Discussion Paper

Modelling mobility trends - update including 2021 ODiN data and Covid effects

Harm Jan Boonstra (Statistics Netherlands)
Jan van den Brakel (Statistics Netherlands)
Hans Wüst (KiM Netherlands Institute for Transport Policy Analysis)

October 17, 2022

Abstract

This work is carried out by Statistics Netherlands in collaboration with the Netherlands Institute for Transport Policy Analysis (KiM)/Rijkswaterstaat as an extension to the trend series projects carried out in 2018-2021, in which time series multilevel models have been developed for estimating mobility trends. In the current extension, new data from the Dutch Travel Survey (DTS) over 2021 are added to the series, providing a fourth year of data under the new ODiN design of the DTS. This means there is more data on which estimates of the level break coefficients associated with the latest redesign can be based. However, the mobility in 2020 and 2021 was strongly influenced by the Covid-19 pandemic. For that reason, the model also includes Covid effects. The model has been updated to allow for different Covid effects for 2021 compared to 2020.

The model has been extended to distinguish five instead of four travel purposes. In particular, trend series for purpose leisure are now estimated separately. At the same time estimation of trends for young children, age class 0 - 5, has been discontinued, as this group is no longer observed under ODiN the design. We describe the updated models for the two target variables considered: the number of trip legs per person per day and the distance travelled per trip leg. A few other small improvements are also addressed, including the incorporation of a bias correction in the input estimates instead of in the back-transformations after model estimation, removing a few small but noisy wiggles in the trend estimates.

The models are specified in a hierarchical Bayesian framework and estimated using a Markov Chain Monte Carlo simulation method. From the model outputs trend estimates can be computed at various aggregation levels for the mean number of trip legs per person per day and the mean distance traveled per trip leg, as well as for derived quantities such as the mean distance per person per day. Starting from this year, the trend series are benchmarked towards the measurement level of the ODiN design. We motivate this change by means of a revision analysis, and discuss its implications.

1 Introduction

The Dutch Travel Survey (DTS) is a long-standing annual survey on mobility of residents of the Netherlands. It is carried out by Statistics Netherlands (CBS) and important users of the data are, among others, Rijkswaterstaat and The Netherlands Institute for Transport Policy Analysis (KiM, Kennisinstituut voor Mobiliteitsbeleid), both part of the Ministry of Infrastructure and Water Management.

Since 1985, the DTS survey has undergone several redesigns. The redesigns in 1999, 2004, 2010 and 2018 have caused major discontinuities in the time series of mobility estimates. In 2004 the design actually remained largely unchanged, but its implementation was transferred to another agency, causing several changes in the observed series. For brevity, however, we will mostly also refer to this transition as a 'redesign'.

For users of mobility estimates the changes due to redesigns are very inconvenient as they hamper the temporal comparability. For the redesign of 1999 direct information was available on the sizes of the discontinuities, based on a parallel conducted pilot study. This has been used to correct the series of estimates prior to 1999 to the level of

the estimates under the new design. For the redesigns of 2004, 2010 and 2018 such parallel studies have not been carried out, so in order to estimate the discontinuities a time series model is needed, see van den Brakel et al. (2020). The time series models developed in the current trend estimation project aim to account for the discontinuities due to the redesigns, such that reliable series of trend estimates are obtained with good comparability over time.

Another issue addressed in the trend estimation project is the fact that estimates are desired for a breakdown into many domains, meaning that for each domain determined by both person characteristics (sex and age class) and trip characteristics (purpose and transportation mode) reliable time series of estimates are to be produced. So not only the discontinuities in each of these series should be accounted for, but in addition the amount of data directly relevant to an estimation domain in a specific year is often so small that direct estimates are very noisy and unreliable. The time series of such direct estimates display a lot of volatility caused by the large variances. The time series models developed are able to reduce the noise and yield model estimates that are more precise than the direct estimates by 'borrowing strength' over time as well as over multiple domains. Here the 'borrowing of strength' over domains is brought about by using multilevel time series models with random effects for several levels defining the domains. Within the field of official statistics, the framework of using models to improve on the accuracy of direct estimates for domains of interest is known as small area estimation, see Rao and Molina (2015) for an overview.

The overall purpose of the mobility trends project has been described by the initiators of the project (KiM, Rijkswaterstaat, CBS) as 'Development of a statistical methodology that can derive reliable trend estimates from OVG-MON-OViN-ODiN sample data for the most prevalent mobility data and that deals in a robust way with discontinuities due to redesigns of the survey process and sample noise.' Here OVG, MON, OViN and ODiN refer to the various names used for the DTS during periods with different survey designs. To achieve the purpose as described, time series multilevel models are employed to fit the input data, consisting of direct estimates and estimated standard errors compiled from the DTS survey data. The resulting trend estimates are used by KiM for example in their publication 'Mobiliteitsbeeld' containing actual figures, trends, and expectations about mobility in the Netherlands. The trend estimates are also published on Statistics Netherlands' publication database StatLine, along with the regular annual output based on the DTS. Starting from this year the trend estimates are corrected to the ODiN design measurement level, where previously all estimates were corrected to the OViN design measurement level. This means that many trend series of number of trip legs and distance per trip leg are now estimated slightly higer than before. One plausible reason for these changes is that under the OViN design there was more under-reporting of trips. Another reason is that under ODiN the definition shifted to also include domestic holiday mobility.

The two target variables that are modeled using time series multilevel models are:

- number of trip legs per person per day (pppd)
- distance per trip leg (in hectometers)

A trip for a certain purpose may consist of several trip legs characterized by different transportation modes. Estimates are computed for domains defined by a cross-classification of some or all of the following classification variables:

- sex (male, female)
- age class (6-11, 12-17, 18-24, 25-29, 30-39, 40-49, 50-59, 60-64, 65-69, 70+)

- purpose (”Work”, ”Shopping”, ”Education”, ”Leisure”, ”Other”)
- mode (”Car driver”, ”Car passenger”, ”Train”, ”BTM (bus/tram/metro)”, ”Cycling”, ”Walking”, ”Other”)

Starting from this year, age group 0-5 is no longer taken into account, as this group is not observed anymore under the ODiN design, which started in 2018. Up until last year we still included this age group, but extrapolating trends for this age group becomes more and more uncertain, and it was therefore decided to drop this group. This also means that aggregate estimates now always refer to the population of people aged 6 and over. Another change that has been made since last year is that the level of detail regarding the purpose of travel has been expanded. The category ’Leisure’, which includes family and social visits and recreation, was previously part of the remainder category ’Other’, but is now treated as a separate category. Note that the remaining category ’Other’ carries the same name but is now more narrowly defined than before.

The time series multilevel models are defined at the most detailed level, corresponding to the full cross-classification of sex, age class, purpose and mode, giving rise to $2 \times 10 \times 5 \times 7 = 700$ estimates for a particular year, although some of them such as car-driving or working for young children are structurally zero.

As observed in an earlier project phase, modeling all level break effects as fixed effects generally results in overestimated discontinuities (Bollineni-Balabay et al., 2017). To reduce the risk of overestimated discontinuities and overfitting in general, we model many effects including discontinuities associated with the design transitions as random effects instead. In particular, a regularization method that employs non-normally distributed random effects is used to suppress noisy model coefficients and at the same time allow large effects sufficiently supported by the data. Outliers in the input direct estimates are also modelled, either by adopting a sampling distribution with broader than normal tails or by modelling them explicitly as additional random effects, which are subsequently removed from the trend estimates.

In Boonstra et al. (2019, 2021) and Boonstra et al. (2021) earlier results on model development and mobility trend estimates based on the DTS are described. In this report we update these results after including the latest DTS data over 2021. Besides, the datasets of the first three ODiN years have been slightly updated in the meantime, to correct some information on distances of certain professional travel, but this should hardly affect the trend estimates as professional travel is largely filtered out for that purpose.

The effects that the Covid-19 pandemic has had on mobility in 2020 turned out to be too sudden and too large for the models developed before 2021 to accommodate them. Therefore in Boonstra et al. (2021) the models have been extended by including Covid effects for the year 2020. Omission of such effects would adversely affect the estimation of smooth trend model components and level breaks corresponding to the ODiN redesign. The inclusion of the new ODiN 2021 data made it clear that Covid had slightly different (generally smaller) effects in 2021, making it necessary to differentiate between Covid 2020 and 2021 effects in the model.

Several other small improvements have been made. By adding fixed Covid effects for purposes ’Shopping’ and ’Leisure’ in combination with mode ’Walking’, a better separation of covid effects is achieved. These effects were too large to be picked up by the model’s random covid effects. This way it is possible to obtain more plausible covid-corrected trends. Another improvement is that a bias correction, necessary because of the non-linear data transformations used, is now incorporated in the input

estimates instead of in the back-transformation after model estimation. This has the advantage that a small amount of noise previously introduced by the bias correction is now filtered out by the model. A further model extension was needed because of the separation of purpose 'Leisure' from purpose 'Other', which revealed some noticeable discontinuities in the input series in 2004, when the MON design period started. To account for these discontinuities fixed effects for MON breaks for purposes 'Shopping', 'Leisure' and 'Other' have been added to the model.

The remainder of this report is organized as follows. Section 2 describes the data sources used including a brief overview of the different redesigns the DTS has undergone. In Section 3 the computation of direct estimates and variance estimates from the DTS survey data is discussed, along with transformations of direct estimates and the Generalized Variance Function approach for smoothing the variance estimates, which both improve model fitting. Section 4 describes the hierarchical Bayesian time series multilevel modeling framework. The (updated) models for trip legs and distance are presented in Section 5. Section 6 provides a discussion of the trend estimates based on the estimated models. Section 7 describes a revision analysis that has been carried out in order to explore how stable trend estimates are when they are corrected towards the ODiN level as compared to the OViN level. The paper concludes with a discussion in Section 8. The appendices contain figures of the estimated trend series based on the models developed.

2 Data sources

The DTS is an annual survey that attempts to measure the travel behaviour of the Dutch population. Each year, a sample is drawn with sampling units being defined either as households (before 2010), or persons (since 2010). The variables of interest considered in this study are the number of trip legs and the distance traveled. Direct estimates for these quantities can be obtained using the survey weights that are computed for each year's response data. The survey weights reduce the bias due to non-response, and the estimates based on them correspond to the general regression (GREG) estimator (see e.g. Särndal et al. (1992)).

The DTS started in 1978, and originally was known under the (Dutch) name Onderzoek Verplaatsingsgedrag (OVG). It started off as a face-to-face household survey where each household member of 12 years or older was asked to report his/her mobility for two days. In 1985 the first large redesign took place. Interview modes changed to telephone and postal, and respondents reported their mobility on one specific day. This redesign led to discontinuities in the annual series of some of the statistics based on OVG. In 1994 the sample size of the DTS was substantially increased and from that year on children under 12 years old have also been included in the surveyed population of interest. In 1999, the DTS went through the second major redesign that featured some response motivation and follow-up measures. In preparation to this redesign a pilot based on the new design was conducted in 1998 in parallel with the survey under the old design. Based on the parallel surveys, correction factors were computed to correct the 1985-1998 OVG to the level of the new OVG. In 2004, the data collection for the survey was transferred to another agency. The survey design remained largely unchanged except for smaller sample sizes and some methodological changes. This 2004 transition also gave rise to discontinuities in some of the series, notably those disaggregated by

purpose. The DTS during the period from 2004 until the next major redesign in 2010 is referred to as MON (Mobiliteitsonderzoek Nederland). Since 2010 the DTS has been conducted by Statistics Netherlands again. In 2010 the survey changed to a person survey, and a sequential mixed mode design with face-to-face, telephone and web modes was established. This changeover led to sizeable discontinuities in many series. The years 2010 to 2017 constitute the OViN (Onderzoek Verplaatsingen in Nederland) period of the DTS. For more information on the history of the DTS and the changes made by the redesigns, we refer to Konen and Molnár (2007), Molnár (2007) and Willems and van den Brakel (2015).

Starting from 2018 another design is in place, named ODiN (Onderweg in Nederland). In ODiN several changes have been adopted, most of which can have systematic effects on the level of the observed mobility characteristics. Among the most important changes are

- The questionnaire has been completely redesigned.
- Children aged 0-5 are no longer surveyed.
- The definition of 'regular' mobility, which is the definition used in most DTS-based publications, has changed, and now includes domestic holiday and professional mobility. Flight trips are no longer observed in ODiN.
- ODiN is a CAWI-only survey, where OViN used a sequential mixed-mode CAWI-CAPI/CATI strategy. To boost response rates incentives in the form of (a chance to win) tablets are now being used.
- A new weighting scheme is used to better account for the changed composition of the response due to the design changes, including changes in regional oversampling.

The DTS considers only mobility within the Netherlands. Also, the DTS uses the concept of regular mobility, which until 2017 meant the mobility excluding holiday mobility (both domestic and abroad) and professional transportation mobility. Therefore, trips with purpose holiday and professional transportation have been removed from the survey data, as far as possible. As mentioned, however, the definition of regular mobility has changed under the ODiN redesign and now includes both domestic holiday and professional mobility. In the case of professional mobility, the data allow to identify such trips, and it was decided to exclude professional mobility, to be more consistent with the definition of mobility used previously in the trend estimation project. However, the ODiN questionnaire design does not easily allow to remove the observed domestic holiday mobility. This means that the ODiN discontinuities will include the effects of this change in mobility definition.

For the OViN years the data contain a small number of trips for children under the age of 12 with purpose "Work", and we have changed this to purpose "Other". Flight trips are also removed from the data, because they are no longer reported in ODiN and because they gave rise to some unstable estimates of distance travelled for mode "Other". Until last year, data about all age classes have been used, but it has now been decided to drop age group 0-5 years, since it is no longer observed under the ODiN design. It would be possible to still use the older estimates for this age class and keep extrapolating, while also borrowing some strength from neighbouring age class 6-11, but the resulting time series would become more and more uncertain.

Even though the DTS dates back to 1978, it was decided in Boonstra et al. (2019) to only use DTS data starting from 1999, the first year of the new OVG survey. This turned out to be sufficient for the purpose of obtaining reliable trends over the last 15 years or so. It also means that the large discontinuities arising from the 1999 redesign need not be

modelled.

It is considered important that mobility trend estimates based on the DTS are in line with external data sources on mobility. Such information has been used in a plausibility analysis, and a few external sources have also been considered for use as auxiliary information in the time series models used for trend estimation. One such source is a time series of annual total passenger train kilometers based on passenger surveys run by the Dutch railways NS. These series also include data on train rides by other private companies active in the Netherlands. Another relevant data source is the annual series of road intensities, compiled by Statistics Netherlands from road induction loop data. In addition, there is a time series of registered kilometers driven by cars from Nationale Autopas (NAP). This series includes kilometres driven abroad but nevertheless it is a potential covariate in the time series models based on the DTS. Finally, annual figures for a set of weather characteristics collected by the Royal Netherlands Meteorological Institute (KNMI) have been considered as additional auxiliary series.

3 Direct estimates

Basing a time series model for mobility trends directly on the micro-data from all years would require a very complex model that must account for non-response, different aggregation levels of interest, discontinuities, time trends, etc, all at once, which would be computationally intractable. Instead we follow a two-step estimation procedure often used in small area estimation. In the first step, estimates and variance estimates of the target variables are obtained directly from each year's micro-data, at the most detailed aggregation level of interest. Here we make use of the existing survey weights, accounting for the sampling design and non-response. In the second step these so-called 'direct estimates' serve as input for a time series model, which can be used to compute improved estimates of mobility accounting for possible discontinuities caused by the redesigns. This section outlines the computation of the direct estimates from the OVG-MON-OViN-ODiN survey data, see also Boonstra et al. (2019).

The direct estimates are computed for all years from 1999 until 2021 for trip legs pppd and distance per trip leg. This results in two tables of 700 series of direct estimates at the most detailed breakdown level considered.

3.1 Point estimates

Point estimates are readily computed using the existing survey weights. First consider the number of trip legs, and let r_i denote the number of trip legs reported by person i for the surveyed day. The average number of trip legs pppd is then estimated by

$$\hat{R} = \frac{\sum_{i \in s} w_i f_i r_i}{\sum_{i \in s} w_i f_i}, \tag{1}$$

where the sums run over respondents, w_i are person weights satisfying $\sum_{i \in s} w_i = N$ with N the total population size, and f_i is a so-called vacation factor. The latter take values slightly less than 1, and are used to account for vacation mobility. The vacation factors are based on estimates obtained from the CVO (Continu Vakantieonderzoek) survey. They can be derived from the 'trip weights' v_i as

$$f_i = \frac{v_i}{D w_i}, \tag{2}$$

where D is the number of days in a year. The estimates (1) can be written more compactly in terms of the trip weights as

$$\hat{R} = \frac{\sum_{i \in s} v_i r_i}{\sum_{i \in s} v_i}. \tag{3}$$

The vacation factors have been used for official publications based on OVG, MON and OViN. In ODiN the vacation factors are no longer used, since a correction for non-response bias regarding vacation mobility is now integrated into the person weights w_i by using estimated population totals from CVO in the weighting scheme directly.

For the second target variable of interest, distance, we estimate the average distance per trip leg by

$$\hat{A} = \frac{\sum_{i \in s} w_i f_i a_i}{\sum_{i \in s} w_i f_i r_i} = \frac{\sum_{i \in s} v_i a_i}{\sum_{i \in s} v_i r_i}, \tag{4}$$

where a_i is the total distance for person i for all trip legs.

For estimates by mode and/or purpose, each particular category defines specific variables r and a referring only to the trip legs in that category, so that equations (1) and (4) still apply. For (further) subdivisions with regard to the person characteristics sex and age class, it is convenient to introduce a dummy variable δ_i for each combination of sex and age class, being 1 if person i belongs to this group and 0 otherwise, and then write instead of (1) and (4),

$$\begin{aligned} \hat{R} &= \frac{\sum_{i \in s} w_i f_i \delta_i r_i}{\sum_{i \in s} w_i f_i \delta_i}, \\ \hat{A} &= \frac{\sum_{i \in s} w_i f_i \delta_i a_i}{\sum_{i \in s} w_i f_i \delta_i r_i}. \end{aligned} \tag{5}$$

By using δ_i also in the denominator of $\hat{R}$, we obtain estimates of the means per sex, age class combination. Note that the denominator of $\hat{R}$ does not depend on any selection of purpose or mode.

As mentioned in the Introduction, at the most detailed level, each target variable gives rise to a set of 700 estimates per year, corresponding to the full cross-classification of person characterics sex and age class and trip characteristics purpose and mode. Some of the 700 domains are, however, non-existent. We refer to these domains as structural zeros, since the number of trips in these domains is zero by definition. This concerns the following domains: age 6-11 in combination with mode ”Car driver” or purpose ”Work” and age 12-17 mode car driver before 2011. Starting from 2011 it is possible to drive a car from age 17, and this can be seen in the data. Distances per trip leg corresponding to structural zero trip legs are undefined, and therefore missing in the set of direct estimates. Other occasional zeros for trip legs and missings for distance per trip leg occur in some years for ’rare domains’ such as education for the elderly. These accidental zeros and missings will be filled in by the predictions based on the time series models.

3.2 Variance estimates

For variance estimation we distinguish between person surveys (OViN, ODiN) and household surveys (OVG, MON). For the latter, the household is the unit of sampling. Observations from persons from the same household cannot be regarded as independent. For example, distances travelled by young children and their parents are often correlated, depending on purpose and mode. Variance estimates should account for the dependence between persons clustered within households.

First write estimates (1) and (4) in the general form

$$\hat{Y} = \frac{\sum_{i\in s} w_i y_i}{\sum_{i\in s} w_i z_i}, \tag{6}$$

which is a ratio of two population total estimates based on person weights w_i. For the average number of trip legs pppd, $y_i = f_i r_i$ and $z_i = f_i$; for the average distance per trip leg, $y_i = f_i a_i$ and $z_i = f_i r_i$.

Basic estimates of the sampling variances of $\hat{Y}$ that ignore the variation of the weights, finite population corrections and the variance of the denominator, are given by

$$v_0(\hat{Y}) = \frac{1}{(\sum_{i\in s} w_i z_i)^2} \frac{N^2}{n} S^2(y), \tag{7}$$

where n is the number of respondents, $S^2(y) = \frac{1}{n-1}\sum_{i\in s}(y_i - \bar{y})^2$ is the sample variance of y, with $\bar{y} = \frac{1}{n}\sum_{i\in s} y_i$ the sample mean of y.

These variance estimates are improved by taking into account (i) the variance of the denominator, (ii) the variance inflation due to variation of the weights (Särndal et al., 1989), and (iii) the variance reducing effect of some covariates used for stratification or weighting. The variance estimates incorporating all three improvements are computed as (see e.g. Särndal et al. (1992))

$$\begin{aligned}
v(\hat{Y}) &= \frac{n}{\left(\sum_{i\in s} w_i z_i\right)^2} S^2(we), \\
e_i &= e_i^y - \hat{Y} e_i^z, \\
e_i^y &= y_i - x_i'\hat{\beta}^y, \\
\hat{\beta}^y &= \left(\sum_{i\in s} x_i x_i'/u_i\right)^{-1} \sum_{i\in s} x_i y_i/u_i, \\
e_i^z &= z_i - x_i'\hat{\beta}^z, \\
\hat{\beta}^z &= \left(\sum_{i\in s} x_i x_i'/u_i\right)^{-1} \sum_{i\in s} x_i z_i/u_i.
\end{aligned} \tag{8}$$

Here $S^2(we)$ is the sampling variance of $w_i e_i$, where e_i are generalized residuals, defined in terms of regression residuals e_i^y for y and e_i^z for z. The regressions are based on vectors of covariates x_i and a positive variance factor u_i. For the persons survey case we use $u_i = 1$.

For the regressions defining the residuals in (8), the following covariate model is used:

$$hhsize + province + sex * ageclass + urbanisation + month + weekday + fuel$$

in which $hhsize$ is the number of persons in a household, $ageclass$ is as defined in the Introduction, $urbanisation$ is the degree of urbanisation of the residential municipality in 5 classes, $month$ is the survey month, $weekday$ the day in the week the response refers to, and $fuel$ is the fuel type of the car used by the respondent in three classes: petrol, other or none if the respondent doesn't use a car. These covariates represent an important subset of variables that have been used for stratification and weighting of the survey data over the years.

The variance formula (8) can be used for any variables y and z in (6) so it applies to all estimates by any combination of trip characteristics purpose, mode and person characteristics sex and age class.

We have compared the simple variance estimates computed with (7) with the refined ones based on (8), and observed that the differences are mostly modest but not generally negligible. The most important refinement turns out to be the variance inflation due to the variation of weights. This clearly increases the variance estimates for domain estimates based on widely varying weights.

