6.1.3. Datensatzbereinigung und Umgang mit fehlenden Werten

Eine erste Sichtung der Häufigkeitsverteilungen innerhalb der verwertbaren Datensätze zeigte eine als gering einzuschätzende Quote fehlender Werte innerhalb der einzelnen Fälle (*item non-response*) von unter fünf Prozent bei der überwiegenden Zahl der Indikatoren. Nur bei fünf Indikatoren[141] sind deutlich höhere Ausfallquoten zwischen zehn und 16 Prozent der Fälle beobachtbar. Beim Umgang mit fehlenden Werten wurde auf ein modifiziertes *Hot Deck* Verfahren zurückgegriffen. Die Funktionsweise der hier verwendeten Spielart einer *Hot Deck* Imputation ist denkbar einfach: Fehlende Werte werden durch Werte eines „Spenderfalls" aus dem Datensatz ersetzt, welcher zufällig aus jenen Fällen ausgewählt wurde, die dem „Empfänger" auf einer Reihe vorgegebener Variablen[142] ähneln (vgl. grundsätzlich Sande 1983; Andridge/Little 2010 und speziell Myers 2011). Damit hat jeder Spenderfall die gleiche Chance, seine Werte einem unvollständigen Fall zu *borgen*.

Die Verwendung dieses im Vergleich der zur Verfügung stehenden Imputationsmethoden[143] verhältnismäßig simplen Ansatzes kann mit seiner Einfachheit, der Struktur der fehlenden Daten, den Eigenschaften der erfassten Variablen sowie den angestrebten Analysemethoden begründet werden. So zeigt der *Little's MCAR-Test*, dass die fehlenden Werte im Datensatz keinem systematischen Muster entsprechen, also die MCAR-Annahme[144] Bestätigung findet (Chi-

141 Diese erhöhten Werte können nicht als zufällig angesehen werden, da sie geballt in spezifischen Konstrukten auftreten. Diese scheinen für einen Teil der Befragten schwer oder ungern zu schätzen gewesen zu sein. Es sind die beiden Indikatoren des Konstrukts *Reputation*: „Wir werden regelmäßig von relevanten Politikern als wirtschaftsfeindlich kritisiert", „Die verantwortlichen Politiker stehen hinter uns", die unmittelbare Einschätzung der „Wirtschaftlichkeit" sowie die beiden Indikatoren der *Rekrutierung* „Die Verwaltungseinheit sucht neue Kollegen selbst aus" und „Neue Kollegen haben eine adäquate Qualifikation".

142 Als diese sog. *Deck*-Variablen wurden hier *Bundesland, Behördentyp, Arbeitsbereich, Reformopfer* und *Ausbildungshintergrund* gewählt.

143 Zum Umgang mit fehlenden Werten bieten sich verschiedene Vorgehensweisen an. Die anerkanntermaßen schlechteste (vgl. Schnell 1986: 214ff.; King et al. 1998; Schafer/Graham 2002: 155-157) ist der in den Sozialwissenschaften übliche *listenweise Fallausschluss* – das schlichte Eliminieren aller unvollständigen Fälle. Hierbei wird einerseits die für eine Analyse zur Verfügung stehende Fallzahl oft stark reduziert, andererseits auch eine Verzerrung aufgrund möglicher systematischer Andersartigkeit der ausgeschlossenen Fälle in die Auswertung getragen. Modernere Verfahren, die sich nicht mit der Ersetzung, sondern der modellbasierten Schätzung befassen sind bspw. der *Expectation Maximization Ansatz* oder Formen der *Multiplen Imputation* (vgl. Schafer/Graham 2002; Acock 2005).

144 In der Literatur wird üblicherweise zwischen drei Ausfallmustern unterschieden: So können Daten (1) völlig zufällig, (2) zufällig oder (3) nicht-zufällig fehlen. Die Erfüllung der ersten, als *missing completly at random* (MCAR) bezeichneten Annahme ist aus statisti-

Quadrat = 12002,3; df = 11769; Sig. = .065). Diese Feststellung ist die Voraussetzung für die verzerrungsarme Imputation fehlender Werte (Schafer/Graham 2002: 154). Darüber hinaus folgt die Hot-Deck Imputationen keinem parametrischen Modell und ist robust gegenüber Verletzungen der Normalverteilungsannahme. Auch ist es nicht auf metrische Daten[145] angewiesen. Beide Vorteile ermöglichen seine Verwendung im vorliegenden Fall – was zu zeigen sein wird. Schließlich können mit einem derart ergänzten Datensatz unmittelbar alle beliebigen Analysen vorgenommen werden, er ist nicht durch die begrenzte Verfügbarkeit von Analysetools beschränkt, wie dies bei komplexen Formen multipler Imputation der Fall ist. Ob seiner Einfachheit unterliegt das Hot Deck Verfahren auch deutlichen Limitationen (vgl. Myers 2011: 302).[146]

6.2 Analyseschritte und Methodik

Vor jeder Analyse wurden die Erfüllung der jeweils grundlegenden Annahmen hinsichtlich der Datenstruktur (Normalverteilungsannahme, Linearitätsannahme, Multikollinearität, Homoskedastizität und Autokorrelation) überprüft. Ergebnisse und ggf. vorliegende Einschränkungen werden im Kontext der Analysen berichtet. Mittels des statistischen Vergleichs der vier Modelle soll untersucht werden, ob eventuell zu beobachtende Unterschiede lediglich zufälliger oder systematischer Natur sind. Konkret werden hierzu die folgenden Analysen[147] durchgeführt:

scher Sicht am wünschenswertesten. Sie beinhaltet, dass der Prozess, welcher das Fehlen der Daten verursachte, in keinem statistischen Zusammenhang mit den in der Untersuchung interessierenden Variablen steht (Collins et al. 2001: 333). Die zweite, als *missing at random* (MAR) bezeichnete Annahme besagt, dass zwar ein Zusammenhang zwischen dem Fehlen von Daten und den in der Untersuchung interessierenden Variablen besteht, dieser jedoch über weitere in die Analyse einbezogene Kontrollvariablen erklärt werden kann (Acock 2005: 1014f.). Im Dritten, als *missing not at random* (MNAR) bezeichneten Fall liegt ein systematischer Ausfallmechanismus vor, der jedoch auch auf unbeobachteten Daten beruht. Dieses Ausfallmuster ist nicht durch Kontrollvariablen kontrollierbar und darf nicht ignoriert werden, die Ausfälle werden deshalb auch als *non-ignorable missing values* (NI) bezeichnet (Acock 2005: 1013, 1015).

145 Der vorliegende Datensatz umfasst mit lediglich zwei Ausnahmen bei den Kontrollvariablen ausschließlich latente kategoriale Daten und bietet damit keinen Ansatzpunkt für einen *Expectation Maximiziation* Ansatz.

146 So können einmalige Fälle, für die kein äquivalent im Datensatz vorliegt, nicht passgenau ersetzt werden, was insbesondere bei kleineren Datensätzen ein Problem darstellen kann. Auch führen Vervollständigungen mit einem Hot Deck Verfahren zu Verzerrungen von Korrelationen und anderen Beziehungsmaßen (vgl. Schafer/Graham 2002: 159).

147 Die statistischen Analysen wurden mit den Programmen *IBM SPSS Statistics 20* sowie *STATA 11* durchgeführt.