Table 1: Constructing the BUNMD from NARA Numident Records

Thurgood Marshall

	ssn	fname	lname	birth date							sex	race		bpl		
Application Entry 1	131074264	THURGOOD	MARSHALL	7/2/1908							1	**2**		**MD**		
	ssn	**fname**	**lname**	**birth date**			**death date**				**sex**				**zip_residence**	
Death Entry	**131074264**	**THURGOOD**	**MARSHALL**	**7/2/1908**			**1/24/1993**				**1**				**220411335**	
	ssn	**fname**	**lname**	**byear**	**bmonth**	**bday**	**dyear**	**dmonth**	**dday**	**death_age**	**sex**	**race_first**	**race_last**	**bpl**	**zip_residence**	**number_apps**
BUNMD Entry	**131074264**	**THURGOOD**	**MARSHALL**	**1908**	**7**	**2**	**1993**	**1**	**24**	**84***	**1**	**2**	**2**	**###**	**220411335**	**1***

Lana Turner

	ssn	fname	lname	birth date							race	sex		bpl		
Application Entry 1	567183907	LANA	TURNER	2/8/1921							**1**	2		ID		
Application Entry 2	567183907	LANA	TOPPING	2/8/1921							1	2		**ID**		
Application Entry 3	567183907	LANA	BARKER	2/8/1921							**1**	2		–		
Application Entry 4	567183907	LANA	DANTE	2/8/1921							–	2		–		
	ssn	**fname**	**lname**	**birth date**			**death death**					**sex**			**zip_residence**	
Death Entry	**567183907**	**LANA**	**TURNER**	**2/8/1921**			**6/29/1995**					**2**			**900255240**	
	ssn	**fname**	**lname**	**byear**	**bmonth**	**bday**	**dyear**	**dmonth**	**dday**	**death_age**	**sex**	**race_first**	**race_last**	**bpl**	**zip_residence**	**number_apps**
BUNMD Entry	**567183907**	**LANA**	**TURNER**	**1921**	**2**	**8**	**1995**	**6**	**29**	**74***	**2**	**1**	**1**	**###**	**900255240**	**4***

Note: Bolded values were selected for in the BUNMD. Starred values represent contructed variables not in the original records. Various features of the BUNMD creation algorithm can been seen here. For example, we select a person's first and last name from their death entries. We select the race and birthplace (bpl) from the application records. We use a crosswalk to recode the original two-letter character birthplace codes into a numeric code schema. We select race information from the application files to construct the race_first and race_last variables. The death_age and number_apps variables aren't included in the original records, but were constructed post-hoc using information in the original records.