For the years before 2010, when the surveys were conducted as household surveys, the same formulas can be used, with the understanding that the unit index i refers to households. In that case y_i, z_i refer to weighted household totals, the weights w_i to the average of the person weights within a household, and x_i to household totals of the weighting covariates. The regression variances u_i are taken equal to the household size, and n in (8) becomes the number of responding households. We refer to Boonstra et al. (2018) for further details as well as for plots of the direct estimates and their standard errors for trip legs and distance at several aggregation levels. These plots show that for 'common domains' such as purpose "Work" for age classes 30-39, 40-49, standard errors are stable and rather small. For rare domains the standard errors are on average much larger, and, like the point estimates, volatile and sometimes missing. Boonstra et al. (2018) also contains a short discussion about the covariances/correlations between the direct estimates within each year. Most of these cross-sectional correlations are small, but there are some large positive and negative ones. The largest positive correlations occur between estimates for modes that are often combined in a single trip, like "Walking" and "Train", while most negative correlations occur between modes that are rarely combined such as "Car driver" and "Cycling". Furthermore, in OVG/MON years, there are some more positive correlations induced by the household clustering, for example between estimates for parents ("Car driver") and children ("Car passenger") and purpose "Shopping" or "Other". The effect of the cross-sectional correlations on the (standard errors of) the trend estimates was tested using a simple multilevel time series model and found there to be quite small. However, due to computational problems we have not been able to use the full correlation matrices of the input estimates in the selected time series models, although we expect to see only a small effect there as well.

3.3 Transformations of input series

The direct estimates and standard errors of the number of trip legs and the distances serve as input for the multilevel time series models used to obtain more accurate and robust trend series. In Boonstra et al. (2019) it was found that instead of directly modeling the direct estimates it is better to first apply a transformation to these input estimates. Using a square-root transformation for trip legs and a log-transformation for distance was seen to improve both model fits as well as the convergence of the simulation-based model fitting procedure. These transformations also reduce the dependence between direct point estimates and standard errors. After fitting the model the inverse transformation is applied to produce the trend estimates, as detailed in Subsection 5.3.

Let $\hat{Y}_{it}$ denote the direct estimate for year t and domain i of the number of trip legs or distance. For the trip legs variable a square-root transformation is used: $\hat{Y}_{it}^{\text{sqrt}} \equiv \sqrt{\hat{Y}_{it}}$. A first-order Taylor linearisation yields approximated standard errors $se(\hat{Y}_{it}^{\text{sqrt}}) = se(\hat{Y}_{it})/(2\sqrt{\hat{Y}_{it}})$. Note that these standard errors are undefined for domains without observed trips (zero point estimate and standard error), but this is no problem as they are imputed using a Generalized Variance Function (GVF) smoothing model, as

described below.

For the distance variable a logarithmic transformation is used: $\hat{Y}_{it}^{\log} \equiv \log \hat{Y}_{it}$, which is well-defined as all distance input estimates are positive. Standard errors for the transformed data are approximated by first-order Taylor linearisation: $se(\hat{Y}_{it}^{\log}) = se(\hat{Y}_{it})/\hat{Y}_{it}$.

3.4 Smoothing the standard errors of the direct estimates

The time series models considered regard the (transformed) direct point estimates as noisy estimates of a true underlying signal. However, the accompanying variance estimates are largely treated as fixed, given quantities by the model. As the variance estimates can be very noisy due to the detailed estimation level, it is wise to smooth them before using them in the model. That way they better reflect the uncertainty of the direct estimates. The most obvious defect of the estimated standard errors is that they are zero in case of zero or one contributing sampling unit.[1)] This is correct for the structural zero domains, but it does not reflect the actual uncertainty about the accidental zero estimates for number of trip legs. For distance, the most problematic estimates are the zero variance estimates in case of a single contributing sampling unit. If there are no contributing sampling units for a certain domain then the direct distance estimate is treated as missing.

The models considered for smoothing the variance estimates are simple regression models relating the variance estimates to a few predictors such as sample size, design effects, and point estimates. Such models are known as Generalized Variance Function (GVF) models in the literature, see Wolter (2007), Chapter 7. As in Boonstra et al. (2019) the GVF smoothing models are applied to the transformed standard errors. The predictions from the GVF models are then used as (smoothed) standard errors accompanying the transformed direct estimates as input for the time series multilevel models. In particular, this yields reasonable standard errors for domains with no observed trips.

Let $\hat{Y}_{tijkl}^{\text{tr}}$ denote either the sqrt-transformed direct estimates for trip legs or the log-transformed estimates for distance, for year t, sex i, age class j, purpose k and mode l. For both target variables we use the same GVF smoothing model

$$\log se(\hat{Y}_{tijkl}^{\text{tr}}) = \alpha + \beta \log \tilde{Y}_{tijkl}^{\text{tr}} + \gamma \log(m_{tijkl} + 1) + \delta \log(\text{deff}_{tijkl}) + \epsilon_{tijkl}, \quad (9)$$

where m_{tijkl} is the number of sampling units (households or persons, depending on the survey year) contributing to domain (i, j, k, l) in year t, and

$$\text{deff}_{tijkl} = 1 + \frac{var(w)_{tijkl}}{\bar{w}_{tijkl}^2}, \quad (10)$$

is the design effect of the survey weights, in which the second term is the squared coefficient of variation of the weights of the contributing units to a specific year and domain.[2)] This factor accounts for the variance inflation due to the variation of the weights. Since we cannot trust the direct estimates for very small m_{tijkl}, the $\tilde{Y}_{tijkl}^{\text{tr}}$ on

1) This means that there is at most one unit (person or household) in a specific sex, age class domain, who reported a trip leg for a specific purpose, mode combination.

2) In case of 0 or 1 contributing units we have defined deff to equal 1.

predictor	coefficient	estimate	se
1	α	-0.625	0.011
$\log \tilde{Y}^{\text{sqrt}}_{tijkl}$	β	0.955	0.003
$\log(m_{tijkl}+1)$	γ	-0.507	0.001
$\log(\text{deff}_{tijkl})$	δ	0.363	0.006

Table 3.1 Estimated coefficients of the GVF model (9) for number of trip legs.

predictor	coefficient	estimate	se
1	α	-0.963	0.019
$\log \tilde{Y}^{\log}_{tijkl}$	β	0.209	0.010
$\log(m_{tijkl}+1)$	γ	-0.330	0.002
$\log(\text{deff}_{tijkl})$	δ	0.122	0.026

Table 3.2 Estimated coefficients of the GVF model (9) for distance per trip leg.

the right hand side of (9) are simple smoothed estimates

$$\begin{aligned} \tilde{Y}^{\text{tr}}_{tijkl} &= \lambda_{tijkl}\hat{Y}^{\text{tr}}_{tijkl} + (1-\lambda_{tijkl})\hat{Y}^{\text{tr}}_{\cdot\cdot jkl}, \\ \lambda_{tijkl} &= \frac{m_{tijkl}}{m_{tijkl}+1}, \end{aligned} \tag{11}$$

where $\hat{Y}^{\text{tr}}_{\cdot\cdot jkl}$ denotes the mean of $\hat{Y}^{\text{tr}}_{tijkl}$ over the years and sexes. For $m_{tijkl}=0$ this replaces the estimate by the mean over year and sex for the same age class, purpose and mode. For $m_{tijkl}=1$ the average of this mean and the estimate itself is used, and for large m_{tijkl} essentially the original point estimate is used.

The regression errors ϵ_{tijkl} are assumed to be independent and normally distributed with a common variance parameter σ^2. The GVF models are fitted to the positive standard errors of the transformed direct estimates. Summaries of the estimated model coefficients for trip legs and distance are given in Tables 3.1 and 3.2. The predicted (smoothed) standard errors based on the fitted models are

$$se_{\text{pred}}(\hat{Y}^{\text{tr}}_{tijkl}) = \exp\left(\hat{\alpha} + \hat{\beta}\log\tilde{Y}^{\text{tr}}_{tijkl} + \hat{\gamma}\log(m_{tijkl}+1) + \hat{\delta}\log(\text{deff}_{tijkl}) + \hat{\sigma}^2/2\right), \tag{12}$$

where $\hat{\sigma}$ is 0.11 for trip legs and 0.44 for distance. The R-squared model fit measures for both models are quite high: 0.89 for trip legs and 0.60 for distance. Note that the exponential back-transformation in (12) includes a bias correction, which in this case has only a small effect.

3.5 Bias correction

Direct estimates are usually constructed to be approximately unbiased. However, by applying a non-linear transformation, the estimates become slightly biased. Previously, we corrected for this bias in the back-transformation to the original scale (Boonstra et al., 2021). For the log transformation, a bias correction $\frac{1}{2}se_{\text{pred}}(\hat{Y}^{\log}_{it})^2$ was added to the predictions at the log-scale, and then the exponential back-transformation was applied, similar to the bias correction in (12). See also Fabrizi et al. (2018). On average, this indeed largely removed the small negative bias due to the log transformation. However, in several small domains this bias correction gave rise to visible small irregularities over time caused by variation in the (smoothed) standard errors at the log

scale. To avoid such irregularities, we now incorporate the bias correction already in the transformed input estimates. That is, for distance per trip leg, to which a log transformation is applied, we now define our input estimates as

$$\hat{Y}^{\log} \equiv \log \hat{Y} + \frac{1}{2} se_{\text{pred}}(\hat{Y}^{\log})^2 , \tag{13}$$

where $se_{\text{pred}}(\hat{Y}^{\log})$ are the GVF smoothed standard errors as discussed in Subsection 3.4 (here we simplify notation by suppressing subscripts). This way, any remaining irregularities in $se_{\text{pred}}(\hat{Y}_{it}^{\log})$ will now mostly get filtered out by the model, while on average the small negative bias is still largely removed.

For trip legs, where a square root transformation is applied, we can similarly incorporate a bias correction in the transformed input estimates, instead of in the back-transformation of the model predictions. The bias correction is derived from the fact that the design expectation of the direct estimates can be written as

$$\begin{aligned} E(\hat{Y}) &= E((\hat{Y}^{\text{sqrt}})^2) = E((\eta + e^{\text{sqrt}})^2) = \eta^2 + 2\eta E(e^{\text{sqrt}}) + E((e^{\text{sqrt}})^2) \\ &= \eta^2 + se_{\text{pred}}(\hat{Y}^{\text{sqrt}})^2 , \end{aligned}$$

where η denotes the model predictions at the square root scale, and e^{sqrt} is the vector of sampling errors after transformation, assumed to be normally distributed with (GVF smoothed) standard errors $se_{\text{pred}}(\hat{Y}^{\text{sqrt}})$. Here, the bias correction is not additive at the transformed scale, so in order to incorporate it in the input estimates we need to make a further approximation,

$$\eta = \sqrt{E(\hat{Y}) - se_{\text{pred}}(\hat{Y}^{\text{sqrt}})^2} \approx \sqrt{E(\hat{Y})} - \frac{1}{2} se_{\text{pred}}(\hat{Y}^{\text{sqrt}})^2 / \sqrt{E(\hat{Y})}, \tag{14}$$

justified by the fact that $se_{\text{pred}}(\hat{Y}^{\text{sqrt}})^2$ is relatively small compared to $\hat{Y}$. If we replace $E(\hat{Y})$ in the second term by $\hat{Y}$, we can use this term as a bias correction in the transformed input estimates, so that

$$\hat{Y}^{\text{sqrt}} \equiv \sqrt{\hat{Y}} + \frac{1}{2} se_{\text{pred}}(\hat{Y}^{\text{sqrt}})^2 / \sqrt{\hat{Y}} . \tag{15}$$

For domains with incidental zero estimates, $\hat{Y} = 0$, this does not work. For those domains we replace the bias correction term in (15) by a smoothed version, computed using the predictions obtained by GVF model (9) applied to the target variable $se_{\text{pred}}(\hat{Y}^{\text{sqrt}})^2 / \sqrt{\hat{Y}}$. We verified that the bias correction term is always relatively small compared to $\sqrt{\hat{Y}}$, and that to a large extent it removes the small negative bias due to the square root transformation, without introducing irregularities in the final trend estimates that were seen before when the bias correction was incorporated in the back-transformation.

4 Time series multilevel modelling

The time series multilevel models considered are extensions of the popular basic area level model proposed by Fay and Herriot (1979). The models are defined at the most detailed level, i.e. the full cross-classification of sex, age class, purpose, mode and year. Let us again denote by $\hat{Y}_{it}^{\text{tr}}$ the transformed direct estimates for either trip legs or distance in year t and domain i. Here domain i refers to a particular combination of sex, age class, purpose and mode, so that i runs from 1 to $M_d = 700$ and t from 1 to T

corresponding to the years 1999 to 2021. We further combine these estimates into a vector $\hat{Y} = (\hat{Y}_{11}, \ldots \hat{Y}_{M_d 1}, \ldots \hat{Y}_{1T}, \ldots \hat{Y}_{M_d T})'$. Note that $\hat{Y}$ is a vector of dimension $M = M_d T$. Structural zero domains are not modeled, and it is implicitly understood that they are removed from all expressions. The number of modeled initial estimates is thereby reduced from $M = M_d T = 700 \times 23 = 16100$ to a total of 15474. For distance per trip leg there are some additional domains without initial estimates due to the (coincidental) absence of observed trips. The total number of available initial distance estimates is 14840. For both target variables model estimates are eventually produced for all 15474 non-structurally-zero domains.

4.1 Model structure

The multilevel models considered take the general linear additive form

$$\hat{Y}^{\text{tr}} = X\beta + \sum_{\alpha} Z^{(\alpha)} v^{(\alpha)} + e\,, \tag{16}$$

where X is a $M \times p$ design matrix for a p-vector of fixed effects β, and the $Z^{(\alpha)}$ are $M \times q^{(\alpha)}$ design matrices for $q^{(\alpha)}$-dimensional random effect vectors $v^{(\alpha)}$. Here the sum over α runs over several possible random effect terms at different levels, such as transportation mode and purpose smooth trends, white noise at the most detailed level of the M domains, etc. This is explained in more detail below. The sampling errors $e = (e_{11}, \ldots, e_{M_d 1}, \ldots e_{M_d T})'$ are taken to be normally distributed as

$$e \sim N(0, \Sigma) \tag{17}$$

where $\Sigma = \text{diag}\left(se_{\text{pred}}(\hat{Y}^{\text{tr}}_{tijkl})^2\right)$, i.e. a diagonal matrix with values equal to the square of the smoothed standard errors computed as discussed in Subsection 3.4.

Equations (16) and (17) define the likelihood function

$$p\left(\hat{Y}^{\text{tr}} | \eta, \Sigma\right) = N\left(\hat{Y}^{\text{tr}} | \eta, \Sigma\right)\,, \tag{18}$$

where $\eta = X\beta + \sum_{\alpha} Z^{(\alpha)} v^{(\alpha)}$ is called the linear predictor. A Student-t distribution for the sampling errors in (17) has been considered instead of the normal distribution to give smaller weight to more outlying observations. This is a traditional approach for handling outliers in Bayesian regression, see e.g. West (1984). We allow the degrees of freedom parameter of the Student-t distribution to be inferred from the data. It has been assigned a Gamma$(2, 0.1)$ prior distribution, which was recommended as a default prior in Juárez and Steel (2010).

The fixed effect part of η contains an intercept and main effects and possibly the second-order interactions for linear trends, discontinuities and the breakdown variables sex, age, purpose and mode. The vector β of fixed effects is assigned a normal prior $p(\beta) = N(0, 100I)$, which is very weakly informative as a standard error of 10 is very large relative to the scales of the transformed direct estimates and the covariates used.

The second term on the right hand side of (16) consists of a sum of contributions to the linear predictor by random effects or varying coefficient terms. The random effect vectors $v^{(\alpha)}$ for different α are assumed to be independent, but the components within a vector $v^{(\alpha)}$ are possibly correlated to accommodate temporal or cross-sectional correlation. To describe the general model for each vector $v^{(\alpha)}$ of random effects, we suppress superscript α in what follows.

Each random effects vector v is assumed to be distributed as

$$v \sim N(0, A \otimes V)\,, \tag{19}$$

where V and A are $d \times d$ and $l \times l$ covariance matrices, respectively, and $A \otimes V$ denotes the Kronecker product of A with V. The total length of v is $q = dl$, and these coefficients may be thought of as corresponding to d effects allowed to vary over l levels of a factor variable, e.g. purpose effects ($d = 5$) varying over time ($l = 23$ years). The covariance matrix A describes the covariance structure among the levels of the factor variable, and is assumed to be known. Instead of covariance matrices, precision matrices $Q_A = A^{-1}$ are actually used, because of computational efficiency (Rue and Held, 2005). The covariance matrix V for the d varying effects is parameterized in one of three different ways:

- an unstructured, i.e. fully parameterized covariance matrix
- a diagonal matrix with unequal diagonal elements
- a diagonal matrix with equal diagonal elements

The following priors are used for the parameters in the covariance matrix V:

- In the case of an unstructured covariance matrix the scaled-inverse Wishart prior is used as proposed in O'Malley and Zaslavsky (2008) and recommended by Gelman and Hill (2007).
- In the case of a diagonal matrix with equal or unequal diagonal elements, half-Cauchy priors are used for the standard deviations. Gelman (2006) demonstrates that these priors are better default priors than the more common inverse gamma priors for the variances.

The following random effect structures are considered in the model selection procedure:

- Random intercepts for the M_d domains obtained by the full cross clasification of age, gender, purpose and mode. In this case $A = I_{M_d}$ and V is a scalar variance parameter, and the corresponding design matrix is the $M \times M_d$ indicator matrix for domains. This can be extended to a vector of random domain intercepts, random slopes for linear time effects and discontinuities due to the redesigns in 2004, 2010 and 2018. In that case V is a 5×5 covariance matrix, parameterized by variance parameters for the intercepts, linear time slopes and the coeffcients for the level interventions, and possibly 10 correlation parameters.
- Random effects that account for outliers. The data for some years appear to be of lesser quality. This is the case especially for data on the number of trip legs in 2009. In order to deal with such less reliable estimates, random effects can be used to absorb some of the larger deviations in such years. The corresponding effects are removed from the trend prediction. This is an alternative to the use of fat-tailed sampling distributions such as the Student-t distribution for dealing with outliers.
- Random walks or smooth trends at aggregated domain levels (e.g. purpose by mode). See Rue and Held (2005) for the specification of the precision matrix Q_A for first and more smooth second order random walks. A full covariance matrix for the trend innovations can be considered to allow for cross-sectional besides temporal correlations, or a diagonal matrix with equal or different variance parameters to allow for temporal correlations only.
- White noise. In order to allow for random unexplained variation, white noise at the most detailed domain-by-year level can be included. In this case $A = I_M$ and V a scalar variance parameter, and the design matrix is $Z = I_M$.

We also investigate generalisations of (19) to non-normal distributions of random effects. Relevant references are Carter and Kohn (1996) in the state space modelling context, Datta and Lahiri (1995), Fabrizi and Trivisano (2010) and Tang et al. (2018) in the small area estimation context, and Lang et al. (2002) and Brezger et al. (2007) in the

context of more general structured additive regression models. In particular, the following distributions are considered for various random effect terms:

- Student-t-distributed random effects
- Random effects with a so-called horseshoe prior (Carvalho et al., 2010).
- Random effects distributed according to the Laplace distribution. This corresponds to a Bayesian version of the popular lasso shrinkage, see (Tibshirani, 1996; Park and Casella, 2008).

These alternative distributions have fatter tails allowing for occasional large effects. The Laplace and particularly the horseshoe distribution have the additional property that they shrink noisy effects more strongly towards zero.

4.2 Model estimation

The models are fitted using Markov Chain Monte Carlo (MCMC) sampling, in particular the Gibbs sampler (Geman and Geman, 1984; Gelfand and Smith, 1990). See Boonstra and van den Brakel (2018) for a specification of the full conditional distributions. The models are run in R (R Core Team, 2015) using package `mcmcsae` (Boonstra, 2021). The Gibbs sampler is run in parallel for three independent chains with randomly generated starting values. In the model building stage 1000 iterations are used, in addition to a 'burn-in' period of 500 iterations. This was sufficient for reasonably stable Monte Carlo estimates of the model parameters and trend predictions. For the selected models we use a longer run of 5000 burn-in plus 10000 iterations of which the draws of every fifth iteration are stored. This leaves $3 * 2000 = 6000$ draws to compute estimates and standard errors. The convergence of the MCMC simulation is assessed using trace and autocorrelation plots as well as the Gelman-Rubin potential scale reduction factor (Gelman and Rubin, 1992), which diagnoses the mixing of the chains. For the longer simulation of the selected models all model parameters and model predictions have potential scale reduction factors below 1.1 (only a few parameters have diagnostic values > 1.02) and sufficient effective numbers of independent draws.

Several models of the form (16) have been fitted to the data. For the comparison of models using the same input data we use the Widely Applicable Information Criterion or Watanabe-Akaike Information Criterion (WAIC) (Watanabe, 2010, 2013) and the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). We also compare the models graphically by their model fits and trend predictions at various aggregation levels.

5 Model building, selected models, and model prediction

The transformed direct estimates for the 700 domains over the 1999-2021 period along with their smoothed standard errors, as described in Section 3, serve as input data for the time series models. Note that these input estimates already include a small bias correction to counter the small biasing effect of the non-linear transformation. The variables defining the domains (sex, $ageclass$, $purpose$, $mode$) and the years (yr) have been used in the model development in many ways, e.g. using different interactions of various orders. Some additional covariates have been constructed in order to model the discontinuities that occur as a result of the change-over to MON in

2004, to OViN in 2010, and ODiN in 2018, as well as to reduce the influence of some lesser quality 2009 input estimates. For the MON level breaks in 2004, a variable br_mon is introduced taking values 1 between 2004 and 2009, and 0 otherwise. For the OViN level breaks in 2010 a variable br_ovin is defined, taking values 1 for the years 2010 and later and 0 otherwise. Likewise, for the ODiN level breaks in 2018 a variable br_odin is defined, taking values 1 from 2018 and 0 otherwise. A slight modification of the break variables was necessary in order not to introduce artificial level breaks in the age 12-17 car driver domains, which are structurally zero domains before 2011. Also, for year 2009 a dummy variable $dummy\_2009$ has been created being 1 only for year 2009 and 0 otherwise. The variable $covid$ is now defined as taking the value 1 for the Covid pandemic years 2020 and 2021 and zero otherwise. This variable and its corresponding effects were introduced in Boonstra et al. (2021) to accommodate the large changes in mobility caused by the pandemic. The new ODiN 2021 data show that the effects on mobility in 2021 are often different from those in 2020, and smaller for most domains. Therefore, we introduce a further intervention variable $covid2021$, taking the value 1 for year 2021 only, and 0 otherwise. Finally, the year variable itself is also used quantitatively to define linear time trends, and for that purpose we use a scaled and centered version denoted $yr.c$. All mentioned variables are used in different parts of the model, with associated fixed and random slope effects.

Some other covariates extracted from other sources like Statistics Netherlands' Statline and KNMI meteorological annual reports have also been used as candidate covariates in the model development. The weather variables considered concern annual averages at a central measurement location in The Netherlands (De Bilt). From these weather variables only a variable $snowdays$ representing the number of snow days by year is used in the selected trip-legs model. In Boonstra et al. (2019) an administrative variable km_NAP representing annual registered car kilometers collected from Nationale Autopas (NAP) was used in the distance model. As this variable is usually not available in time for producing new mobility trend estimates it is no longer considered for inclusion in the models.

