR==2007 *7721 deletions

*Tabout: ssc install tabout *Example: tabout FAC_PROMOTE FAC_RESEARCH using "C:\Users\RHT Research\Desktop\TREMUR_DIS\Data_DoFiles\temporary1.txt", replace

> *Horns Parallel Analysis: ssc install paran

```
*Lengthy algorythm
       *Bartlett's test and KMO:
       ssc install factortest
       *Anova Contrasts
       ssc install anovacontrast
*Data Exploration and Predictor Creation*
*****
*1. Gender
       tab SRS SEX
       generate student sex=0
       replace student_sex=1 if SRS_SEX=="F"
       tab student sex
       *No missing data for gender.
       label define sex_LBL 1 "Female" 0 "Male"
       label values student sex sex LBL
*2. Race/Ethnicity
       tab PMQ_SRS_race_ethnicity_r
       tabulate PMQ SRS race ethnicity r, gen (student race)
       generate student_race_4cat=4
       replace student_race_4cat=2 if PMQ_SRS_race_ethnicity_r==4
replace student_race_4cat=3 if PMQ_SRS_race_ethnicity_r==2 |
PMQ SRS race ethnicity r==3 | PMQ SRS race ethnicity r==5 | PMQ SRS race ethnicity r==6
       replace student_race_4cat=1 if PMQ_SRS_race_ethnicity_r==1
       tabulate student_race_4cat, gen (student_race_4cat)
       *Labels
       label define URM LBL 1 "White" 2 "Asian" 3 "URM" 4 "Other"
       label values student race 4cat URM LBL
       tab student race 4cat
       tabulate student_race_4cat, gen (student_race_4cat)
       ren student race 4cat1 White
       ren student_race_4cat2 Asian
       ren student_race_4cat3 URM
ren student_race_4cat4 Other
       *No missing data for race/ethnicity using the 8 category variable.
*3. MCAT score
       summarize first mcat total, detail
       *About 30 students with missing data.
*4. Parent Occupation from PMQ*
       /* Codes for the PMQ Parental Occupation Question*
       01 'Physician (M.D., M.D./Ph.D. or D.O.) [Health Professional]'
       02 'Dentist [Health Professional] 03 Veterinarian [Health Professional]'
       04 'Optometrist [Health Professional] 05 Podiatrist [Health Professional]'
       06 'Pharmacist [Health Professional] 07 Registered Nurse [Health Professional]'
       08 'Health Technologist [Health Professional] 09 Other licensed health
professional [Health Professional] '
       10 'Other health worker [Health Professional] 11 Architect [Business and Other
Professional] '
       12 'Clergy [Business and Other Professional] 13 Engineer [Business and Other
Professional]'
       14 'Government Official [Business and Other Professional]'
       15 'Lawyer or Judge [Business and Other Professional]'
       16 'Mathematician or Computer Scientist [Business and Other Professi'
       17 'Natural Scientist [Business and Other Professional]'
```

