Summary of ILRR replication results, see tables and memo in the replication tables folder for further details.

	Table 2. Pre-migration characteristics in Norway in 1900.		Table 3. Fathers' characteristics in Norway in 1865.		Table 4. Occupation-based earnings in US, controlling for pre-migration occupations.	
Main result(s):	Return migrants earned less on average than permanent migrants even before moving.	No significant earnings gap within urban or rural areas.	Fathers of return migrants more likely to live in rural areas	Fathers of return migrants more likely to be owner-occupier farmers	Return migrants had lower earnings than permanent migrants	Controlling for own occupation in 1900 reduces earnings gap
Original census data:						
Updated name cleaning & age restriction (standard ABE matching)	✓	✓	✓	✓	✓	✓
Newly available data:						
I. ABE/ Ferrie Algorithm						
Standard	✓	✓	✓	✓	✓	✓
Conservative	✓	✓	✓	✓	✓	✓
Conservative w/ exact name	✓	✓	✓	✓	✓	✓
Conservative w/ exact age					✓	✓
II. ABE w/ JW adjustment						
Unique by exact age	✓	✓	✓	✓	✓	✓
5 yr uniqueness band	✓	✓	✓	✓	✓	✓
5 yr unique band, exact age	✓	✓	✓	✓	✓	✓
III. EM Algorithm						✓
p = 0.10; l = 0.05	✓	✓	✓	✓	✓	✓
p= 0.70; l = 0.65	✓	✓	✓	✓	✓	✓
p=0.70; l = 0.65, exact age	✓	✓	✓	✓	✓	✓

	Table 6. Occupation-based earnings, return migrants vs. non-movers		Table 7. Occupation & location in Norway 1910	Table 8. Top occupations in 1910 of men who were farm laborers in 1900	Table 9. Marital status
Main result(s):	Longer-term migrants experienced a larger earnings penalty.	w/ pre-migration controls, earnings penalty eliminated for short term migrants, but a small negative penalty remains for long-term migrants	Return migrants more likely to be rural and owner occupied farmers w/in municp. of birth	Return migrants more likely to be farmers, less likely to be farm laborers compared with those who stayed in birth municipality	Return migrants more likely to be married before & after migration
Original census data:					
w/ updated name cleaning & age restriction (Standard ABE/Ferrie)	✓	Small short-term penalty remains	✓	✓	✓
Newly available data:					
I. ABE/ Ferrie Algorithm					
Standard	✓	Small short-term penalty remains	✓	✓	✓
Conservative	✓	Small short-term penalty remains	✓	✓	✓
Conservative w/ exact names	✓	Small short-term penalty, no long-term penalty	✓	✓	✓
Conservative w/ exact age	✓	✓	✓	✓	✓
II. ABE w/ JW adjustment					
Unique by exact age	Size of penalty not sig. different	Small short-term penalty remains	✓	✓	✓
5 yr uniqueness band	Penalty larger for short-term migrants	Small short-term penalty remains	✓	✓	✓
5 yr unique band, exact age	Penalty larger for short-term migrants	Small short-term penalty remains	✓	Return migrants more likely to be farm laborers	✓
III. EM Algorithm					
p = 0.10; l = 0.05	Size of penalty not sig. different	Small short-term penalty remains	✓	✓	✓
p= 0.70; l = 0.65	Penalty larger for short-term migrants	Small short-term penalty, no long-term penalty	✓	✓	✓
p=0.70; l = 0.65, exact age	Penalty larger for short-term migrants	Penalty disappears for both groups	✓	✓	✓

Memo on ILRR replication results

Summary

In the original ILRR paper we used the standard ABE/Ferrie matching algorithm with NYSIIS standardized names to link men in the 1910 US Census and the 1910 Norway Census to the 1865 and 1900 Norwegian censuses. We refer to the data linked between 1910 and 1900 as the "early adulthood sample" and the data linked between 1910 and 1865 as the "childhood sample." Since our original analysis, a more detailed version of the US 1910 full-count census has been made available through the NBER. Additionally, several "second generation" matching methods have been developed. We first replicate our original analysis using the newly-available 1910 US Census and find that all main results replicate well. We then present results using several alternative versions of the ABE/Ferrie algorithm such as restricting names to be unique within ± 2 years of age, using exact names, and using the Jaro-Winkler name adjustment. We additionally present results that use only pairs with the same reported age and a set of results using the EM matching algorithm. For more details on each of these linking algorithms see the "matching methods" document. We find that our results generally replicate well when using these alternative matching methods. The replication steps and results are discussed below, and for a visual summary of the replication results see the replication summary document.