In order to improve the modelling of Covid effects on mobility over the years 2020 and 2021 we have tried to incorporate indicators for policy measures that were imposed over time to combat the pandemic. For this purpose we have used the stringency indicator for The Netherlands, as compiled and disseminated for all countries by the Oxford COVID-19 Government Response Tracker[3)] (Hale et al., 2021). This indicator is based on several underlying indicators about such measures as school closing, workplace closing, public transport restrictions, restrictions on gatherings, etc. Unfortunately, these indicators turned out not to be good predictors for the covid effects on mobility in 2020 and 2021. Where the stringency indicator shows more and stronger measures on average in 2021, the data show that effects on mobility are larger in 2020. A possible explanation is that the compliance with Covid measures might have eroded over time. We also tried the underlying indicators about different measures but it didn't help.

In the following two sub-sections, time series models developed for the number of trip legs and the distance per trip leg are discussed. Along the way, we emphasize the differences with the previously developed models as described in Boonstra et al. (2021), which have been taken as a starting point. Following that, it is described how the target

[3)] See https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker

trend estimates are derived from the developed time series models. The models are expressed as time series multilevel models in a hierarchical Bayesian framework and fit using a Markov Chain Monte Carlo (MCMC) simulation method, as described in Section 4.

5.1 Time series multilevel model for the number of trip legs

As described in Section 3, we model the square-root-transformed direct estimates (including bias correction) of the number of trip legs pppd, using the corresponding transformed and GVF-smoothed standard errors to define the variance matrix Σ of the sampling errors.

The model parameters in (16) are separated in fixed and random effects. The fixed effects specification for the updated model is

$$\begin{aligned} & sex * ageclass + purpose * mode + mode * snowdays \\ & \quad + purpose * br\_mon\_SLO \\ & \quad + (purpose + mode) * (br\_odin + covid + covid2021) \\ & \quad + covid\_shopping\_walking + covid2021\_shopping\_walking \\ & \quad + covid\_leisure\_walking + covid2021\_leisure\_walking \end{aligned} \tag{20}$$

Compared to the model described in Boonstra et al. (2021) the set of fixed effects has been extended in several ways:

1. The interaction term $purpose * br\_mon\_SLO$ has been added, in which br_mon_SLO is defined similarly to br_mon as equal to 1 only during the MON years, except that it is restricted to be 1 only for purposes 'Shopping', 'Leisure' and 'Other'. So the interaction means that different fixed MON break effects are included for these three purposes only. For the other two purposes 'Work' and 'Education' MON break effects for trip legs are hardly noticeable and so these effects are not modelled. The main reason for including this interaction term is to account for the sizeable MON break effects for 'Leisure' and 'Other'. Previously, these effects were hidden because the two purposes were not distinguished and their net effect was very small. Indeed, Figure A.8 shows MON break effects for 'Leisure' and 'Other' of opposite sign and of the same order of magnitude.
2. The interaction term on the 3rd line in (20) has been extended with the additional $covid2021$ effects to allow for different covid effects specific to 2021, related for example to different measures or compliance with these measures in the second Covid year.
3. More detailed Covid interactions for the combination of purposes 'Shopping' or 'Leisure' and mode 'Walking' have been added as fixed effects, corresponding to the last two lines of (20). These effects have been added to avoid the trend model components becoming too much influenced by what are most likely covid effects and their changes in 2021. So this way model estimates for the trends for number of trip legs corrected for Covid effects become more plausible.

The random effects part of the model for number of trip legs consists of the components listed in Table 5.1. The first component 'V_2009' is included to account for some very influential outliers in 2009. It uses a horseshoe prior distribution, which turned out to work well as for most domains the outlier effects are negligible, but for a few domains, notably those for children and purposes 'Shopping' and 'Other', they are very large, as can be seen from Figure A.19. The variable $dummy\_2009\_SO$ is a limited version of $dummy\_2009$, and only equals 1 for year 2009 in combination with age class 6-11 and purposes 'Shopping' or 'Other'. Associated with $dummy\_2009\_SO$ the variable

$dummy\_2009\_classes$ is the categorical variable with classes defined by the cross-classification of sex, age class, mode and purpose classes for which $dummy\_2009\_SO$ is nonzero, all other classes being grouped into a single remainder category.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_2009	$dummy\_2009\_SO$	scalar	$dummy\_2009\_classes$	horseshoe	29
V_BR	$1 + yr.c + br\_mon\_SO + br\_ovin + br\_odin + covid + covid2021$	unstructured	$sex * AR1(ageclass, 0.75) * purpose * mode$	Laplace	4900
RW1AMM	$ageclass * purpose * mode$	scalar	RW1(yr)	normal	8050
RW2MM	$purpose$	unstructured	$sex * mode * RW2(yr)$	normal	1610
WN_MM	$mode$	unstructured	$purpose * yr$	normal	805
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	16100

Table 5.1 Summary of the random effect components for the updated model for trip legs. The second and third columns refer to the varying effects with covariance matrix V in (19), whereas the fourth and fifth columns refer to the factor variable associated with A in (19). The last column contains the total number of random effects for each term.

The random effects component 'V_BR' includes MON, OViN and ODiN level break random effects, as well as random intercepts, linear time trends, and covid effects, varying over all domains (the cross-classification of sex, age class, purpose and mode). Compared to last year's model specific covid2021 effects have been added. The full covariance matrix V in (19) is now a 7 x 7 matrix parameterised in terms of standard deviation and correlation parameters. Here too the variable br_mon_SO indicates that MON breaks are only considered for purposes 'Shopping' and 'Other'. This component uses a Laplace prior distribution for the random effects, which shrinks noisy effects more while large significant effects are shrunk less. Note also that the effects follow an AR1 autoregressive process as a function of ordered age class. The autoregressive parameter is fixed at 0.75, an approximately optimal value arrived at by grid search. This way of exploiting the order of age classes further helps to stabilize the (difficult) estimation of break effects.

The next two model components are time trend components. The component 'RW1AMM' adds a local level trend component for each age class, purpose, mode combination. Component 'RW2MM' represents smooth trends for each combination of sex, mode and purpose, where different variances corresponding to different degrees of smoothness as well as correlations are allowed among the 5 purposes distinguished. The 'RW1AMM' component only uses a single ('scalar') variance parameter. and can be interpreted as a correction to the 'RW2MM' trends, allowing for differences between age classes. The contribution of the 'RW1AMM' effects to the overall signal turns out to be much smaller than that of the 'RW2MM' effects.

The component 'WN_MM' allows non-gradual time dependence at the mode-purpose level. It is modelled with a general covariance matrix among the modes. The motivation for this term is that it prevents the smooth 'RW2MM' component becoming too volatile.

Finally, the white noise component named 'WN' in Table 5.1 accounts for remaining

variation of the true average number of trip legs pppd over the domains and the years. For analyses that focus on long term evolutions it can be useful to compute specific smooth trends by excluding non-gradual time-dependent components 'WN', 'WN_MM' as well as the explicit $snowdays$ effects.

5.2 Time series multilevel model for distance per trip leg

For distance we model the log-transformed direct estimates of distance per trip leg, including bias correction, using the corresponding transformed and GVF-smoothed standard errors discussed in Section 3 to define the variance matrix Σ of the sampling errors. The use of Student-t distributed sampling errors in this case succeeds in reducing the influence of outliers sufficiently. The degrees of freedom parameter of the Student-t distribution is assigned a weakly informative prior and is inferred from the data.

Similar to the model for number of trip legs pppd, only main effects and second order interaction effects are used in the fixed effects part of the selected model. In the updated model the following fixed effects components are included:

$$sex * ageclass + purpose * mode + mode * yr.c + br\_ovin\_walking + (purpose + mode) * (br\_odin + covid + covid2021)$$

The term $mode * yr.c$ represents linear time trends by mode. The term $br\_ovin\_walking$ represents a single OViN break fixed effect for mode 'Walking', and is included because the overall discontinuity for mode 'Walking' appears to be too large to model using random effects only. There are two differences compared to last year's model. First the $covid2021$ effects are now included, in the same way as in the model for trip legs. Second, an interaction term $sex * mode$ that was included previously has now been removed. It turned out that this term was mostly redundant and it seemed to occasionally cause a numerical instability in the Gibbs sampler.

Higher order interactions are modelled as random effects. The components included in the updated model are listed in Table 5.2.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_BR	$1 + yr.c + br\_mon + br\_ovin + br\_odin + covid + covid2021$	unstructured	$sex * AR1(ageclass, 0.75) * purpose * mode$	Laplace	4900
RW2MM	$purpose$	unstructured	$sex * mode * RW2(yr)$	normal	1610
WN_MM	$mode$	unstructured	$purpose * yr$	normal	805
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	16100

Table 5.2 Summary of the random effect components for the selected model for distance per trip leg. The second and third columns refer to the varying effects with covariance matrix V in (19), whereas the fourth and fifth columns refer to the factor variable associated with A in (19). The last column contains the total number of random effects for each term.

The only change made to the random effects components are the additional random coefficients for $covid2021$ in the 'V_BR' component. The components 'RW2MM', 'WN_MM' and 'WN' are the same as in Table 5.1. The distance model does not include a component for outliers, since there are no such extreme outliers in 2009 as in the series

for number of trip legs. Moreover, any outliers in the distance estimates are dealt with by the Student-t sampling distribution. Otherwise, the distance model does not include the 'RW1AMM' component of Table 5.1 to prevent overfitting the more noisy input series.

5.3 Trend estimation and derived estimates

The trend estimates of main interest are computed based on the MCMC simulation results as follows. First, simulation vectors of model linear predictions are formed, i.e.

$$\eta^{(r)} = X\beta^{(r)} + \sum_{\alpha} Z^{(\alpha)} v^{(\alpha,r)}\,, \tag{21}$$

where superscript r indexes the retained MCMC draws, and each $\eta^{(r)}$ is of dimension M. Consequently, the level break effects are removed or added, depending on the choice of benchmark level.

Until last year, the OViN level was used as benchmark level, i.e. all trend estimates were computed at the OViN measurement level, by including OViN effects but removing MON and ODiN effects. This year the decision was made that from now on the trend estimates are benchmarked to the level of the new ODiN design. There are currently 4 years of ODiN data available, so it is expected that ODiN break effects can now be estimated with sufficient accuracy. The estimation of these effects is admittedly hampered by the Covid pandemic and the necessary introduction of Covid effects in the model, but it is not expected that by waiting much longer the ODiN break effect estimates will change by a large amount. See Section 7 for the results of a revision analysis supporting this decision.

Given the way the level break dummies are coded, benchmarking to the ODiN level means that we need to add all ODiN break effects to the predictions referring to the OVG, MON and OViN years, add the OViN effects to the predictions of pre-OViN years, and in addition need to remove the MON effects from the predictions referring to the MON years. Also, the dummy effects for outliers ('V_2009' component in the model for trip legs) are removed. Covid effects are *not* removed. We further note that the survey errors e in (16) are absent from the linear predictor (21). The simulation vectors of linear predictors thus obtained are

$$\tilde{\eta}^{(r)} = \tilde{X}\beta^{(r)} + \sum_{\alpha} \tilde{Z}^{(\alpha)} v^{(\alpha,r)}\,, \tag{22}$$

where $\tilde{X}$ and $\tilde{Z}^{(\alpha)}$ are modified design matrices that accomplish the stated correction for level breaks and possibly outlier effects. Back-transformation of these vectors to the original scale yields the MCMC approximation to the posterior distribution of the trends. For the square root transformation as used for modelling the number of trip legs pppd, the back-transformation amounts to

$$\theta^{(r)} = (\tilde{\eta}^{(r)})^2\,. \tag{23}$$

For the log transformation, as used in modelling distance per trip leg, back-transforming $\tilde{\eta}^{(r)}$ to the original scale yields the MCMC approximation to the posterior distribution of the distance trends. In this case, the back-transformation is

$$\theta^{(r)} = e^{\tilde{\eta}^{(r)}}\,. \tag{24}$$

The means over the MCMC draws $\theta^{(r)}$ are used as trend estimates, whereas the standard deviations over the draws serve as standard error estimates.

Recall that η and θ are vector quantities with components for all year-domain combinations. We have computed the trends at the most detailed level, but also for

aggregates over several combinations of the domain characteristics. Aggregation of distance per trip leg involves the number of trip legs, and so requires combining the MCMC output for both target variables. By multiplying the distance per trip leg results by the number of trip leg pppd results we obtain the results for distance pppd. Aggregation amounts to simple summation over trip characteristics purpose and mode, and to population weighted averaging over person characteristics sex and age class. Inference for other derived quantities like total number of trip legs per day and total distance per day at different aggregation levels can also be readily conducted using the simulation results for the two modeled target variables.

6 Results

The appendices contain a rather complete set of time series plots for the number of trip legs pppd (A.3), distance per trip leg (A.4), and distance pppd (A.5) at different aggregation levels, including the most detailed level, based on the models for the number of trip legs pppd and the distance per trip leg described in Section 5. In addition, trend estimates of total number of trip legs per day at overall, purpose and mode levels are shown in Figures A.1, A.2, and A.3, respectively, in Appendix A.1. For total distance per day, these plots are shown in Figures A.4, A.5, and A.6 in Appendix A.2. The latter plots for (average) total trip legs and distance per day also involve the effects of changes in population sizes of all distinguished demographic subgroups (the $sex, ageclass$ combinations) over the years.

In all these plots the black lines correspond to the series of direct estimates, the red lines to the model fit based on all model components, i.e. the back-transformation of (21), and green lines are derived from the trend series (23) or (24) at the level observed under the latest ODiN survey design. So the green lines correspond to the trends benchmarked to the ODiN level with the effect of some strong outliers for trip legs in 2009 removed, but including the Covid effects on mobility, as explained in Subsection 5.3. However, it is also possible to compute 'Covid-corrected' trends by removing the Covid effects.

In the following sub-sections, the trend estimates obtained for number of trip legs and distance are discussed.

6.1 Trip legs

Trend estimates of number of trip legs per person per day are given in Appendix A.3.

Almost all trend series show clear Covid effects. In most cases the effect is slightly stronger (and downward) in 2020 as compared to 2021. An exception is the trend series for 'Walking' where the Covid effect is upward and much stronger in 2021, see Figure A.9.

Since purpose 'Leisure' is now distinguished from purpose 'Other', several MON break effects are now visible, most clearly in Figure A.8. The trend estimates correct for these relatively large and opposite break effects for purposes 'Leisure' and 'Other'. For young children aged 0-5 there are also large MON break effects but as the trends for this age group are no longer extrapolated, these results are not shown here. However, MON break effects are generally smaller than OViN and ODiN break effects. This is consistent with the fact that the OVG and MON designs are largely the same except that they were carried out by different data collection organisations.

Overall, the OViN level for number of trip legs is lower than the levels observed during MON, OVG and ODiN design periods, if we disregard the large drop due to Covid, as is clear from Figure A.7. The lower OViN level is especially clear for modes ”Car driver”, ”Train” and ”Cycling”, see Figure A.9.

There are some noteworthy differences in discontinuities between men and women trend lines, particularly for 30-39 and 40-49 age groups, see e.g. Figure A.23. In these particular cases the differences in the levels of the direct estimates between men and women are much larger during the OViN period. The particular OViN measurement errors for purpose ”Education”, female, age groups 30-39 and 40-49 are perhaps due to incorrect purpose assignment for mothers taking their children to school. The ODiN level for these domains seems to revert to approximately the OVG-MON level. Relatedly, an opposite movement can be discerned for purpose ”Other”. From this perspective it is an improvement that the trend estimates are now benchmarked to the ODiN instead of OViN level.

Since the trends are now defined at the ODiN level, the outcomes during the preceding OVG, MON and OViN design periods are corrected for the discontinuities induced by the redesigns relative to the ODiN design. It implies that due to the uncertainty of the estimated discontinuities the standard errors for the trend estimates in these periods are larger than in the ODiN period. The standard errors of the trend estimates in especially OVG and MON periods are therefore often larger than the variances of the direct estimates, which only measure sampling variation. See for example Figure A.7 for estimates at the overall level and Figures A.8 and A.9 for estimates by purpose and mode.

The input estimates for children, purposes ”Shopping” and ”Other” in 2009 are very different from those in other years, as can be seen from Figure A.19. There is a clear exchange between both purposes in 2009 for children, presumably due to systematic classification errors in the 2009 data. These effects have been captured by the random effect term ’V_2009’ of the model for trip legs. The trend lines show that the 2009 ’outliers’ are indeed neutralized by excluding the ’V_2009’ effects. This effect is even larger for the youngest children aged 0-5, but as mentioned before this group is not modelled anymore because it is no longer observed under ODiN.

Tables 6.1, 6.2 and 6.3 list the posterior means and standard errors of several variance components of the trip leg model. It is to be noted from Table 6.1 that some correlation parameters are large and negative, i.p. between random intercepts and OViN break and (overall) covid effects, between linear random slopes ($yr.c$) and OViN, ODiN break effects as well as (overall) covid effect, and between MON and ODiN break effects. The dependencies might indicate that it is not easy to disentangle these effects, which is to be expected from a series with relatively few observations per design period, and final years affected by the Covid pandemic. Note also that the covid2021 effects are small compared to the covid effects, which is reasonable as the covid2021 effects only capture the deviation from the already large covid effects in 2020.

Table 6.2 shows that the smooth trend components ’RW2MM’ are most flexible for purpose ’Leisure’. Table 6.3 shows that the white noise component at purpose-mode level ’WN_MM’ is most volatile for modes ’Walking’, ’Cycling’ and ’Car driver’. This model component can explain some of the irregularities of especially the walking and cycling series of direct estimates, which given their relatively small standard errors are not well described by smooth trends, see figures A.9 and A.10.

The differences in volatility by purpose and mode are clearly visible in Figure A.10.

Where the series for cycling and purpose 'Leisure' is highly volatile, which may be a real phenomenon caused e.g. by weather effects, other domains display more smooth trends despite volatile direct estimates, as for example 'Train' for purpose 'Leisure'. As the direct standard error estimates are much larger in this case, the model chooses a smoother trend series.

	(Intercept)	yr.c	br_mon_SO	br_ovin2	br_odin2b	covid	covid2021
(Intercept)	7.8 (0.3)	3.9 (6.1)	22.5 (8.9)	-46.4 (4.6)	-22.9 (5.7)	-57.2 (4.3)	7.8 (15.7)
yr.c		1.2 (0.1)	2.1 (11)	-45.5 (6.3)	-30.3 (7.9)	-38.3 (6.7)	-24.6 (21.4)
br_mon_SO			1.3 (0.1)	14.2 (10.0)	-33.5 (11)	-24.4 (11.7)	24.4 (22.9)
br_ovin2			14.2 (10)	2.3 (0.1)	9.5 (8.4)	35.8 (6.8)	16.2 (20.7)
br_odin2b					1.4 (0.1)	32.8 (8.0)	-7.6 (21.3)
covid						1.9 (0.1)	3.3 (21.9)
covid2021							0.3 (0.1)

Table 6.1 Estimated standard deviations and correlations (×100) for the 'V_BR' component. Numbers in parentheses are posterior standard errors.

Work	Shopping	Education	Leisure	Other
0.022 (0.016)	0.030 (0.021)	0.037 (0.017)	0.091 (0.023)	0.050 (0.026)

Table 6.2 Estimated standard deviations (×100) for the 'RW2_MM' component. Numbers in parentheses are posterior standard errors.

Car driver	Car passenger	Train	BTM	Cycling	Walking	Other
0.63 (0.08)	0.46 (0.07)	0.15 (0.07)	0.22 (0.08)	0.66 (0.08)	0.75 (0.09)	0.13 (0.07)

Table 6.3 Estimated standard deviations (×100) for the 'WN_MM' component. Numbers in parentheses are posterior standard errors.

6.2 Distance

Plots of trend estimates of mean distance per trip leg and mean distance per person per day are given in Appendices A.4 and A.5, respectively.

The direct estimates of distance per trip leg are rather volatile, even at the most aggregated level, see Figure A.88. The distance variable is also more affected by outliers, which occur in all years, and usually for domains with few observed trips. Therefore a Student-t distribution is used to fit the (log-transformed) distance variable. The posterior mean of the degrees of freedom parameter of the t distribution is 3.8 with a standard error of about 0.1.

Due to the noisier data, it is harder to detect fine changes in the underlying distance trends. In order to avoid overfitting, the model for distance is more parsimonious than that for the number of trip legs. One exception is that the distance model includes a fixed OViN break effect for mode 'Walking'. This effect was required to capture the very pronounced discontinuity in 2010 for mode 'Walking', as shown in Figure A.90, and is most likely due to the fact that from 2010 onwards walks are more often classified as single tours instead of consisting of a go and return trip.

Tables 6.4 - 6.6 list some parameter estimates (posterior means and standard errors) for the fit of the selected model described in Subsection 5.2. The 'V_BR' component containing varying coefficients by domain for intercept, linear slope over time, MON, OViN and ODiN breaks, and Covid effects, displays, as in the trip legs model, several negative correlations. There is also a strong positive correlation among ODiN breaks and

covid effects. This may indicate that it remains difficult to disentangle these effects. The diagonal values of Table 6.4 show that the variation of OViN and ODiN break effects is large relative to that of the MON break and Covid effects. Note also that the covid2021 effects are small compared to the covid effects, which is reasonable as the covid2021 effects only capture the deviation from the already large covid effects in 2020. The 'RW2_MM' component has a relatively large variance component for purpose 'Other' (Table 6.5). Table 6.6 shows that the mode-dependent scales of the of the 'WN_MM' effects are quite diverse. These effects are most pronounced for mode 'Other'.

	(Intercept)	yr.c	br_mon2	br_ovin2	br_odin2b	covid	covid2021
(Intercept)	17.1 (0.9)	20 (12)	13.7 (22.1)	-27.4 (9.8)	-21.9 (9.6)	-16.5 (13.6)	-9.1 (30.1)
yr.c		4.0 (0.6)	30.9 (29.1)	-39.8 (14.4)	-18.7 (16.4)	-40.5 (18.4)	-6.9 (36.6)
br_mon2			1.6 (0.8)	-15.1 (26.2)	-30.3 (28.2)	-32.5 (30.1)	-1.3 (37.7)
br_ovin2				9.9 (1.1)	-6.5 (15.2)	6.1 (20.2)	-0.5 (32.3)
br_odin2b					8.5 (0.9)	61.3 (14.3)	-1.7 (36.6)
covid						5.2 (1.0)	-2.7 (38.6)
covid2021							1.1 (0.8)

Table 6.4 Estimated standard deviations and correlations (×100) for the 'V_BR' component. Numbers in parentheses are posterior standard errors.