18 'Social Scientist [Business and Other Professional]' 19 'Teacher - College or University [Business and Other Professional' 20 'Teacher - other than coll or univ [Business and Other Professio' 21 'Owner, Executive, Manager - corp bus [Business and Other Profess' 22 'Owner/Manager - small business [Business and Other Professional]' 23 'Other technician, semi-professional [Business and Other Professi' 24 'Other business or professional [Business and Other Professional]' 25 '(Secretary) Administrative Assistant [Clerical/Sales]' 26 'Other Clerical worker [Clerical/Sales] 27 Sales - supervisor [Clerical/Sales]' 28 'Sales Representative [Clerical/Sales] 29 Sales worker, retail/personal services [Clerical/Sales]' 30 'Enlisted Armed Services [Service Occupation/Trades]' 31 'Service (police, fire, etc.) [Service Occupation/Trades]' 32 'Skilled worker or craftsman [Service Occupation/Trades] 33 'Transport or Equipment worker [Service Occupation/Trades]' 34 'Unskilled worker/laborer, household [Service Occupation/Trades]' 35 'Farmer, farm manager, farm foreman [Agriculture]' 36 'Farm laborer [Agriculture] 37 Homemaker [Other]' 39 'Medical Student (added 1991) [Other] 40 Full time student in other he alth field (added 1991) [Other]' 41 'Full time student in non-health field (added 1991) [Other]' 42 'Part time student in other health field (added 1991) [Other]' 43 'Part time student in non-health field (added 1991) [Other]' 44 'Other (added 1991) [Other]' 45 'Unemployed (added 1991; dropped 1999) [Other]' 46 'Do not know (dropped from questionnaire in 1999)' 47 'Best guess is ACT used for missing'. We will use the already reduced codes from D&D "... occ r4"*/ *Father: tab father occ r4, nolabel tab father occ r4 *Low and Missing generate father occupation=1 *Professional or Healthcare Worker replace father occupation=2 if father occ r4==2 replace father occupation=2 if father occ r4==3 *Physician replace father occupation=3 if father occ r4==1 *Mother tab mother_occ_r4, nolabel tab mother occ r4 *Low and Missing generate mother occupation=1 *Professional or Healthcare Worker replace mother occupation=2 if mother occ r4==2 replace mother occupation=2 if mother occ r4==3 *Physician replace mother occupation=3 if mother occ r4==1 *Combined generate highest_parent_occupation=3 replace highest_parent_occupation=2 if father_occupation==2 & mother_occupation==2 replace highest parent occupation=2 if father occupation==2 & mother occupation==1 replace highest parent occupation=2 if father occupation==1 & mother occupation==2 replace highest_parent_occupation=1 if father_occupation==1 & mother occupation==1 tab mother occupation father occupation tab mother_occupation highest_parent_occupation tab father occupation highest parent occupation

*Labels label define highest_parent_occupationLBL 1 "Other/Missing" 2 "Professional/Healthcare" 3 "Physician" label values highest_parent_occupation highest_parent_occupationLBL tab highest parent occupation tabulate highest parent occupation, gen (parent occ dummy) ren parent occ dummyl Other Parent ren parent_occ_dummy2 Prof_Parent ren parent occ dummy3 Phys Parent *5. Hospital Work (PMQ) tab X_HOSPWORK xk label drop X HOSPWORK xk ren X HOSPWORK xk Hospital *6. Lab Expereince: "Indicate any programs you participated in to prepare for a career in medicine or science. (Laboratory research apprenticeship for college students)" tab PART COLL LAB r, missing tab PART_COLL_LAB_r, missing nolabel
replace PART_COLL_LAB_r=0 if PART_COLL_LAB_r!=1 label drop PART COLL LAB r ren PART_COLL_LAB_r College_research ***TABLES*** *Data screening: tab GQ completer MSQ_completer, cell tab student sex GQ completer, r tab PMQ SRS race ethnicity r GQ completer, r *3.1 tabulate PMQ SRS race ethnicity r student race 4cat *3.3 tab father occupation highest parent occupation *3.4 tab Hospital College research, cell *Outcome Variable Creation *PMQ Data: describe NEW SPEC PREF destring NEW_SPEC_PREF, replace tab NEW SPEC PREF, missing *Primary Care Groups gen pmq_internal dummy=0 replace pmq_internal_dummy=1 if NEW_SPEC PREF==200 gen pmq_family_dummy=0 ________replace_pmq_family_dummy=1 if NEW_SPEC_PREF==180 gen pmq peds dummy=0 replace pmq peds dummy=1 if NEW SPEC PREF==500 gen pmq_pc_dummy=0 replace pmq pc dummy=1 if pmq internal dummy==1 | pmq family dummy==1 | pmq peds dummy==1 tab NEW SPEC PREF pmq pc dummy, missing