Replication steps and results

Since the original paper we have improved upon the way we clean names prior to matching, and now use a standardized cleaning method[1]. This change allows us to match slightly more people, but does not significantly change any of our findings. We additionally correct for a small error in our age restriction which allowed for our regression to include men ages 26 to 47 rather than ages 28 to 45 as reported. Fixing this age restriction results in a slightly smaller matched sample size. Additionally, since the original paper was published the NBER has made available a more detailed version of the 1910 US census that includes the variable occ1950 and allows us to more easily assign occupation-based income. We present our replication results using this newly available data.

We additionally replicate all results using various alternative matching algorithms, which are detailed in the "matching methods" document. The original results concerning the pre-migration characteristics of return migrants and permanent migrants, presented in Tables 2-4, are robust to all matching methods. In Table 2 we continue to find that return migrants earned less than return migrants prior to moving, but that this gap is non-significant when comparing migrants within urban or rural areas. In Table 3 we use various versions of the 1865 to 1910 matched sample and continue to find that fathers of return migrants were more likely to live in rural areas and work as owner-occupier farmers within their municipality of birth. In Table 4 we estimate the earnings gap between permanent and return migrants. When using the newly available census data and making the above-mentioned changes we find the earnings gap to be approximately 50 percent smaller when using the standard ABE/Ferrie algorithm. However, the size of this earnings gap doubles when using exact names in the ABE/Ferrie algorithm and when using the Jaro-Winkler and EM algorithm. In all cases the estimated earnings gap is lower when controlling for pre-migration characteristics, and thus our original conclusion that there were some differences by migration status even prior to immigrating to the US remains unchanged.

---

[1] The code for this standardized cleaning process is available in the codes provided to implement the basic matching approach at https://people.stanford.edu/ranabr/matching-codes

The results presented in Table 6 show that among 28-45 year old men in the 1910 Norway census, those who spent some time in the US had lower occupation-based earnings. In the original paper there was an earnings penalty of approximately 4 percent for return migrants who spent 1 to 3 years in the United States, and a larger penalty of 8.6 percent for those who stayed in the US for 4 to 9 years. We continue to find that longer-term migrants experienced a larger earnings penalty when using alternative versions of the ABE/Ferrie algorithm with standardized and exact names. However, when using the ABE/Ferrie algorithm with JW adjustment and the EM algorithm we find the earnings penalty for short term migrants is not significantly less than the penalty for longer-term migrants, and in some cases the earnings penalty is larger for short-term migrants. Another main takeaway from the original results presented in Table 6 was that once controlling for pre-migration characteristics short-term migrants did not experience a significant earnings gap compared with non-movers, while those who spent more than 4 years in the US earned 6 percent less. The finding that longer-term migrants experienced an earnings penalty even once controlling for pre-migration characteristics is replicated well with the alternative matching methods, however in this replication we also find that a small earnings penalty remains for short-term migrants even after controlling for pre-migration characteristics.

The original results reported in Table 7 found that return migrants were much less likely to live in urban areas and more likely to work as owner-occupied farmers, particularly within the migrant's own municipality of birth. These results are robust to all alternative matching methods. Similarly, Table 8 found that return migrants who worked as farm laborers in 1900 were more likely to work as owner-occupier farmers in 1910 compared with non-migrants and less likely to be farm laborers compared with non-migrants who remained in their municipality of birth. Both finding replicate well when using alternative matching methods. Table 9 compares the marriage status of return migrants and non-movers. Using all alternative matching methods we continue to find that men who spent some time in the US were less likely to be married both before migration in 1900, and after return in 1910.