Work	Shopping	Education	Leisure	Other
0.20 (0.13)	0.15 (0.12)	0.15 (0.12)	0.11 (0.09)	0.94 (0.28)

Table 6.5 Estimated standard deviations (×100) for the 'RW2_MM' component. Numbers in parentheses are posterior standard errors.

Car driver	Car passenger	Train	BTM	Cycling	Walking	Other
1.7 (0.5)	2.7 (0.7)	1.1 (0.8)	4.2 (1.1)	0.7 (0.4)	3.1 (0.5)	9.5 (1.4)

Table 6.6 Estimated standard deviations (×100) for the 'WN_MM' component. Numbers in parentheses are posterior standard errors.

6.3 Model alternatives

In this subsection a few of the alternative model specifications that have been attempted are summarized.

First, as mentioned in Section 5 a model was attempted in which the differences between 2020 and 2021 Covid effects on mobility were modelled using indicators for the severity of Covid measures imposed over time. However, it turned out that ODiN data of 2020 and 2021 are not in line with the stringency of Covid measures for the same years, so that there is no gain in using these stringency measures as covariates in the model. Instead, we have introduced fixed and random effects for the deviation of Covid effects in 2021 from the Covid effects in 2020, for both the trip legs and distance models.

We also tried a variant of what was termed 'ordinal model' in Boonstra et al. (2021). The main difference with the models described in Section 5 is that the 'RW2_MM' smooth trend component is replaced by one with separate variance parameters for each mode, purpose combination, at least for the trip legs model, and that the 'WN_MM' component is dropped. The results show that also after including ODiN 2021 data, the 'ordinal model' results in worse model criteria (WAIC), trends that seem too rigid for several domains, and an implausible trend behaviour in the years 2009-2011 for mode 'Walking' trip legs.

A model with additional random instead of fixed covid and covid2021 effects by purpose and mode was attempted. This didn't give sufficient flexibility, in that the large covid2021 effect for mode 'Walking', purpose 'Leisure' (see Figure A.16) could not be sufficiently captured by the random effects, even when a horseshoe prior was used for the distribution of the latter.

For the distance model, we also tried to include the covid/covid2021 interaction fixed effects for purposes 'Shopping' and 'Leisure' and mode 'Walking'. However, in this case it turned out that the random effects in component 'V_BR' are sufficient to capture such interaction effects.

Several models with added effects for some apparent coding problems during the first two MON years 2004 and 2005 for purposes 'Leisure' and 'Other' have been tried. In the end, however, such effects seemed to be strongest for age group 0-5. After this age group was dropped from the modelling, there was no further need for these effects.

Several variants of the GVF model for smoothing input variance estimates have been tried. A shortcoming of the current GVF models is that the smoothed variances for a domain with 0 trip leg observations in a certain year are sometimes smaller than the smoothed variances for the same domain in another year with at least one observation. This is counterintuitive at least, and may sometimes put too much weight on zero trip leg estimates. Fortunately, such domains have little influence on trend estimates at more aggregate levels, but it would nevertheless be better to improve the GVF models in this respect. We tried a GVF model where the covariate $\log(m_{tijkl} + 1)$ in (9) is replaced by $\log(m_{tijkl} + \epsilon)$ with $\epsilon = 1/2$ or $\epsilon = 1/4$. This did not change much and somehow gave rise to a slightly larger negative bias. In another attempt, the variance estimates for each year with zero trip legs for a certain domain were replaced by the maximum variance of estimates for the same domain over all years with a positive number of observed trip legs. This is a somewhat ad hoc method though the results are mostly not implausible. This will be taken as a starting point for a continued search for improved GVF models next year.

7 Revision analysis

In 2018 the survey design of the DTS changed from OViN to ODiN. Until 2020 trend estimates are published on the level of OViN, which is the survey design that is used for data collection from 2010 until 2017. A revision analysis is conducted to investigate whether it is possible to change the level of the trend estimates to ODiN, i.e. the level of the survey design that is currently used for data collection. To this end 'real-time' estimates for the trends are made for the period that data are collected under ODiN. This implies that trend estimates are made for the time series observed until 2018 and the trends are re-estimated repeatedly by sequentially adding one new annual dataset at a time. As a result, four sets of trend estimates are obtained using the series observed until 2018, 2019, 2020 and 2021. Subsequently the size of the revisions of the trend estimates that are the result of adding the observations under ODiN one by one, are analysed. This is done for the trends estimated both at the level of OViN and ODiN.

To visualize the impact of the revisions, the relative absolute revisions (RAR) are calculated. Let $\theta_{i,t}^{k,j}$ denote the trend estimate for domain i estimated at the level of survey design k (OViN or ODiN), based on the time series observed until year $j = 2018, \ldots, 2020$, for year $t = 1999, \ldots, j$. The RAR is defined as the absolute revision

with respect to the estimates based on the most recent time series observed until 2021, i.e.

$$RAR_{i,t}^{k,j} = 100 * \frac{|\theta_{i,t}^{k,j} - \theta_{i,t}^{k,2021}|}{\theta_{i,t}^{k,2021}}, \quad k = \text{OViN, ODiN}, \quad j = 2018, \ldots, 2020. \tag{25}$$

Estimates are produced for number of trip legs per person per day and the mean distance per trip leg. This is done at the national level, the five levels for purpose (Work, Shopping, Education, Leisure, Other) and the seven levels of transportation mode (Car driver, Car passenger, Train, BTM (bus/tram/metro), Cycling, Walking, Other). In Appendix B the following four figures are presented for each variable:

- trend estimates based on the time series observed until 2018, 2019, 2020 and 2021 at the level of OViN,
- trend estimates based on the time series observed until 2018, 2019, 2020 and 2021 at the level of ODiN,
- RAR for the years 2018, 2019, 2020 with respect to 2021 for the trend estimates at the level of OViN,
- RAR for the years 2018, 2019, 2020 with respect to 2021 for the trend estimates at the level of ODiN.

The figures illustrate that the revisions are the smallest for the period where the level of the trend estimates coincides with the data collection method. In other words, the revisions for the trend estimates at the OViN level are the smallest during the time period from 2010 until 2017 where the data are collected under the OViN design. The revisions of the trend estimates at the OViN level are larger during the period that data are collected under the ODiN design between 2018 until 2021 and also during the OVG and MON design from 1999 until 2009. In a similar way, the revisions of the trend estimates at the ODiN level are generally smaller between 2018 and 2021, when the data collection is based on the ODiN design and larger during the period 1999 until 2017, when the data are collected under other designs. Another general pattern is that the revisions increase in 2020 in the trend estimates at the level of OViN as well as ODiN. This is caused by the level interventions that are added to the models to accommodate the effects of the COVID crisis on mobility. It is anticipated that information in future observations will mainly be used to better estimate the covid intervention coefficients. They will to a much lesser extent lead to an improvement of the estimates of the ODiN discontinuities.

8 Discussion

In Boonstra et al. (2019, 2021) and Boonstra et al. (2021) multilevel time series models have been developed to estimate mobility trends based on the Dutch Travel Survey (DTS). This report provides an update on the model development as well as on the resulting trend estimates. Compared to the previous report, data over 2021 have been added, which means that the length of the time series has increased to 23 years, of which the last four years correspond to data observed under the latest ODiN implementation of the DTS.

The last two years of ODiN mobility data are strongly influenced by the Covid pandemic. It turned out (Boonstra et al., 2021) necessary to add explicit Covid effects to the model,

to avoid that the large changes (mostly declines) in mobility in 2020 do not adversely affect the 2018 level break estimates associated with the transition to ODiN. In the current update with ODiN 2021 data it turns out that the model also needs to include components to account for the differences between the Covid effects on mobility in 2020 and 2021. For most domains the effects are larger in 2020, one clear exception being leisure walking trips. The models have therefore been extended to include specific Covid effects for 2021. The explicit modelling of Covid effects also enables to estimate covid-corrected trends, i.e. trends that would likely have resulted if there were no pandemic.

As before, the target variables modelled are the number of trip legs per person per day and the distance per trip leg for domains that are defined by a cross-classification of sex, age, purpose, and transport mode at a yearly frequency. To better match information needs, the level of detail regarding the purpose classification has been increased from four to five categories, now including purpose 'Leisure'. Another change is that age group 0-5 years is no longer modelled. This change was anticipated, because children under 6 years old are not observed under the ODiN design. Predictions for this age group, including extrapolations up to 2020, have been computed until last year. However, to obtain reliable extrapolations becomes more and more difficult without new direct information, and it has now been decided to drop this age class.

Estimation takes place in two stages. In the first stage direct estimates as well as their standard errors are compiled from the DTS data by means of the general regression estimator. The direct estimates are the input for the time series models and are first transformed to better meet normality assumptions. For the number of trip legs a square-root transformation is used and for distance a log transformation. The standard error estimates are also transformed and subsequently smoothed using a generalized variance function (GVF) model. In the second stage, the resulting direct estimates at the level of the aforementioned cross-classification are used as input for the multilevel time series models, which are fitted using MCMC simulations.

A small improvement made to the models is that a bias correction, necessary because of the non-linear transformations applied to the original direct estimates, is now incorporated in the input estimates rather than in the back-transformation used in model prediction. This has the advantage that some amount of noise previously introduced into the model predictions is now filtered out by the model.

The models account for discontinuities due to three redesigns: the change-over from OVG to MON in 2004, the change-over from MON to OViN in 2010 and the change-over from OViN to ODiN in 2018. Discontinuities are predominantly modeled as random effects to reduce the risk of overestimation. Because of the increased level of detail by purpose, the model for trip legs has now been updated to include fixed MON break effects for modes 'Shopping', 'Leisure' and 'Other'. Previously this was not necessary as the largest discontinuities for purposes 'Leisure' and 'Other" are almost opposites and were therefore hidden when 'Leisure' was still part of the remainder category.

The DTS time series are also affected by outliers. The model for trip legs contains random effects to absorb the most dominant outliers in 2009, while the model for distances assumes a Student-t distribution for the sampling errors. The models further contain random intercepts (levels) and time slopes, as well as several trend components at different levels of the hierarchy defined by the sex, age, purpose and mode variables. Some of the random effects are assigned non-normal priors to induce a stronger form of regularization while also allowing occasional larger effects. Fixed effects included in the

models consist of main effects and some second-order interactions of the domain variables sex, age, purpose and mode, as well as an effect for the annual number of snow days in the model for trip legs.

As mentioned, one of the changes made to the models is that covid effects for 2021 have been added. The Covid effects, both overall effects and deviation effects for 2021, are modelled as before in the same way as the ODiN discontinuities. Fixed main effects by mode and separately by purpose capture some of the large effects at these aggregate levels, while strongly regularised random effects are included at the most detailed level. Moreover, fixed Covid effects for mode 'Walking' in combination with purposes 'Shopping' and 'Leisure' have now been added to accomodate some large interaction effects, especially the high number of leisure walking trips in 2021.

Model predictions and trends for the number of trip legs per person per day, distance per trip leg and distance per person per day are obtained at different aggregation levels of the cross-classification by aggregating the model predictions and trends at the most detailed level. Until last year (when the series extended up to 2020), the trends have been estimated at the OViN measurement level. From this year on, trends are instead estimated at the ODiN measurement level. This means that trends are derived from the model predictions including ODiN break effects and excluding the MON and OViN break effects. Whereas the trend estimates are corrected for the level break measurement effects towards the ODiN measurement level, the Covid effects on mobility are real and so the trends are not corrected for them. Nevertheless, it is possible to remove the Covid effects from the trends, and this could be useful for specific analyses.

There are several reasons for the change to correct trend estimates to the ODiN level. First, there are now four years of data under the latest ODiN design so that more data is available to estimate the ODiN discontinuites. Second, a revision analysis shows that the size of the rivisions of the trend estimates that are the result of adding the observations for an additional year are smallest for the time period where the level of the trend estimates coincides with the data collection method. This implies that the revisions during the ODiN design period are smallest if the trends are estimated at the ODiN level. The revision analysis also shows that the inclusion of the level interventions to accommodate Covid effects on mobility increases the revisions if the observations for 2020 and 2021 are added to the time series, regardless whether the trends are estimated at the level of OViN or ODiN. Note that more accurate break estimates generally mean smaller revisions when trend series are updated with new data. It is expected, however, that future observations will hardly improve the estimates for the ODiN discontinuities, since Covid will probably still have a varying and thereafter a lasting effect on mobility. From this point of view there is no reason to postpone the change to the ODiN level as a benchmark for the trend estimates. Third, for the number of trip legs, the levels observed under ODiN are most often similar to the levels observed under OVG and MON, whereas OViN levels are generally lower. There are reasons to believe that OViN suffers more from under-reporting, and altogether this suggests that the ODiN level is probably closer to the truth. There are also a few domains where OViN shows a much higher level than the other designs, most notably purpose 'Education' for women aged 30-39 (see Figure A.23). The latter example is presumably due to a coding problem in OViN. Finally, and perhaps most importantly is that many users are less familiar with the older OViN design, and from their perspective it is much more natural to correct the trend series towards the measurement level of the current design. The decision to as of now correct the trends series towards the ODiN level is supported by a revision analysis, in which the sizes of revisions due to sequentially including new annual

datasets are analysed.

Measured by the estimated standard errors, the trend series are most accurately estimated for the design period whose measurement level is used as the benchmark level. As discussed, this is now the ODiN level. The standard errors of the trend estimates during the preceding OVG, MON and OViN periods are larger, as a result of the uncertainty of ODiN break effects. This is especially true for OVG and MON periods as some large OViN discontinuity estimates also add to their uncertainty. Often, the standard errors in these periods are even larger than the design-based standard errors of the direct (untransformed) input estimates. This is because the latter do not account for (relative) measurement errors.

The modelling of Covid effects may have to be updated again next year, when data over 2022 will be added to the input series, as the effects on mobility in 2022 are likely different from both 2020 and 2021. Possible improvements to the GVF model for the smoothing of input variances may then also be further investigated. One shortcoming of the current GVF model is that variance estimates for zero trip leg estimates sometimes appear to be too small. This hardly affects aggregate level estimates, but it may further improve trend series for some domains with few observations.

References

Bollineni-Balabay, O., J. van den Brakel, F. Palm, and H. J. Boonstra (2017). Multilevel hierarchical bayesian versus state space approach in time series small area estimation: the dutch travel survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 180*(4), 1281–1308.

Boonstra, H. J. (2021). *mcmcsae: Markov Chain Monte Carlo Small Area Estimation*. R package version 0.7.0.

Boonstra, H. J. and J. van den Brakel (2018). Hierarchical bayesian time series multilevel models for consistent small area estimates at different frequencies and regional levels. Statistics Netherlands Discussion Paper, December 4, 2018.

Boonstra, H. J., J. van den Brakel, and S. Das (2018). Computing input estimates for time series modeling of mobility trends. CBS report.

Boonstra, H. J., J. van den Brakel, and S. Das (2019). Multi-level time series modeling of mobility trends - final report. Statistics Netherlands Discussion Paper, 30 October 2019.

Boonstra, H. J., J. van den Brakel, and S. Das (2021). Modeling of mobility trends - 2020 update including new odin data and level breaks. Statistics Netherlands Discussion Paper, February 2021.

Boonstra, H. J., J. van den Brakel, S. Das, and H. Wüst (2021). Modelling mobility trends - update including 2020 odin data and covid effects. Statistics Netherlands Discussion Paper, November 2021.

Brezger, A., L. Fahrmeir, and A. Hennerfeind (2007). Adaptive gaussian markov random fields with applications in human brain mapping. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 56*(3), 327–345.

Carter, C. K. and R. Kohn (1996). Markov chain monte carlo in conditionally gaussian state space models. *Biometrika 83*(3), 589–601.

Carvalho, C. M., N. G. Polson, and J. G. Scott (2010). The horseshoe estimator for sparse signals. *Biometrika 97*(2), 465–480.

Datta, G. S. and P. Lahiri (1995). Robust hierarchical bayes estimation of small area characteristics in the presence of covariates and outliers. *Journal of Multivariate Analysis 54*(2), 310–328.

Fabrizi, E., M. R. Ferrante, and C. Trivisano (2018). Bayesian small area estimation for skewed business survey variables. *Journal of the Royal Statistical Society Series C 67*(4), 861–879.

Fabrizi, E. and C. Trivisano (2010). Robust linear mixed models for small area estimation. *Journal of Statistical Planning and Inference 140*(2), 433–443.

Fay, R. and R. Herriot (1979). Estimates of income for small places: An application of james-stein procedures to census data. *Journal of the American Statistical Association 74*(366), 269–277.

Gelfand, A. and A. Smith (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association 85*, 398–409.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis 1*(3), 515–533.

Gelman, A. and J. Hill (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A. and D. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7*(4), 457–472.

Geman, S. and D. Geman (1984). Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence 6*, 721–741.

Hale, T., N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar, et al. (2021). A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature human behaviour 5*(4), 529–538.

Juárez, M. A. and M. F. J. Steel (2010). Model-based clustering of non-gaussian panel data based on skew-t distributions. *Journal of Business and Economic Statistics 28*(1), 52–66.

Konen, R. and H. Molnár (2007). Onderzoek verplaatsingsgedrag - methodologische beschrijving. Statistics Netherlands.

Lang, S., E.-M. Fronk, and L. Fahrmeir (2002). Function estimation with locally adaptive dynamic models. *Computational Statistics 17*(4), 479–500.

Molnár, H. (2007). Mobiliteitsonderzoek nederland - methodologische beschrijving. Statistics Netherlands.

O'Malley, A. and A. Zaslavsky (2008). Domain-level covariance analysis for multilevel survey data with structured nonresponse. *Journal of the American Statistical Association 103*(484), 1405–1418.

Park, T. and G. Casella (2008). The bayesian lasso. *Journal of the American Statistical Association 103*(482), 681–686.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rao, J. and I. Molina (2015). *Small Area Estimation*. Wiley-Interscience.

Rue, H. and L. Held (2005). *Gaussian Markov Random Fields: Theory and Applications*. Chapman and Hall/CRC.

Särndal, C.-E., B. Swensson, and J. Wretman (1989). The weighted residual technique for estimating the variance of the general regression estimator of the finite population total. *Biometrika 76*, 527–537.

Särndal, C.-E., B. Swensson, and J. Wretman (1992). *Model Assisted Survey Sampling*. Springer.

Spiegelhalter, D., N. Best, B. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B 64*(4), 583–639.

Tang, X., M. Ghosh, N. S. Ha, and J. Sedransk (2018). Modeling random effects using global–local shrinkage priors in small area estimation. *Journal of the American Statistical Association*, 1–14.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

van den Brakel, J., X. Zang, and S.-M. Tam (2020). Measuring discontinuities in time series obtained with repeated sample surveys. *International Statisitcal Review) 88*, 155–175.

Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research 11*, 3571–3594.

Watanabe, S. (2013). A widely applicable bayesian information criterion. *Journal of Machine Learning Research 14*, 867–897.

West, M. (1984). Outlier models and prior distributions in bayesian linear regression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 431–439.

Willems, R. and J. van den Brakel (2015). Methodebreukcorrectie ovin. PPM 210514/12, Statistics Netherlands.

Wolter, K. (2007). *Introduction to Variance Estimation*. Springer.

Appendix

A Time-series plots model-based and direct estimates

A.1 Total number of trip legs per day

Figure A.1 Direct estimates (black), model fit (red) and trend estimates (green) for total number of trip legs per day with approximate 95% intervals.

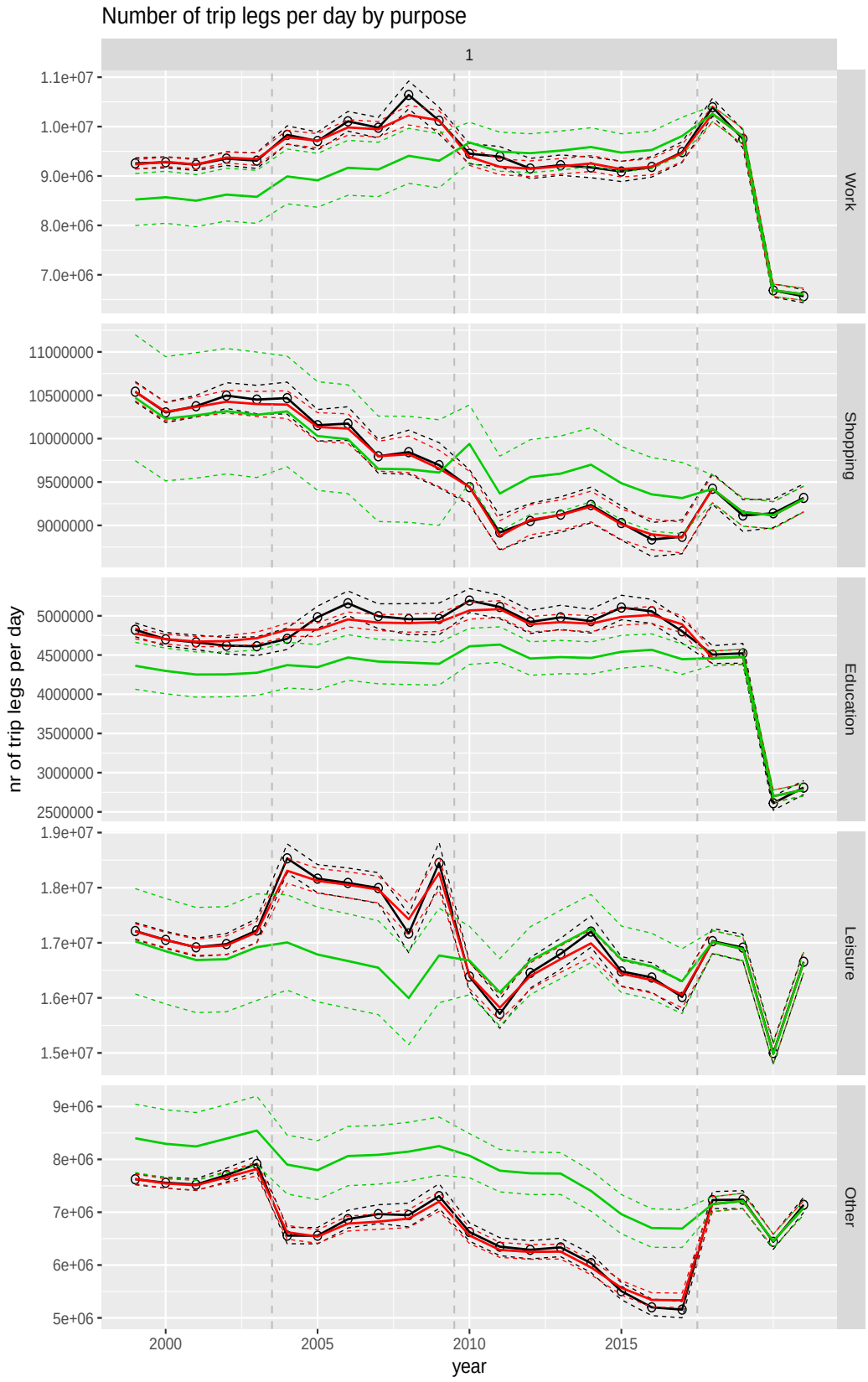

Figure A.2 Direct estimates (black), model fit (red) and trend estimates (green) for total number of trip legs per day by purpose with approximate 95% intervals.