*This creates a variable for whether they chose INT/PEDS/FAM at matriculation, as well as one dummy variable combining ANY of these choices.

```
*Non-PC and Undecided Groups:
       gen pmq_undecided dummy=0
               replace pmq undecided dummy=1 if NEW SPEC PREF==.
       gen pmq nonpc dummy=0
               replace pmq nonpc dummy=1 if pmq pc dummy==0 & pmq undecided dummy==0
       *Confirms that all respondents belong to only one group:
       egen pmq_speccheck= rowtotal(pmq_nonpc_dummy pmq_undecided dummy pmq pc dummy)
       tab pmq speccheck
       drop pmq speccheck
       *Creates a single variable for group analyses
       generate pre group 3cat=1
               replace pre group 3cat=2 if pmq undecided dummy==1
               replace pre_group_3cat=3 if pmq_nonpc_dummy==1
       *Labels
       label define group 3catLBL 1 "PC" 2 "Undecided" 3 "Non-PC"
       label values pre group 3cat group 3catLBL
       tab pre_group_3cat
       *5 group category for matriculation:
       generate pre_group_5cat=.
               replace pre_group_5cat=1 if pmq_internal_dummy==1
               replace pre_group_5cat=2 if pmq_family_dummy==1
               replace pre_group_5cat=3 if pmq_peds_dummy==1
replace pre_group_5cat=4 if pmq_undecided_dummy==1
               replace pre group 5cat=5 if pmq nonpc dummy==1
       *Labels
       label define group 5catLBL 1 "Internal" 2 "Family" 3 "Peds" 4 "Undecided" 5 "Non-
       label values pre group 5cat group 5catLBL
       tab pre group 5cat
*MSQ Data:
       tab MSQ SPEC PLAN r
       tab MSQ SPEC PREF
       tab MSQ SPEC PREF MSQ SPEC PLAN r, missing
       *Primary Care Groups
       gen msq_internal_dummy=0
              replace msq internal dummy=1 if MSQ SPEC PREF==200
       gen msq_family_dummy=0
               replace msq family dummy=1 if MSQ SPEC PREF==180
       gen msq_peds dummy=0
               gen msq_pc_dummy=0
               replace msq pc dummy=1 if msq internal dummy==1 | msq family dummy==1 |
msq peds dummy==1
               tab MSQ SPEC PREF msq pc dummy, missing
```

*This creates a variable for whether they chose INT/PEDS/FAM at matriculation, as well as one dummy variable combining ANY of these choices.

*Non-PC and Undecided Groups: gen msq_undecided_dummy=0 replace msq undecided dummy=1 if MSQ SPEC PREF==.

PC"

```
gen msq nonpc dummy=0
               replace msq nonpc dummy=1 if msq pc dummy==0 & msq undecided dummy==0
        *Confirms that all respondents belong to only one group:
        list msq nonpc dummy msq undecided dummy msq pc dummy in 1/30
        egen msq speccheck= rowtotal (msq nonpc dummy msq undecided dummy msq pc dummy)
        tab msq speccheck
        *Creates a single variable for group analyses
       generate mat group 3cat=1
               replace mat group 3cat=2 if msq undecided dummy==1
               replace mat group_3cat=3 if msq_nonpc_dummy==1
       tabulate mat_group_3cat, gen (mat_group_3cat_dummy)
       ren mat_group_3cat_dummy1 Mat_PC
ren mat_group_3cat_dummy2 Mat_Und
       ren mat group 3cat dummy3 Mat NPC
       *Labels
       label values mat_group_3cat group_3catLBL
       tab mat group 3cat
       *5 group category for matriculation:
       generate mat group 5cat=.
               replace mat_group_5cat=1 if msq_internal_dummy==1
               replace mat_group_5cat=2 if msq_family_dummy==1
replace mat_group_5cat=3 if msq_peds_dummy==1
               replace mat group 5cat=4 if msq undecided dummy==1
               replace mat group 5cat=5 if msq nonpc dummy==1
        *Labels
       label values mat group 5cat group 5catLBL
        tab mat group 5cat
*GQ Data:
        tab GQ overall spec pref r
       tab GQ SPEC PLAN r2
       gen gq_internal dummy=0
               replace gq_internal_dummy=1 if GQ_overall spec pref r==140
        gen gq family dummy=0
               replace gq_family_dummy=1 if GQ_overall spec pref r==120
       gen gq_peds_dummy=0
               replace gq_peds_dummy=1 if GQ_overall spec pref r==320
       gen gq pc dummy=0
               replace gq_pc_dummy=1 if gq_internal_dummy==1 | gq_family_dummy==1 |
gq peds dummy==1
        *This creates a variable for whether they chose INT/PEDS/FAM at graduation,
               *as well as one dummy variable combining ANY of these choices.
        *Non-PC and Undecided Groups:
       gen gq undecided dummy=0
               replace gq undecided dummy=1 if GQ SPEC PLAN r2==2 |GQ SPEC PLAN r2==3 |
GQ SPEC PLAN r2==.
       gen gq nonpc dummy=0
               replace gq nonpc dummy=1 if gq pc dummy==0 & gq undecided dummy==0
        *Confirms that all respondents belong to only one group
       list gq nonpc dummy gq undecided dummy gq pc dummy in 1/30
       egen gq speccheck= rowtotal(gq nonpc dummy gq undecided dummy gq pc dummy)
```

```
tab gq speccheck
       *Notice that there are 9 that are cross coded between PC and Undecided.
       *Creates a single variable for group analyses (3 groups)
       generate grad group 3cat=1
               replace grad_group_3cat=2 if gq_undecided_dummy==1
               replace grad group 3cat=3 if gq nonpc dummy==1
       *Labels:
       label values grad group 3cat group 3catLBL
       tab grad group 3cat
       *5 group category for matriculation:
       generate grad group 5cat=.
               replace grad group 5cat=1 if gq internal dummy==1
               replace grad_group_5cat=2 if gq_family_dummy==1
               replace grad group 5cat=3 if gq peds dummy==1
               replace grad_group_5cat=4 if gq_undecided_dummy==1
replace grad_group_5cat=5 if gq_nonpc_dummy==1
       *Labels
       label values grad group 5cat group 5catLBL
       tab grad group 5cat
*RQ#1 EFA/CFA Modelling*
* Variables for the factor model:
               describe FAC AUTHORTY FAC AUTONOMY FAC CONTACT FAC CONTROL FAC DECISION
FAC DEXTER FAC EVALUATE FAC EXPERT FAC INCOME FAC INNOVATE FAC INTELL CHALL FAC INTERACT
FAC_LEADER FAC_MAKE_DIFF FAC_PROMOTE FAC_PSY_CARE FAC_RESEARCH FAC_SECURE FAC SOCIAL
FAC STATUS FAC STRESS FAC TECHNOLOGY
               * Creates a variable cluster for all 22 variables:
local factor_variablelist_A "FAC_MAKE_DIFF FAC_AUTHORTY FAC_AUTONOMY
FAC_CONTACT FAC_CONTROL FAC_DECISION FAC_DEXTER FAC_EVALUATE FAC_EXPERT FAC_INCOME
FAC INNOVATE FAC INTELL CHALL FAC INTERACT FAC LEADER FAC PROMOTE FAC PSY CARE
FAC RESEARCH FAC SECURE FAC SOCIAL FAC STATUS FAC STRESS FAC TECHNOLOGY"
               *Univeriate Analysis:
               tabstat `factor variablelist A', stat(mean, n, sd, sk , k)
                              tabstat `factor_variablelist_A', stat(n)
tabstat `factor_variablelist_A', stat(sd)
                              tabstat `factor_variablelist_A', stat(sk)
                              tabstat `factor variablelist A', stat(k)
                              swilk `factor_variablelist_A'
               *Creates a variable cluster without 'make difference'
               local factor_variablelist_B "FAC_AUTHORTY FAC_AUTONOMY FAC_CONTACT
FAC CONTROL FAC DECISION FAC DEXTER FAC EVALUATE FAC EXPERT FAC INCOME FAC INNOVATE
FAC INTELL CHALL FAC INTERACT FAC LEADER FAC PROMOTE FAC PSY CARE FAC RESEARCH FAC SECURE
FAC SOCIAL FAC STATUS FAC STRESS FAC TECHNOLOGY"
       *Sets the sample into two randomly divided groups of equal size.
       restore
       drop random EFA_group
       set seed 12324
       generate random = runiform()
```