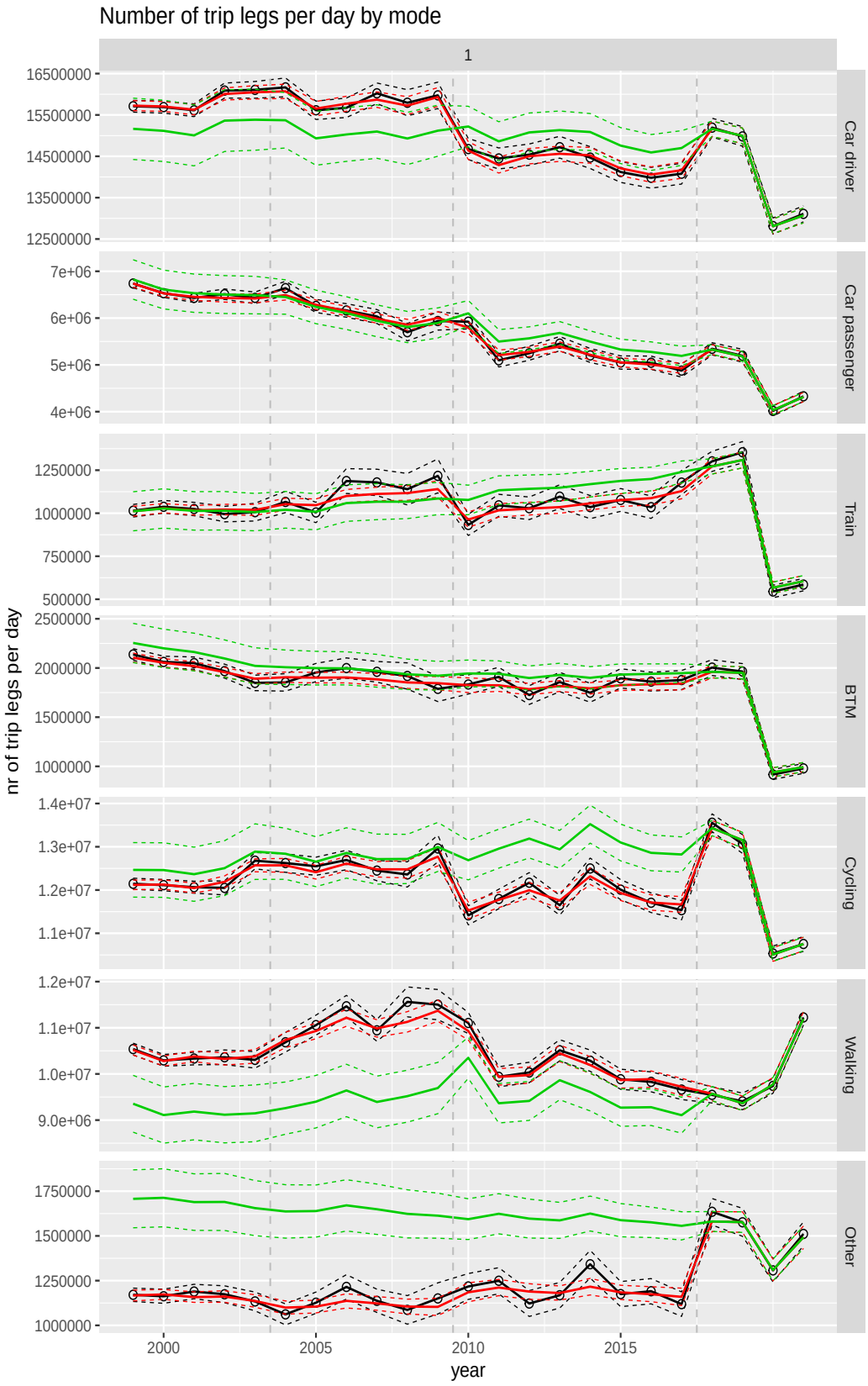

Figure A.3 Direct estimates (black), model fit (red) and trend estimates (green) for for total number of trip legs per day by mode with approximate 95% intervals.

A.2 Total distance per day

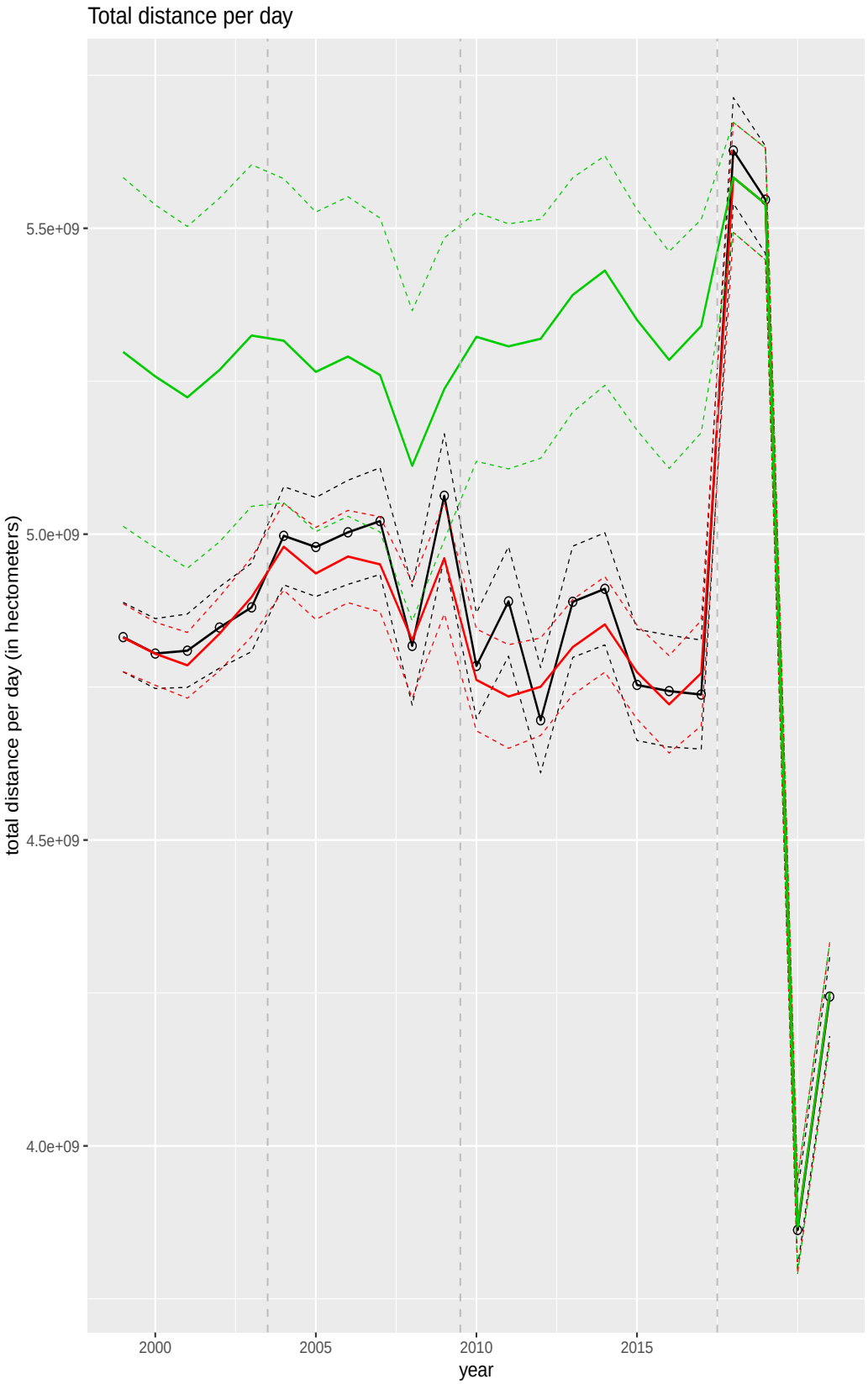

Figure A.4 Direct estimates (black), model fit (red) and trend estimates (green) for total distance per day with approximate 95% intervals.

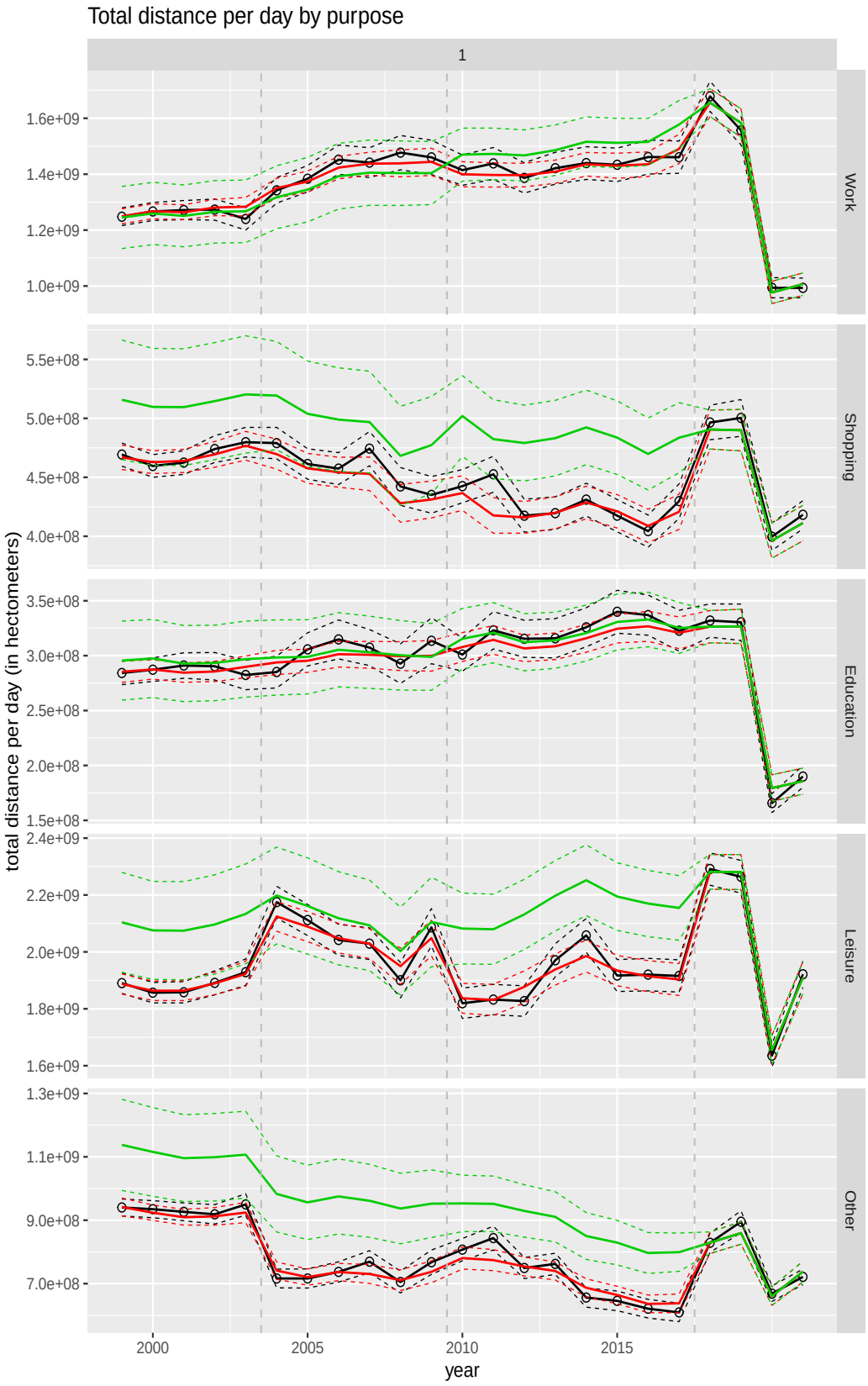

Figure A.5 Direct estimates (black), model fit (red) and trend estimates (green) for total distance per day by purpose with approximate 95% intervals.

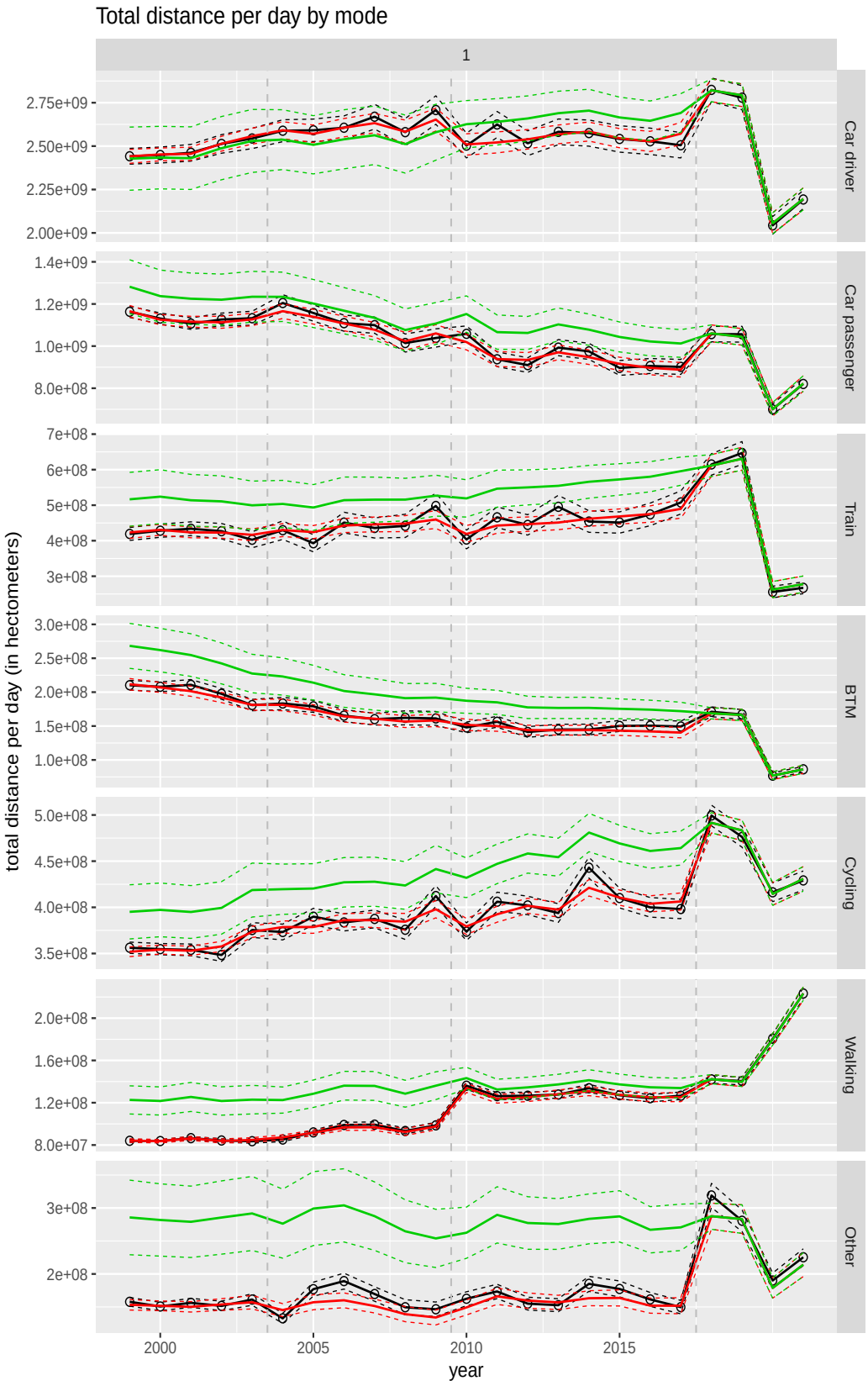

Figure A.6 Direct estimates (black), model fit (red) and trend estimates (green) for total distance per day by mode with approximate 95% intervals.

A.3 Number of trip legs per person per day

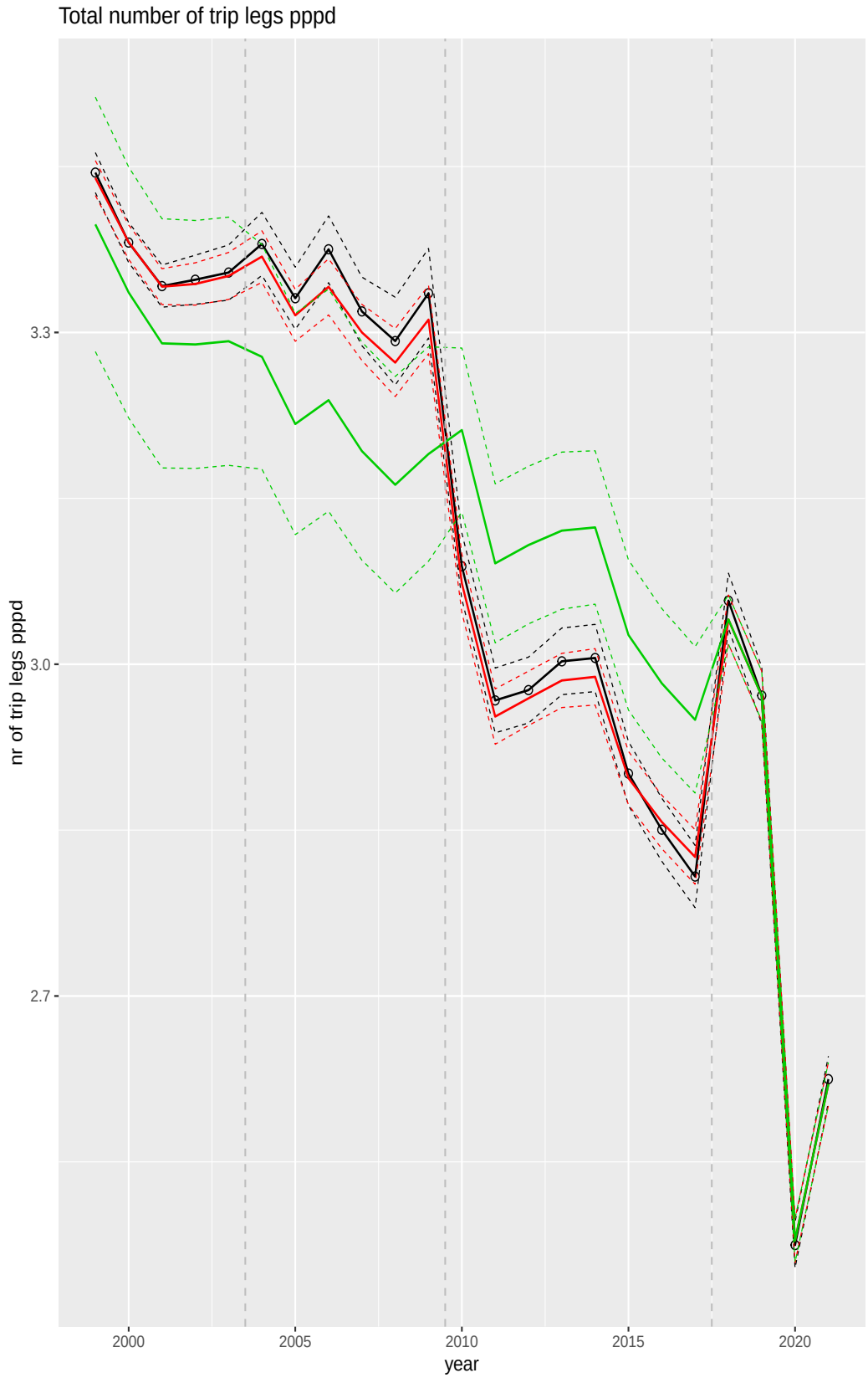

Figure A.7 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

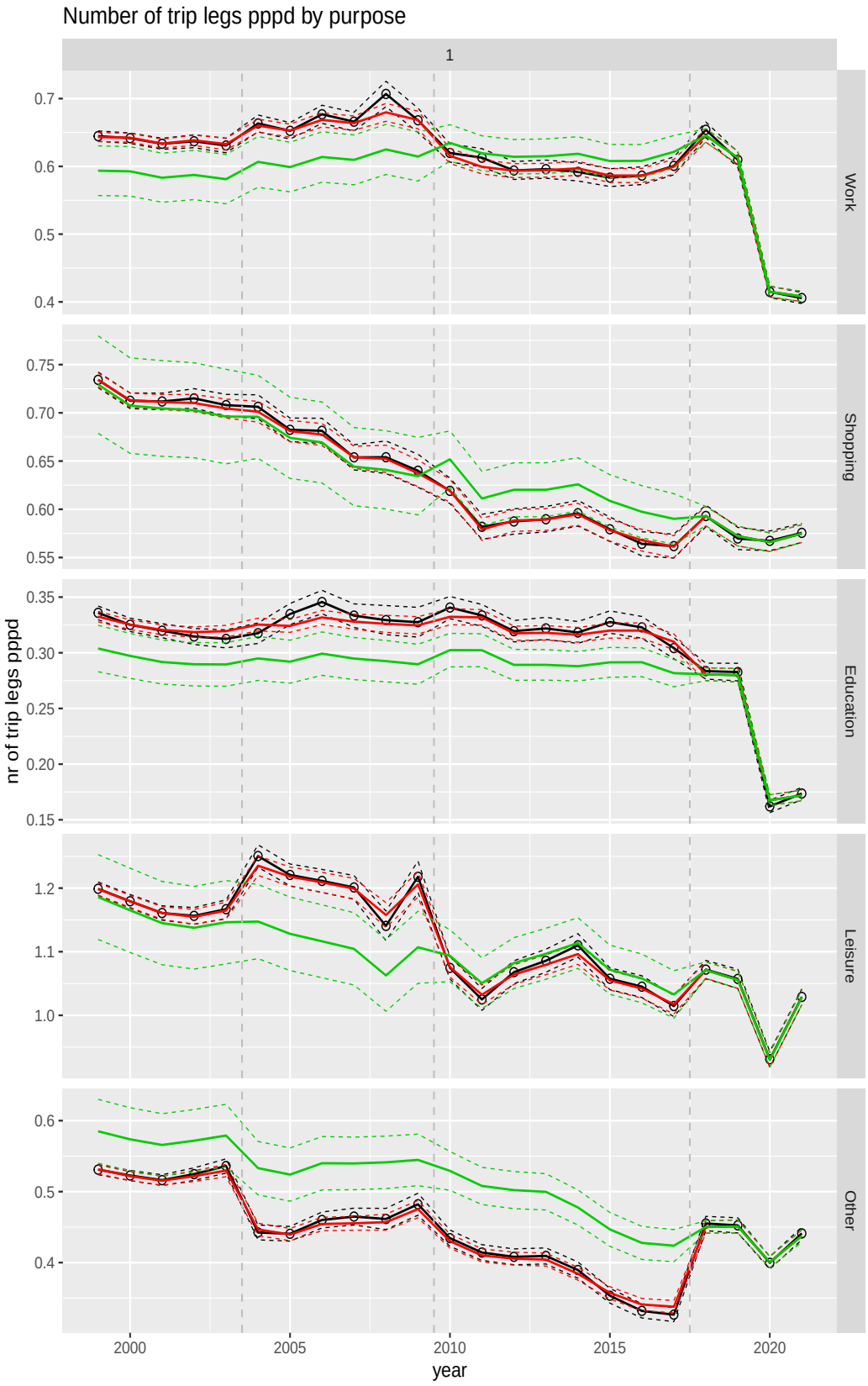

Figure A.8 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

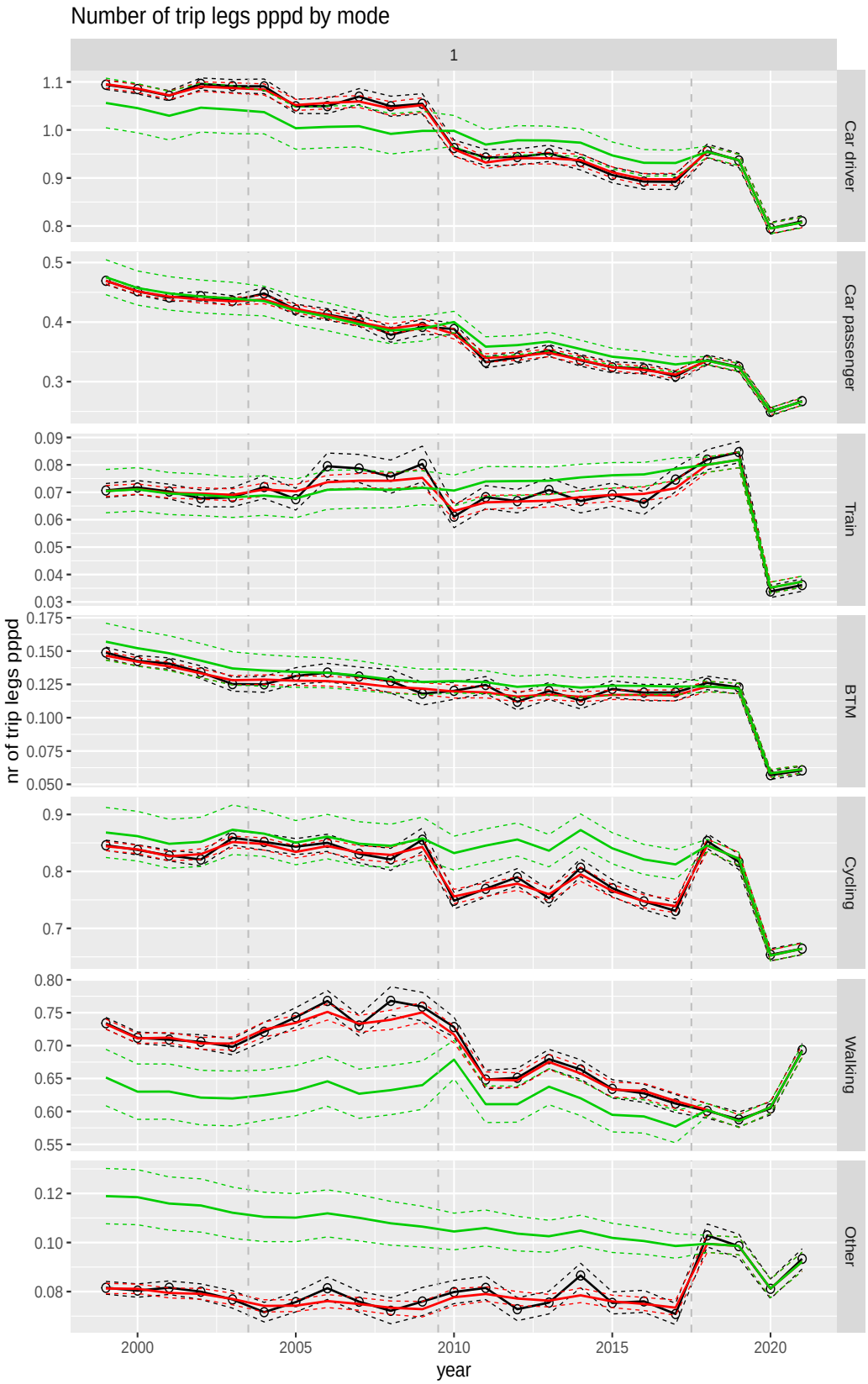

Figure A.9 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

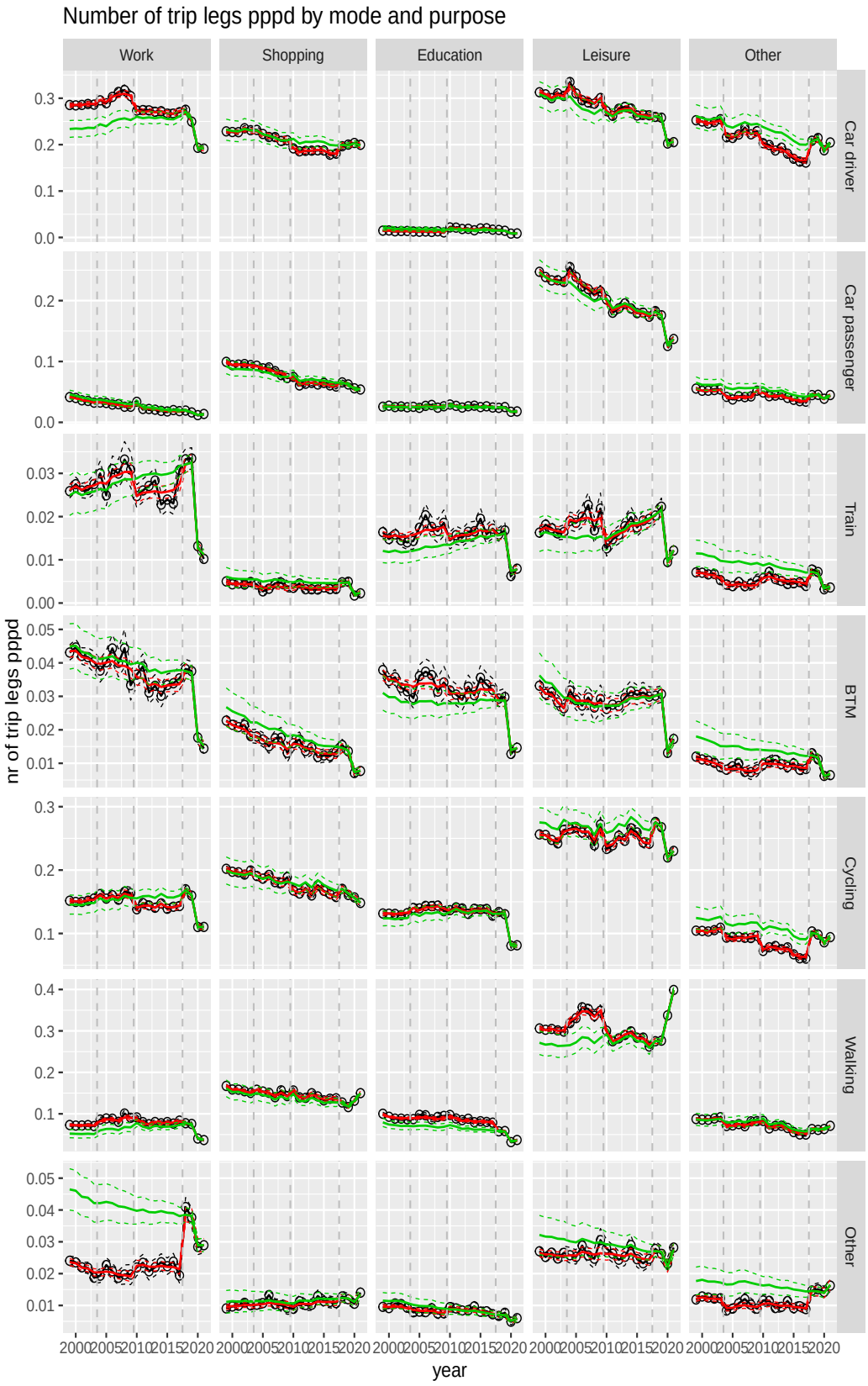

Figure A.10 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

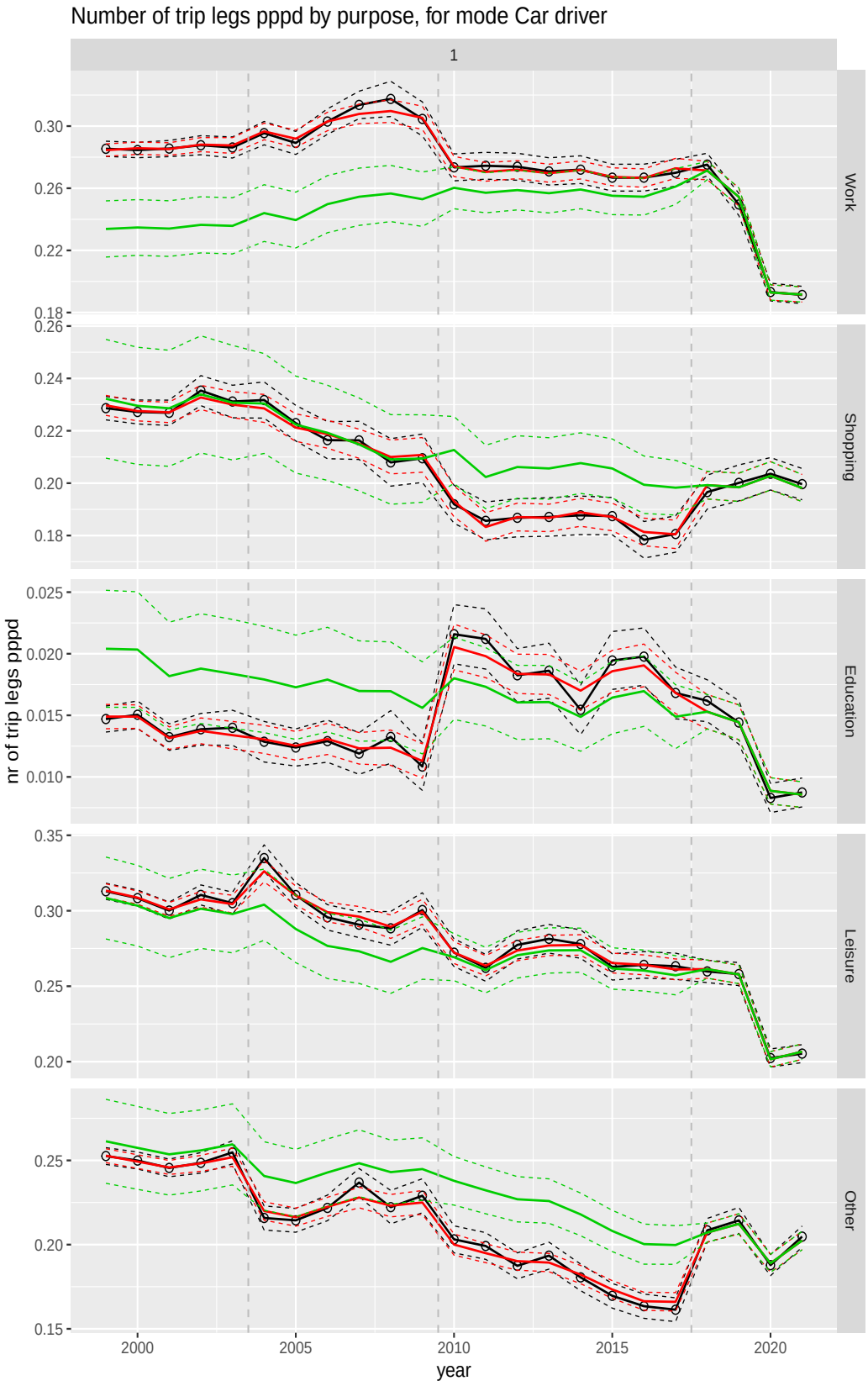

Figure A.11 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

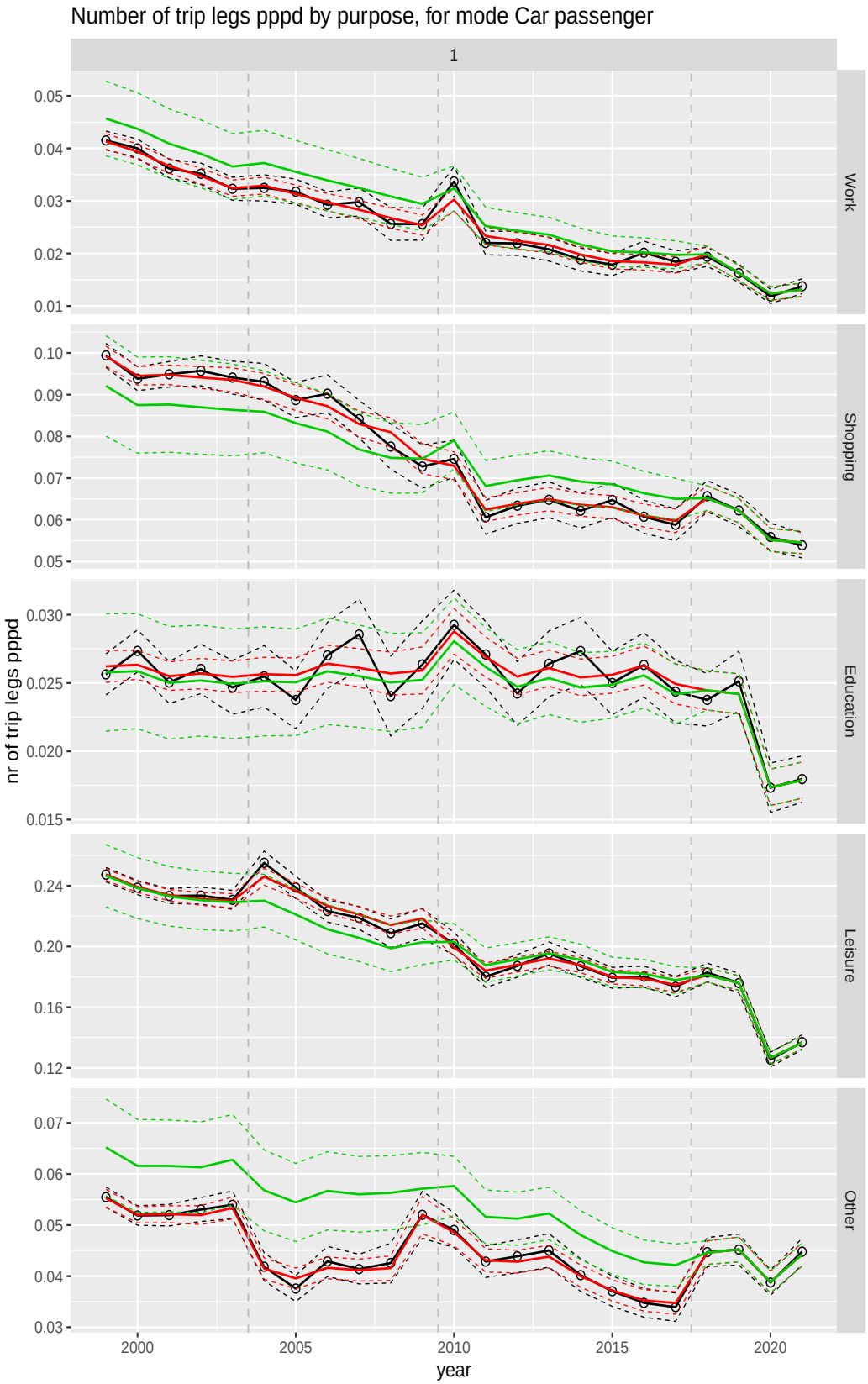

Figure A.12 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

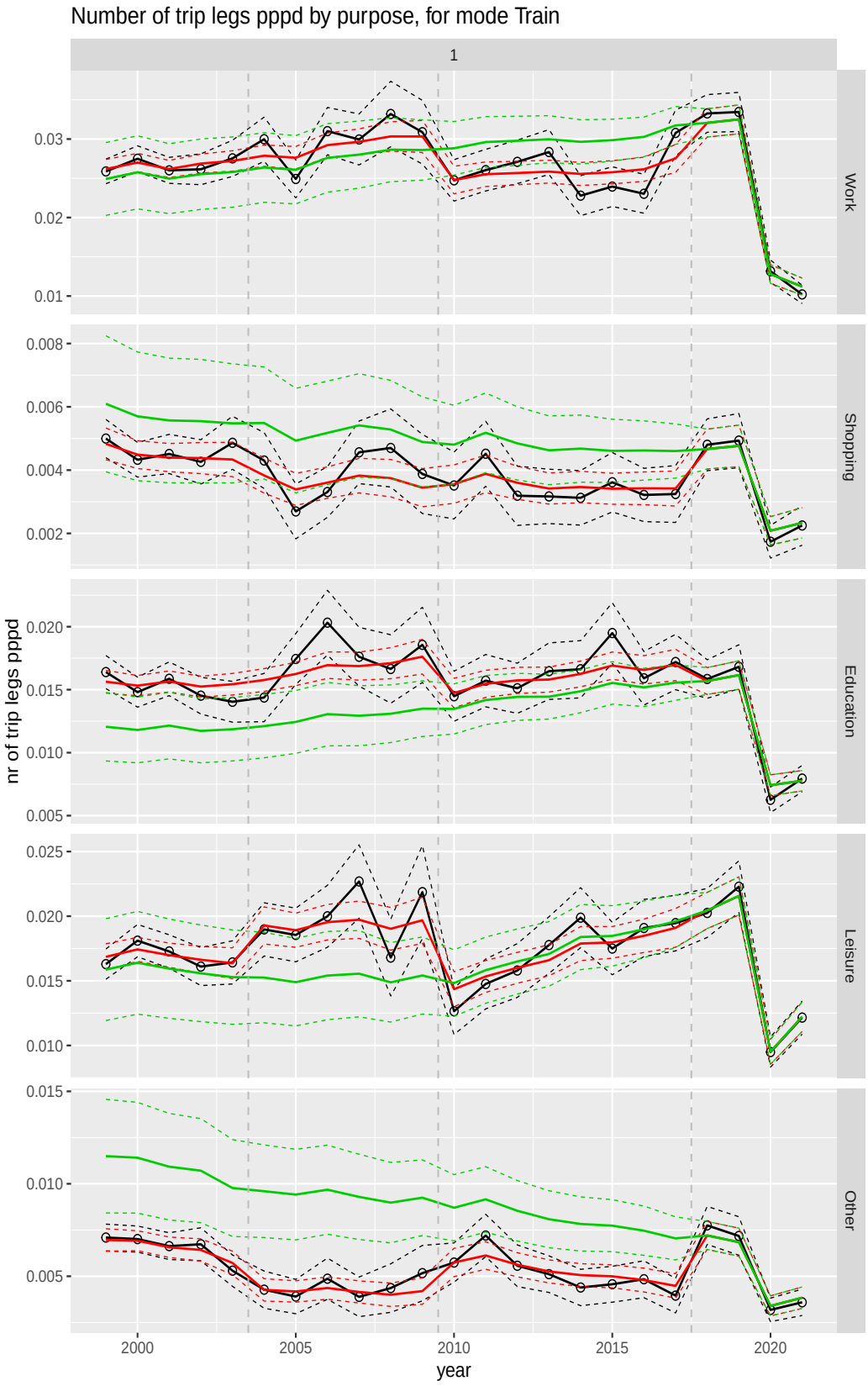

Figure A.13 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