* *

```
summarize random
       sort random
       generate EFA_group = _n \leq 15712
              ttest `factor variablelist A', by(EFA group)
              tabulate EFA group student race 4cat, chi2
              tabulate EFA group student sex, chi2
              tabulate EFA_group highest_parent_occupation, chi2
              tabulate EFA group Hospital, chi2
              tabulate EFA group College research, chi2
              tabulate EFA group mat group 3cat, chi2
              tabulate EFA group grad group 3cat, chi2
       *Eliminates the CFA group temporarily:
       preserve
       drop if EFA group==0
       count
       *
       *
       *
       *RQ#1 EFA (SECTION 1/3)
       *Multivariate Analysis:
       pwcorr `factor variablelist B', sig obs
       alpha `factor_variablelist_B', detail
       factor `factor_variablelist_B'
factor `factor_variablelist_B', pcf
             estat kmo
              factortest `factor_variablelist_B'
       /*Notes:
       1. Used Jeffe et al. (2010) analysis to look at crosstabulations and most show a
consistent trend;
       2. Lots of pair-wise and inter-item correlations above .5;
       3. KMO overall ~.85 and minimum KMO item was .76;
       4. Bartlet significant at the >.001 level. */
       * Principal components factorization to run the diagnostic tests in STATA:
       pca `factor_variablelist_B', mineigen(1) blanks(.35)
              * Scree plot:
              *scree, ci(hetero reps(1000) seed(18228)) mean yline(1)
              * Horn's Parallel Analysis:
              *paran `factor variablelist B', all graph quietly seed(1)
              /*Notes: We find that we can continue with as many as five factors; ie
there are five factors with eigenvalues greater than one. ^{\star/}
              estat loadings
              predict pc1 pc2 pc3 pc4 pc5
       * PCF factorization; 1st Iteration:
       factor `factor_variablelist_B', pcf mineigen(1.00)
              *Orthoganal rotations:
              rotate, ortho varimax blanks(.45) norm
              rotate, ortho quartim blanks(.45) norm
              rotate, ortho equamax blanks(.45) norm
```

rotate, ortho parsimax blanks(.45) norm

*Oblique rotations rotate, oblique promax blanks(.45) norm rotate, oblique oblimax blanks(.45) norm rotate, oblique quartimin blanks(.45) norm estat common

*Creates a variable cluster without 'make difference' 'autonomy' 'interact' and 'innovate'

local factor_variablelist_C "FAC_AUTHORTY FAC_CONTACT FAC_CONTROL FAC_DECISION FAC_DEXTER FAC_EVALUATE FAC_EXPERT FAC_INCOME FAC_INTELL_CHALL FAC_LEADER FAC_PROMOTE FAC_PSY_CARE FAC_RESEARCH FAC_SECURE FAC_SOCIAL FAC_STATUS FAC_STRESS FAC_TECHNOLOGY"

* PCF factorization; 2nd Iteration:

*Diagnostic Evaluation using the PCA command: pca `factor variablelist C', mineigen(1) blanks(.35) * Scree plot: scree, ci(hetero reps(1000) seed(18228)) mean yline(1) * Horn's Parallel Analysis: paran `factor_variablelist_C', all graph quietly seed(1) factor `factor variablelist C', pcf mineigen(1.00) *Orthoganal rotations: rotate, ortho varimax blanks(.45) norm rotate, ortho quartim blanks(.45) norm rotate, ortho equamax blanks(.45) norm rotate, ortho parsimax blanks(.45) norm *Oblique rotations rotate, oblique promax blanks(.45) norm rotate, oblique oblimax blanks(.45) norm rotate, oblique quartimin blanks(.45) norm estat common * Continue with the oblique promax rotation. *Reliability for all factors: alpha FAC STATUS FAC INCOME FAC AUTHORTY FAC SECURE, item alpha FAC_STRESS FAC_CONTROL, item alpha FAC_DEXTER FAC_LEADER FAC_DECISION, item alpha FAC_EXPERT FAC_TECHNOLOGY FAC_EVALUATE FAC_RESEARCH FAC INTELL CHALL, item alpha FAC CONTACT FAC SOCIAL FAC PSY CARE FAC_PROMOTE, item /* Notes: five factor solution is preferable, all are internally reliable (>=.70*/* Generates the factors using the regression method for the preffered solution: factor `factor variablelist C', pcf mineigen(1.00) rotate, oblique quartimin normalize blanks(.45) predict factor1 factor2 factor3 factor4 factor5 *Renames the factors: ren factor1 efa Prestige ren factor2 efa_Intellectual ren factor3 efa Command and Control ren factor4 efa_Altruism ren factor5 efa Lifestyle *RQ#1 CFA (SECTION 2/3) *****

```
restore
       preserve
       drop if EFA_group==1
        *Brings BACK the EFA group and N is back to 31,425.
       restore
               *
        *
        ^{\ast} I the export the reduced (N=30,425) and full (N=31,425) data files into MPLUS*
                       * The reduced data file is used only to confirm the EFA factor
structure,
                              * The full data file is used to create the factor scores
for all respondents.
                                      *I then bring in the full data set with the factors
back into STATA.
                                              * Syntax Below.
/*
DATA:
    FILE IS "E:/EXPORT TO MPLUS.csv";
    FORMAT IS FREE;
VARIABLE:
    NAMES ARE AUTHORTY AUTONOMY CONTACT CONTROL DECISION DEXTER
    EVALUATE EXPERT INCOME INNOVATE
    INTELL INTERACT LEADER MAKE PROMOTE PSY CARE
    RESEARCH SECURE SOCIAL STATUS STRESS TECH R ID;
    IDVARIABLE IS R ID;
    USEVARIABLES ARE AUTHORTY CONTACT CONTROL DECISION DEXTER
    EVALUATE EXPERT INCOME
    INTELL LEADER PROMOTE PSY CARE
    RESEARCH SECURE SOCIAL STATUS STRESS TECH;
ANALYSIS:
    ESTIMATOR= MLR;
MODEL:
    pres BY SECURE@1 STATUS INCOME AUTHORTY;
    life BY STRESS@1 CONTROL;
    deci BY DEXTER@1 LEADER DECISION;
    rese BY EXPERT@1 TECH EVALUATE RESEARCH INTELL;
    altr BY CONTACT@1 SOCIAL PSY CARE PROMOTE;
    pres WITH life;
    pres WITH deci;
    pres WITH rese;
    pres WITH altr;
    life WITH deci;
    life WITH rese;
    life WITH altr;
    deci WITH rese;
    deci WITH altr;
    rese WITH altr;
    DEXTER WITH DECISION;
    INCOME WITH SECURE;
```

```
INCOME WITH STATUS;
    AUTHORTY WITH SECURE;
    EXPERT WITH TECH;
    EVALUATE WITH INTELL;
OUTPUT:
   Standardized Modindices;
Savedata:
    file is "E:/factor scores.txt";
    save = fscores;
* /
        ^{\star} I then use a SQL editor to bring the factor scores for all respondents back into
an accptable file format.
                * I then bring the full datatset back into STATA.
        *Use file MODEL 13A OBLIQE JAN09
       sem (fac1@1 -> FAC STRESS, ) (fac1 -> FAC CONTROL, ) (fac2@1 -> FAC DEXTER, )
(fac2 -> FAC_LEADER, ) (fac2 -> FAC_DECISION, ) (fac3@1 -> FAC STATUS, ) (fac3 ->
FAC INCOME, ) (fac3 -> FAC AUTHORTY, ) (fac3 -> FAC SECURE, ) (fac5 -> FAC SOCIAL, )
(fac5 -> FAC PROMOTE, ) (fac5@1 -> FAC CONTACT, ) (fac5 -> FAC PSY CARE, ) (fac4@1 -> FAC_EXPERT, ) (fac4 -> FAC_TECHNOLOGY, ) (fac4 -> FAC_EVALUATE, ) (fac4 -> FAC_RESEARCH,
) (fac4 -> FAC INTELL CHALL, ), covstruct( lexogenous, diagonal) latent(fac1 fac2 fac3
fac5 fac4 ) cov( e.FAC_DEXTER*e.FAC_DECISION fac1@1 fac1*fac2 fac1*fac3 fac1*fac5
fac1*fac4 e.FAC_STATUS*e.FAC_INCOME e.FAC_INCOME*e.FAC_SECURE e.FAC_AUTHORTY*e.FAC_SECURE
e.FAC_EXPERT*e.FAC_TECHNOLOGY e.FAC_EVALUATE*e.FAC_INTELL_CHALL fac2*fac3 fac2*fac5
fac2*fac4 fac3*fac5 fac3*fac4 fac4*fac5) nocapslatent
       estat gof, stats(all)
       estat mindices
        *Create factor scores for the CFA model:
       predict cfa Lifestyle cfa Command and Control cfa Prestige cfa Altruism
cfa Intellectual, latent
        *
        *
        *
        *
        *RQ#1 ANOVA (SECTION 3/3)
        *****
       *Univariate Analysis:
       local motivational factors "cfa Lifestyle cfa Prestige cfa Command and Control
cfa Intellectual cfa Altruism"
       tabstat `motivational factors ', stat(mean, n, sd, sk , k, range)
                               tabstat `motivational_factors', stat(n)
tabstat `motivational_factors', stat(sd)
tabstat `motivational_factors', stat(sk)
                               tabstat `motivational_factors', stat(k)
       egen cfa Lifestyle z= std(cfa Lifestyle)
        *For the third part of RQ#1, I create anova tests between the groups
               bysort mat group 3cat: tabstat `motivational factors ', stat(mean, sd)
        *ANOVA tests (matriculation)
               anova cfa Prestige mat group 3cat
```