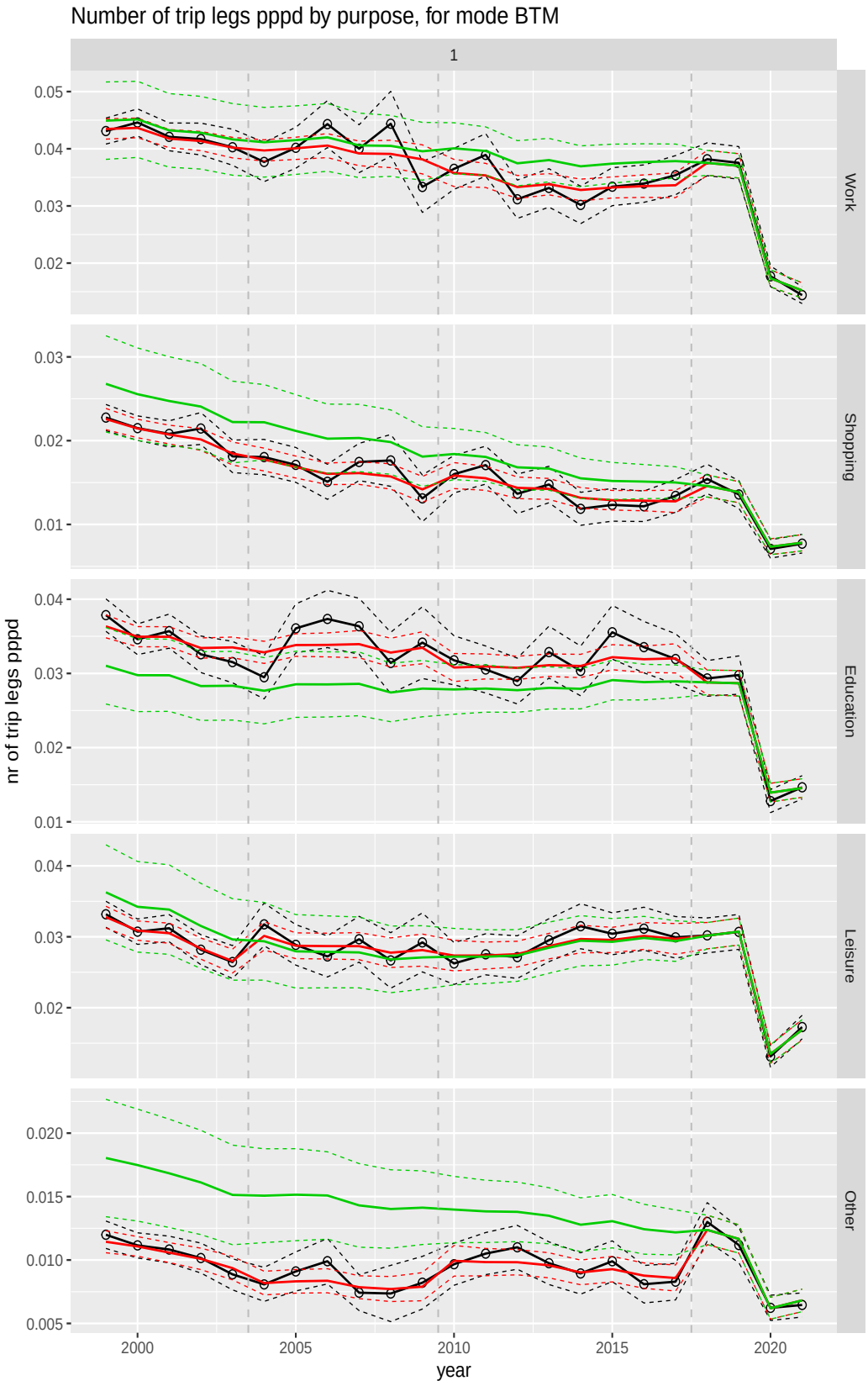

Figure A.14 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

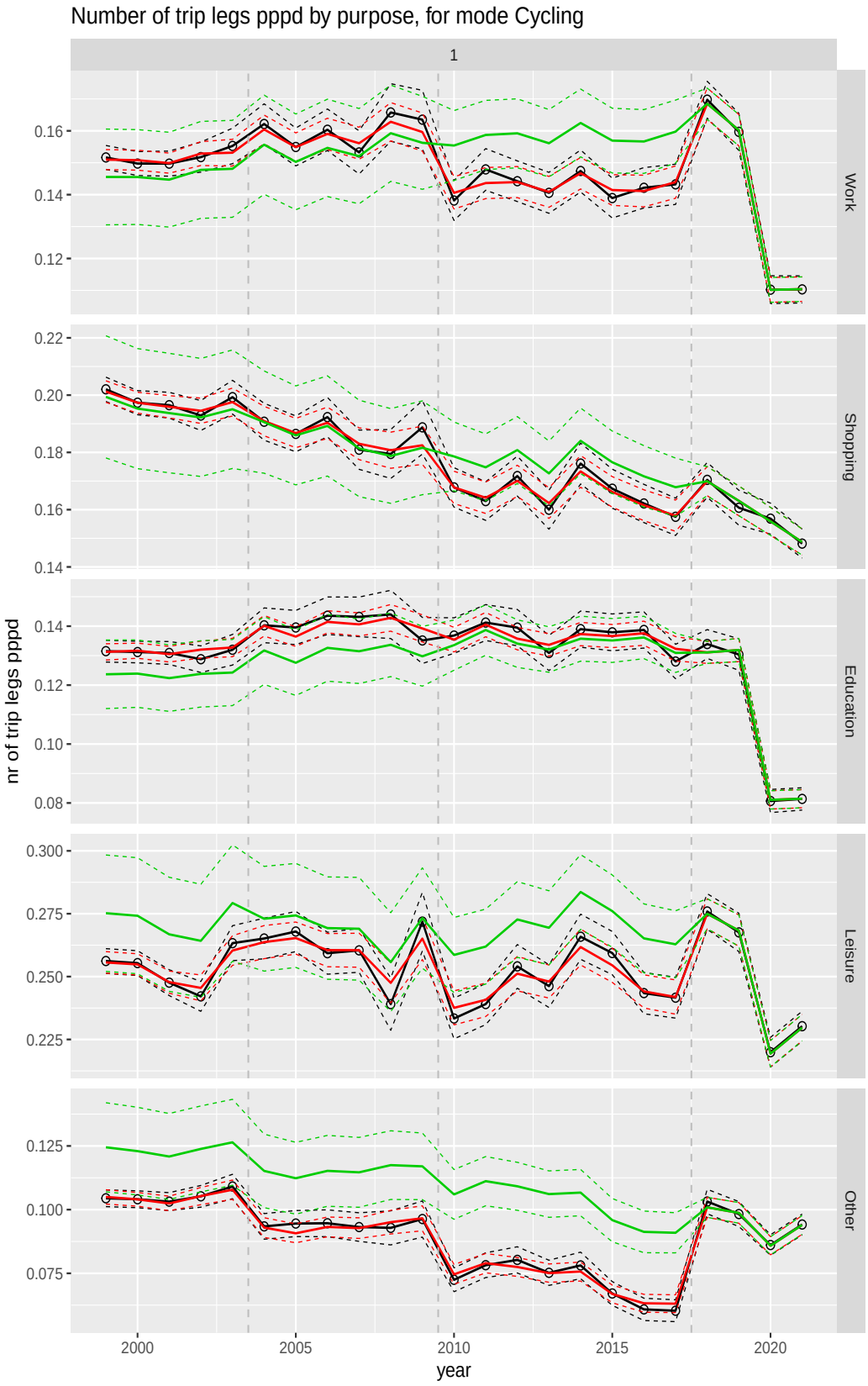

Figure A.15 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

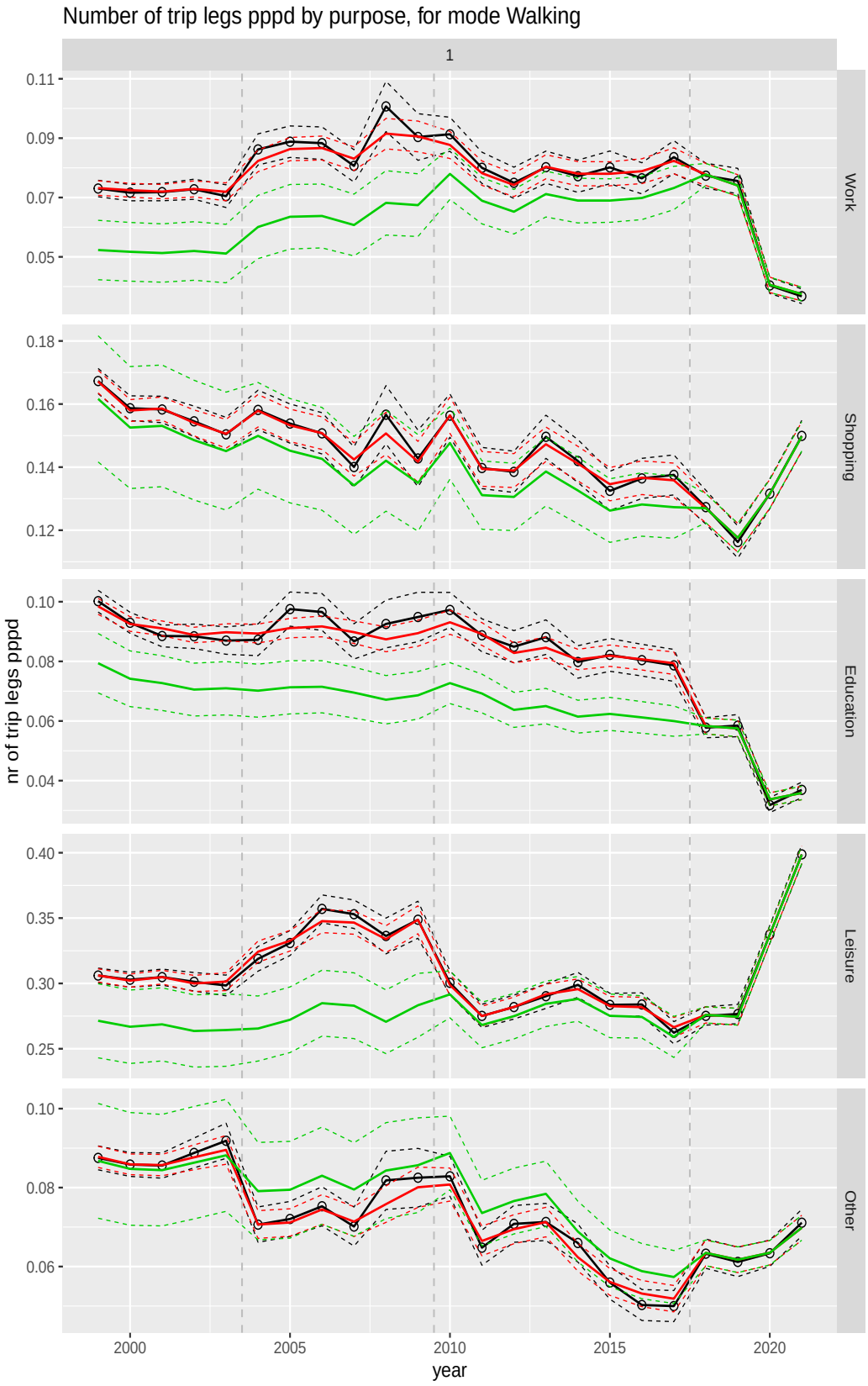

Figure A.16 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

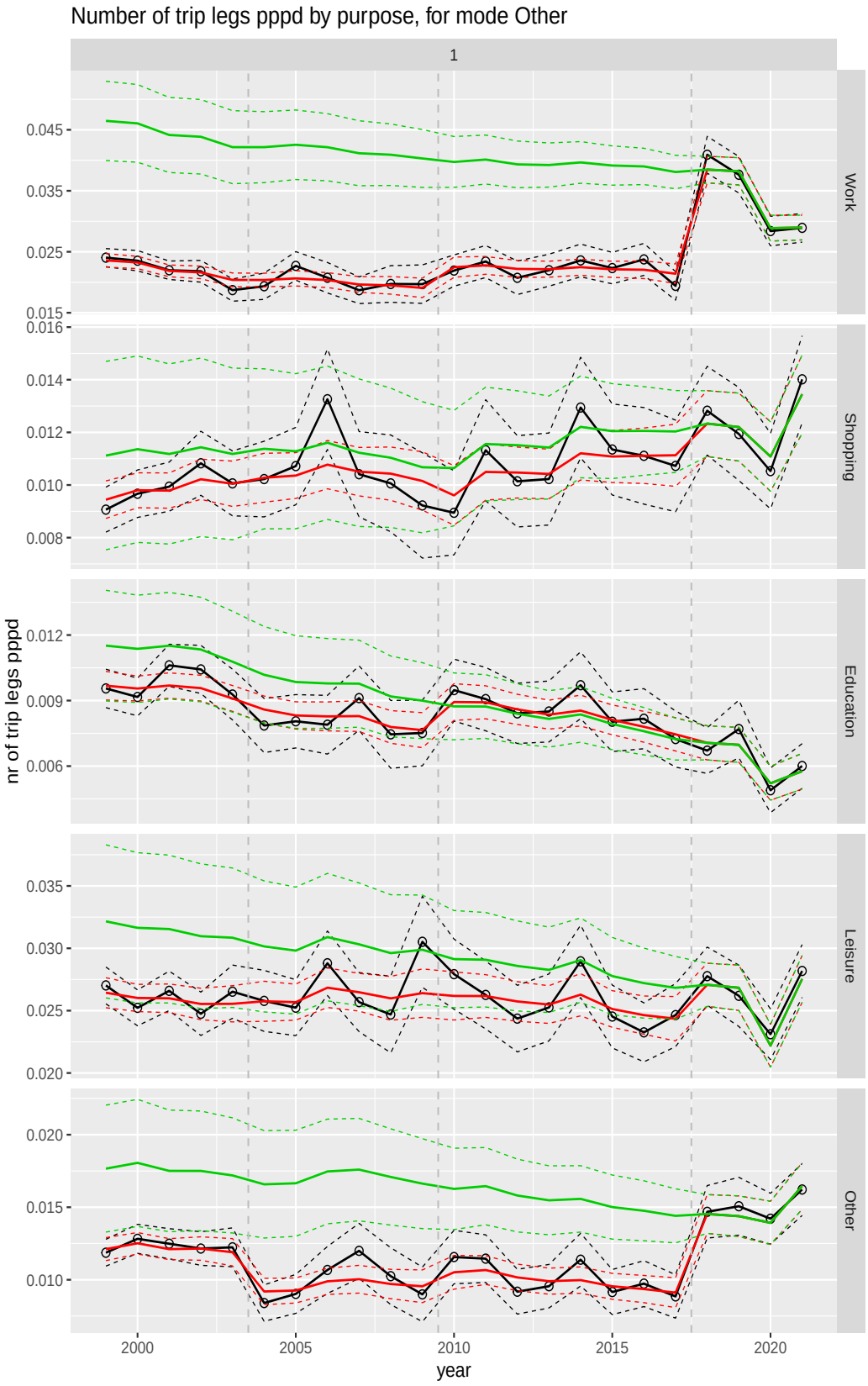

Figure A.17 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

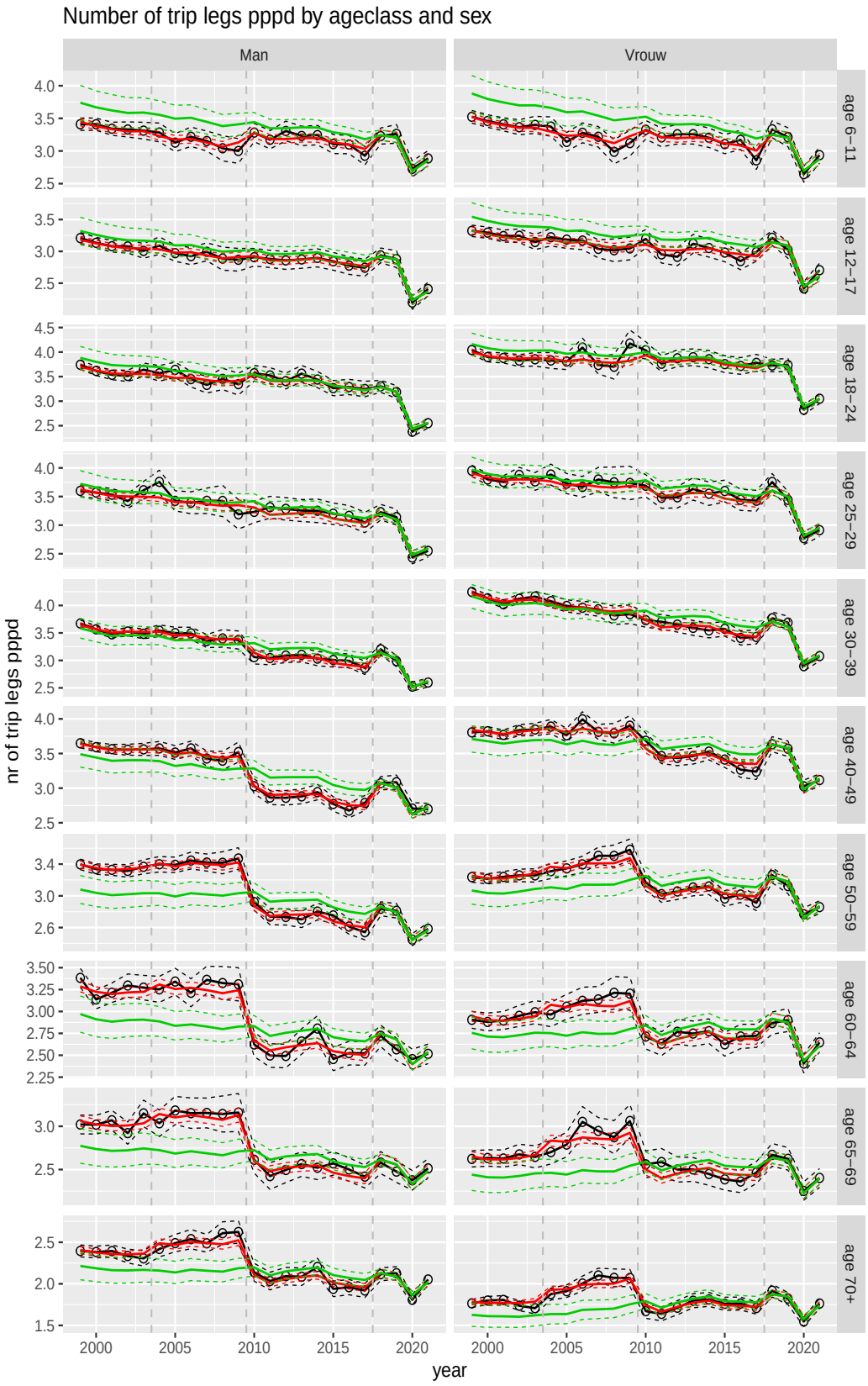

Figure A.18 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

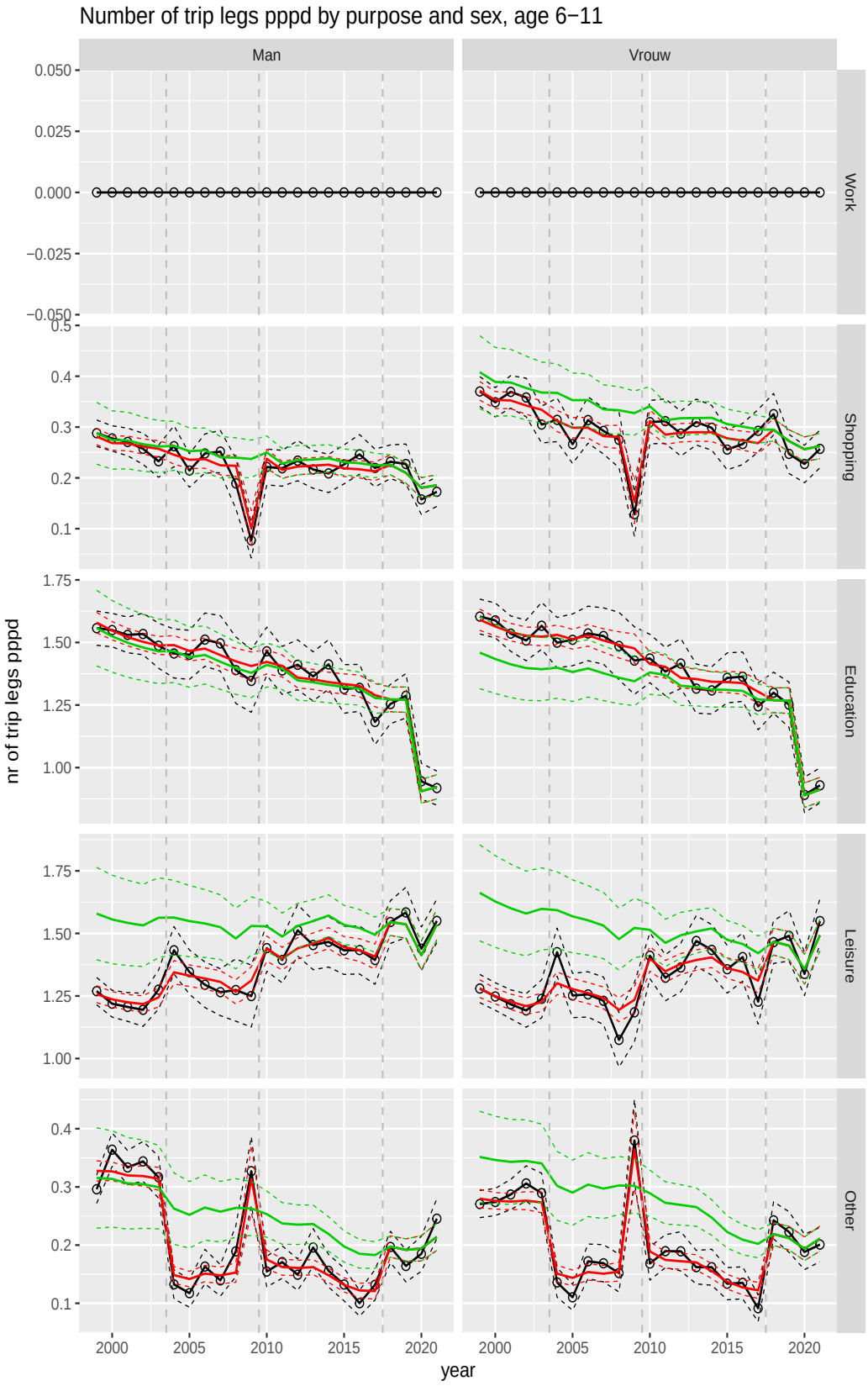

Figure A.19 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.20 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.21 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