anovacontrast mat group 3cat, values(0 1 -1) anovacontrast mat_group_3cat, values(-1 1 0) contrast rb2.mat group 3cat anova cfa_Intellectual mat_group_3cat contrast rb2.mat group 3cat anova cfa Command and Control mat group 3cat contrast rb2.mat group 3cat anova cfa Altruism mat group 3cat contrast rb2.mat_group_3cat anova cfa_Lifestyle mat_group_3cat contrast rb2.mat group 3cat *Notes: All tests were significant at the .01 level.* *RQ#2: Multinomial Logistic Regression* *M1, Baseline: mlogit mat_group_3cat, baseoutcome (2) rrr *M2, Baseline + Personal Characteristics: mlogit mat group 3cat student sex Asian URM Other, baseoutcome (2) rrr *M3, Baseline + Personal Characteristics + Socialization: mlogit mat group 3cat student sex Asian URM Other Prof Parent Phys Parent Hospital College research, baseoutcome (2) rrr *M4, Baseline + Personal Characteristics + Socialization + Motivational Factors: mlogit mat group 3cat student sex Asian URM Other Prof Parent Phys Parent Hospital College research cfa Prestige cfa Intellectual cfa Command and Control cfa Altruism cfa Lifestyle, baseoutcome(2) rrr *RQ#3: Multinomial Logistic Regression* *M1, Baseline: mlogit grad group 3cat, baseoutcome (2) rrr *M2, Baseline + Personal Characteristics: mlogit grad group 3cat student sex Asian URM Other, baseoutcome (2) rrr *M3, Baseline + Personal Characteristics + Socialization: mlogit grad group 3cat student sex Asian URM Other Prof Parent Phys Parent Hospital College research, baseoutcome (2) rrr *M4, Baseline + Personal Characteristics + Socialization + Motivational Factors: mlogit grad_group_3cat student_sex Asian URM Other Prof_Parent Phys_Parent Hospital College research cfa Prestige cfa Intellectual cfa Command and Control cfa Altruism cfa Lifestyle, baseoutcome(2) rrr *M5, Baseline + Personal Characteristics + Socialization + Motivational Factors + Prior Spec. Pref: mlogit grad group 3cat student sex Asian URM Other Prof Parent Phys Parent Hospital College_research cfa_Prestige cfa_Intellectual cfa_Command_and_Control cfa Altruism cfa Lifestyle Mat PC Mat NPC, baseoutcome(2) rrr

```
*4.1/4.2
      tab student sex
             tab student race 4cat
                    tab highest_parent_occupation
                           tab Hospital
                                  tab College_research
      *4.1
      tab student_sex mat_group_3cat, column
             tab student_race_4cat mat_group_3cat, column
                    tab highest parent occupation mat group 3cat, column
                           tab Hospital mat group 3cat, column
                                  tab College_research mat_group_3cat, column
      *4.1
      tab student sex mat group 5cat, column
             tab Hospital mat group 5cat, column
                                  tab College_research mat_group_5cat, column
      *4.2
      tab student sex grad group 3cat, column
             tab student_race_4cat grad_group_3cat, column
                    tab highest parent occupation grad group 3cat, column
                           tab Hospital grad group 3cat, column
                                  tab College_research grad_group_3cat, column
      *4.2
      tab student sex grad group 5cat, column
             tab student_race_4cat grad_group_5cat, column
                    tab highest_parent_occupation grad_group_5cat, column
                           tab Hospital grad group 5cat, column
                                  tab College_research grad_group_5cat, column
      *4.3
      tabstat `factor variablelist A', stat(mean, n, sd, sk , k)
      *4.4
      *See .log file for the split sample data
      *4.5
      *See Mplus output.
      *4.6
      *See Mplus output.
      *4.7
      *See Mplus output.
      *4.8
      bysort mat_group_3cat: tabstat `motivational factors ', stat(mean, sd)
      *4.9
      *See the ANOVA section above.
      *4.10
      *See the final model 'M4' in RQ#2
      *4.11
      *See the final model 'M5' in RQ#3
      log close
 *
```

*