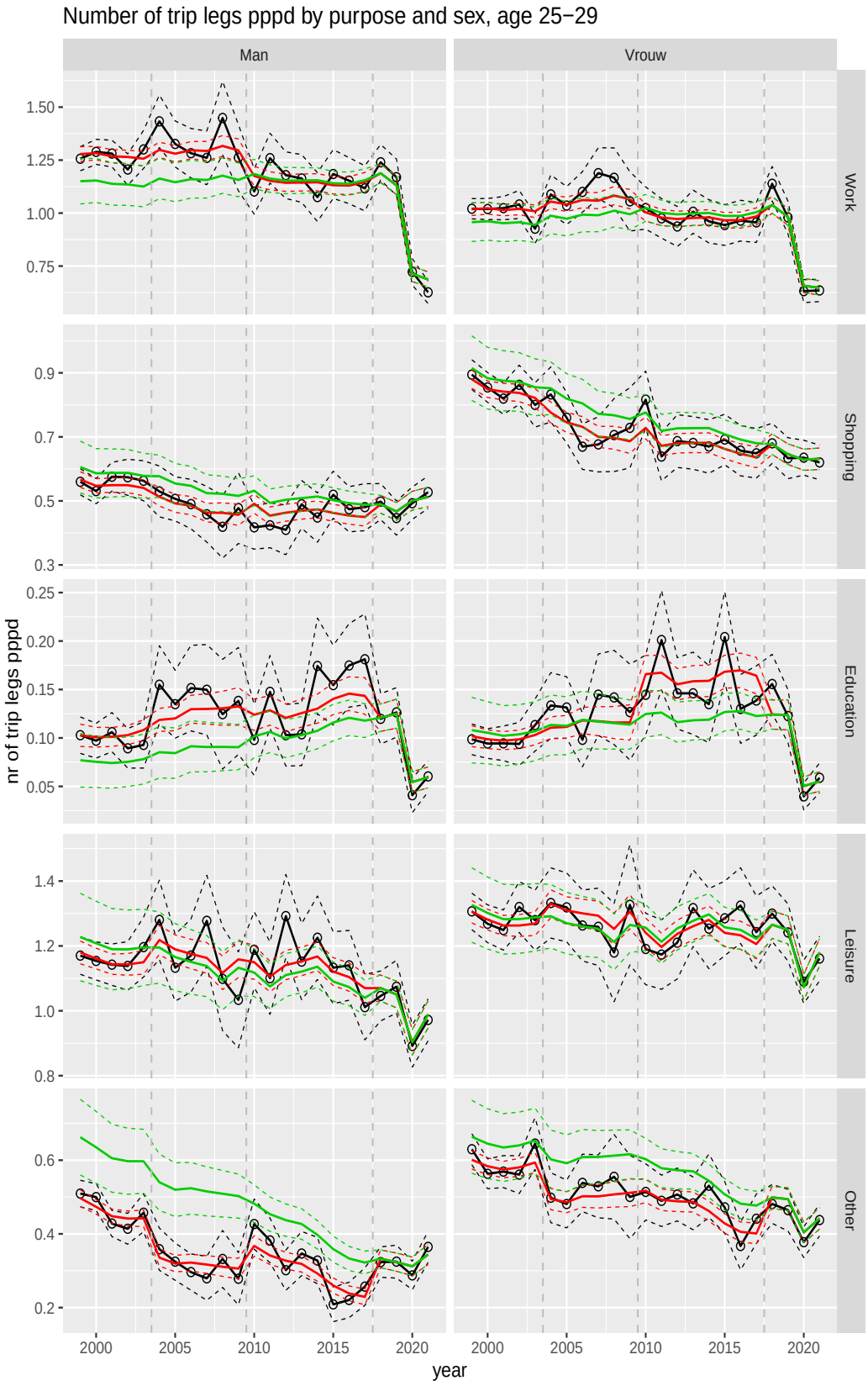

Figure A.22 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

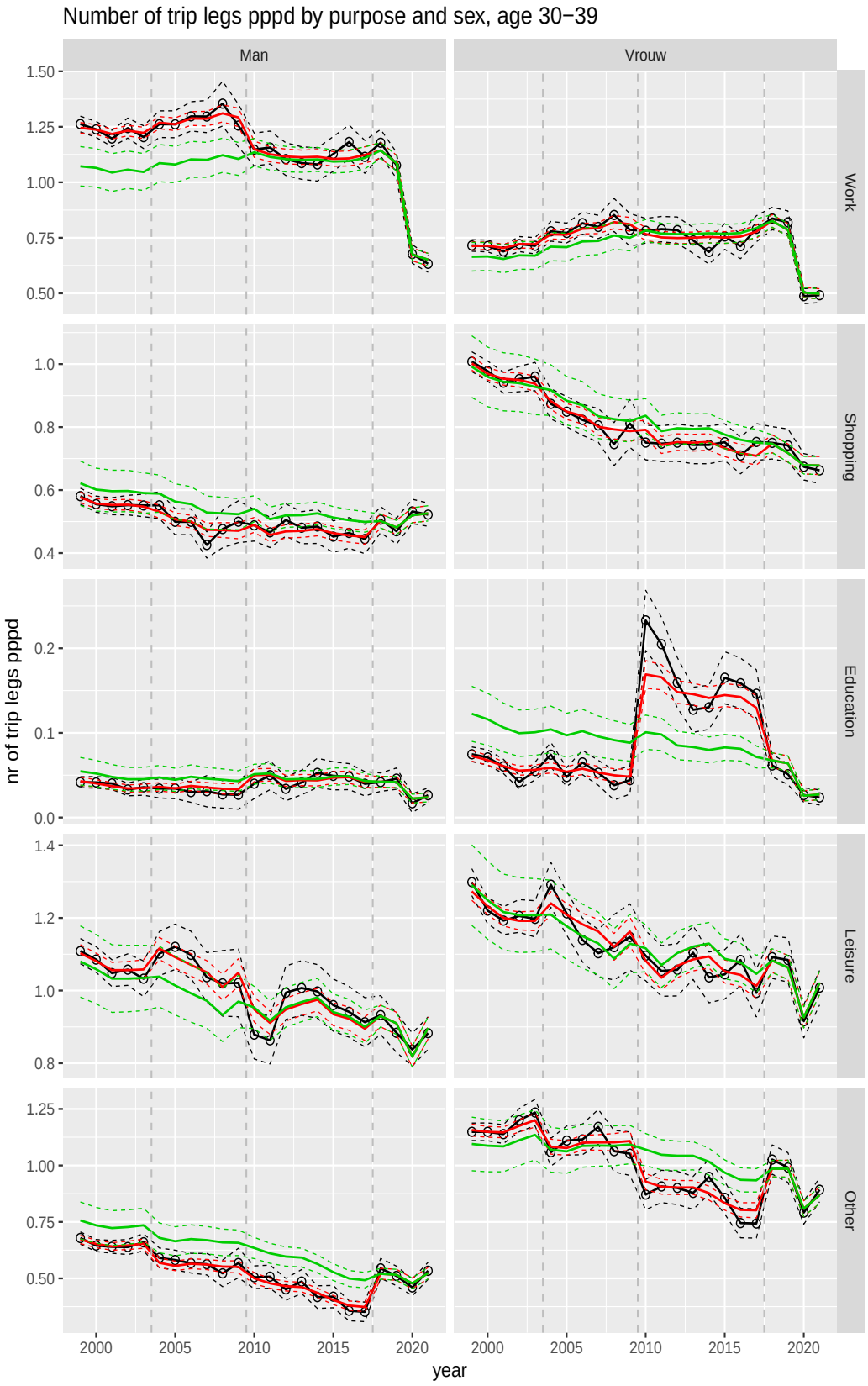

Figure A.23 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.24 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

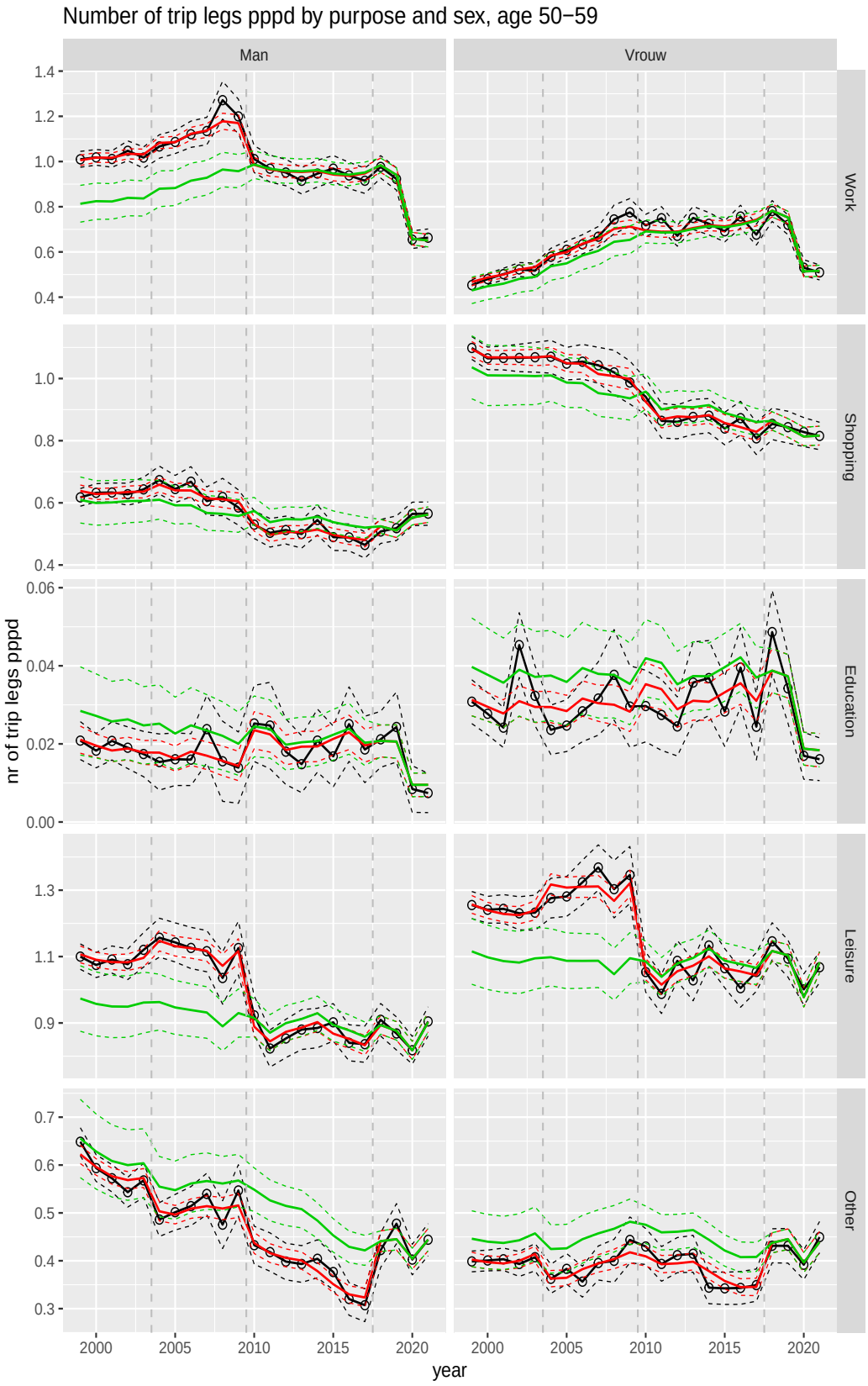

Figure A.25 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.26 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.27 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.28 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

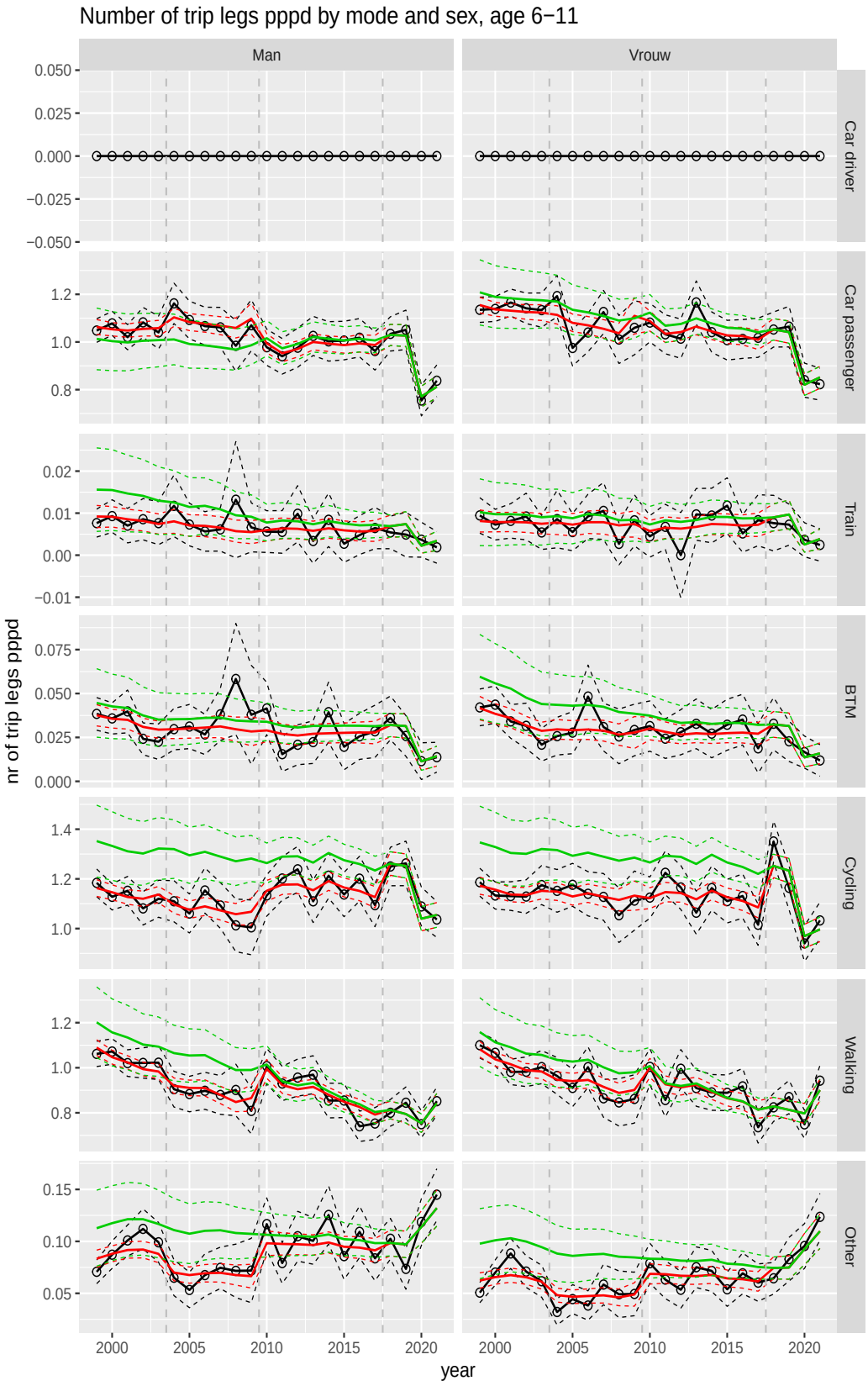

Figure A.29 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

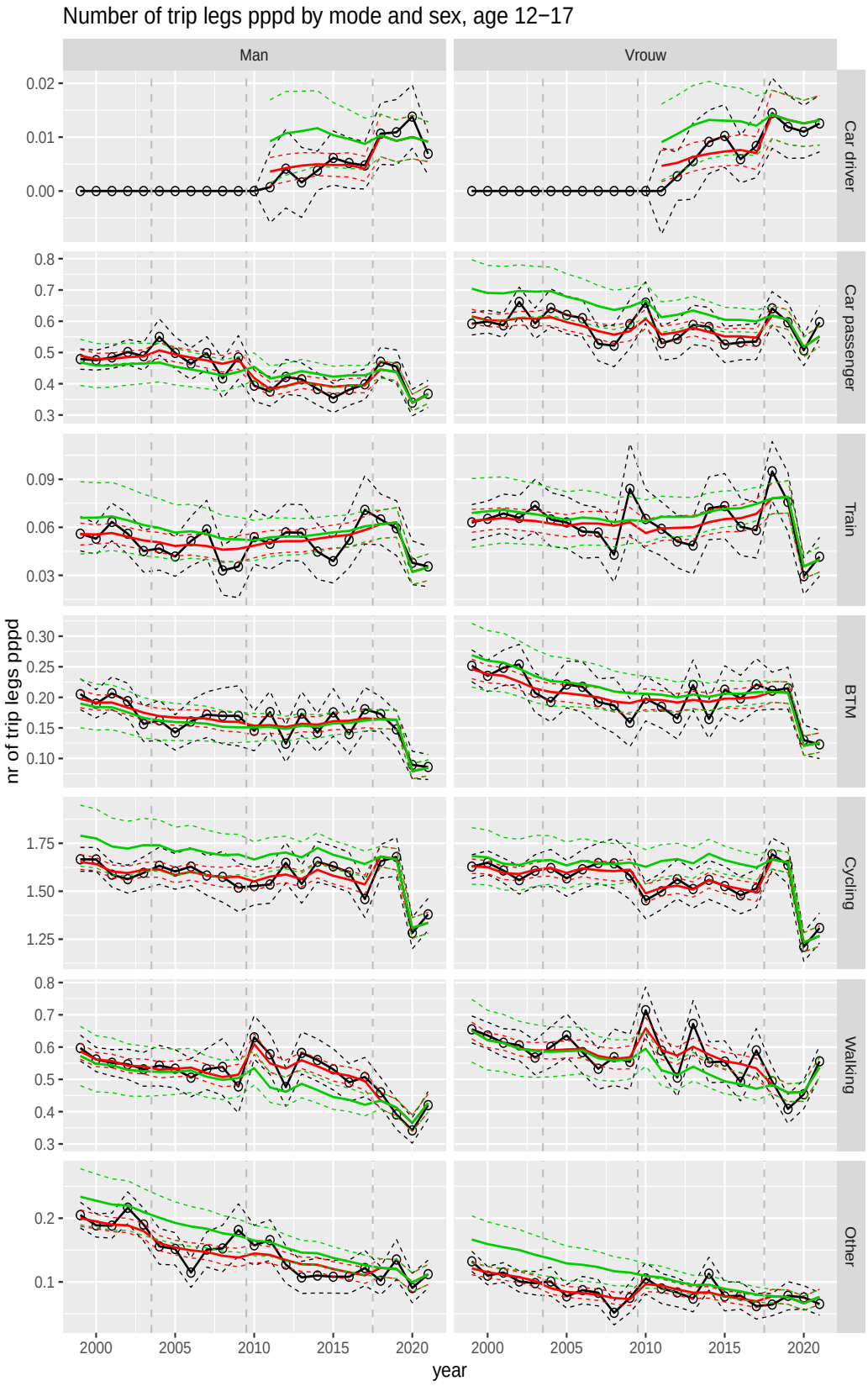

Figure A.30 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.31 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

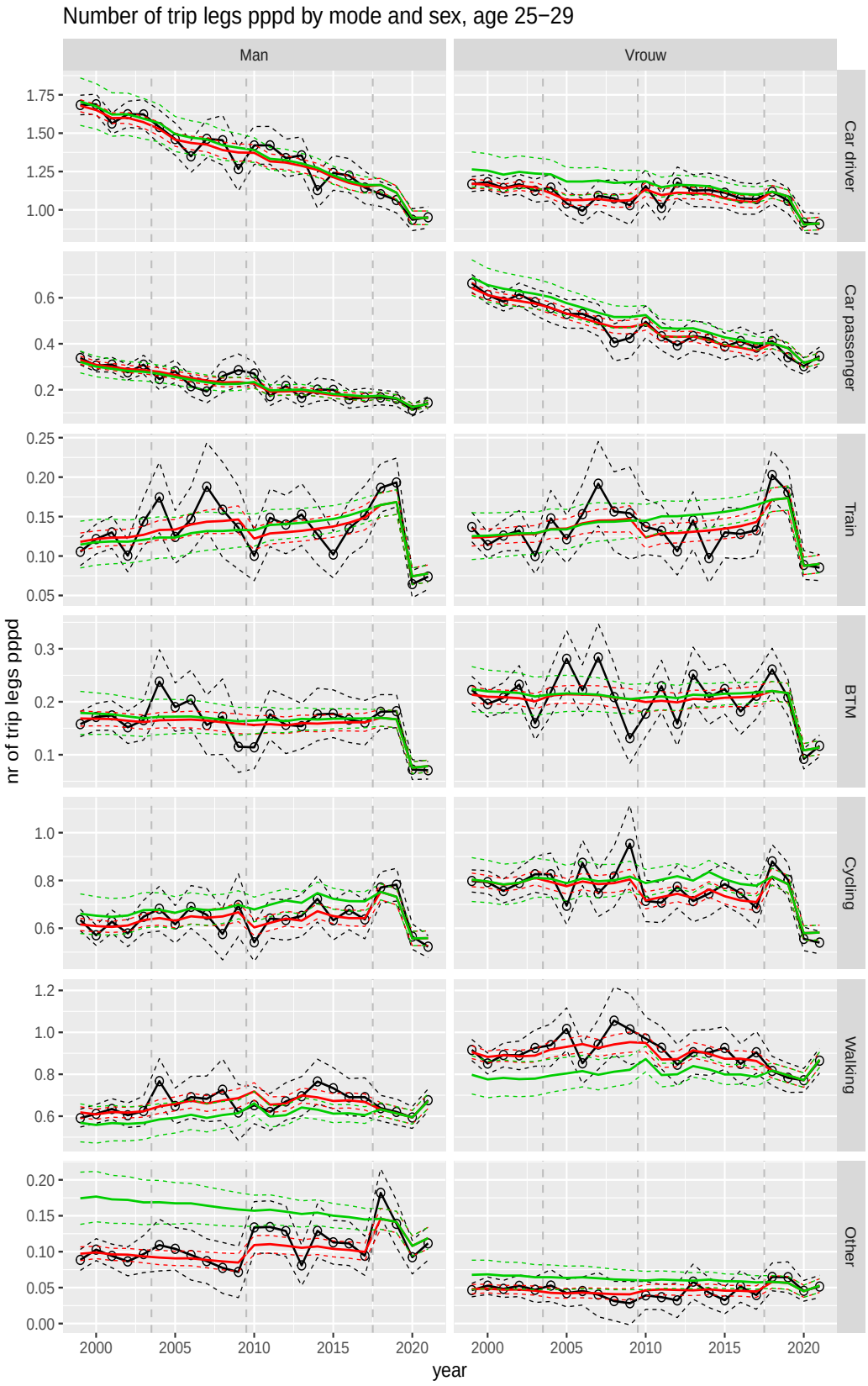

Figure A.32 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

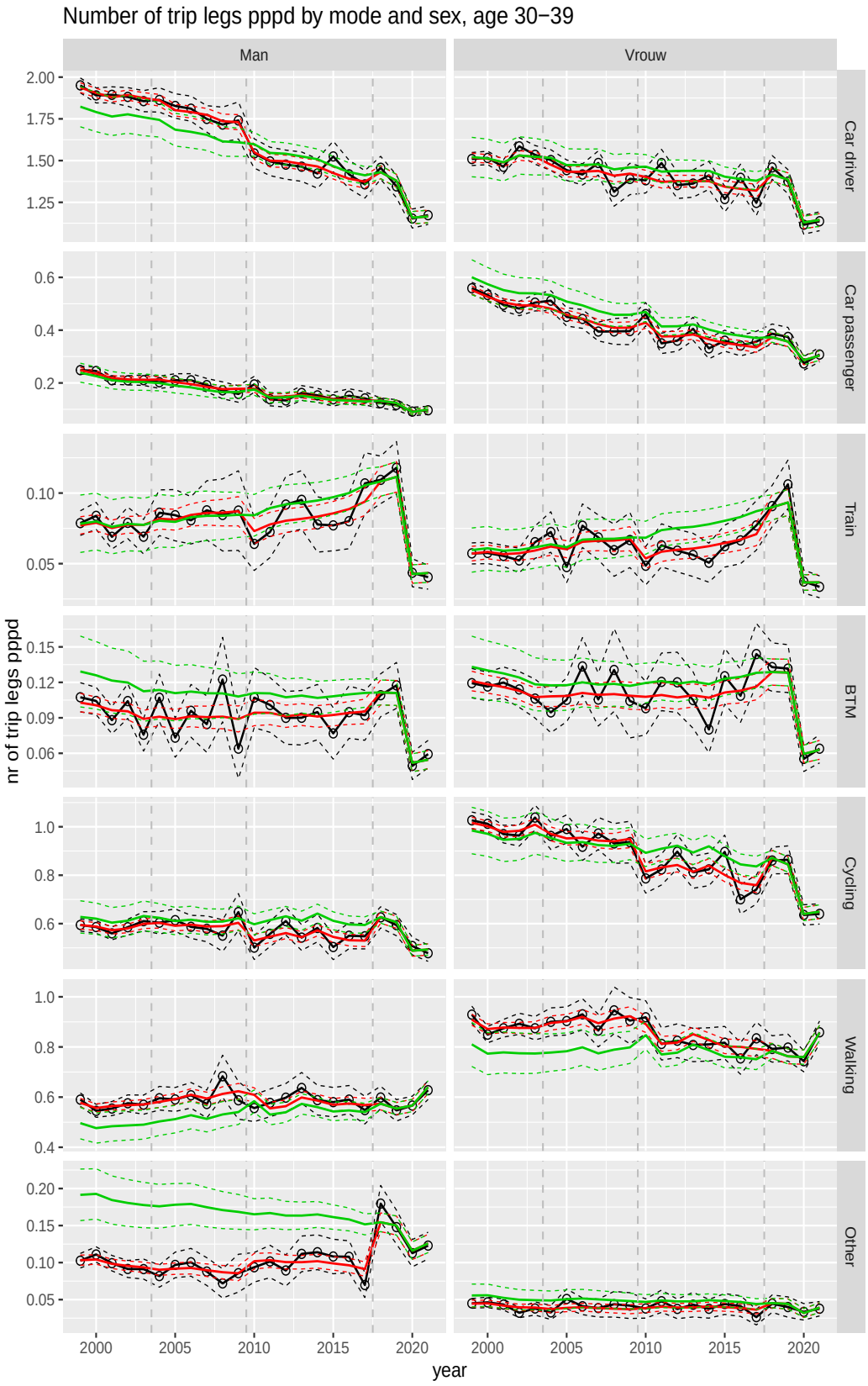

Figure A.33 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

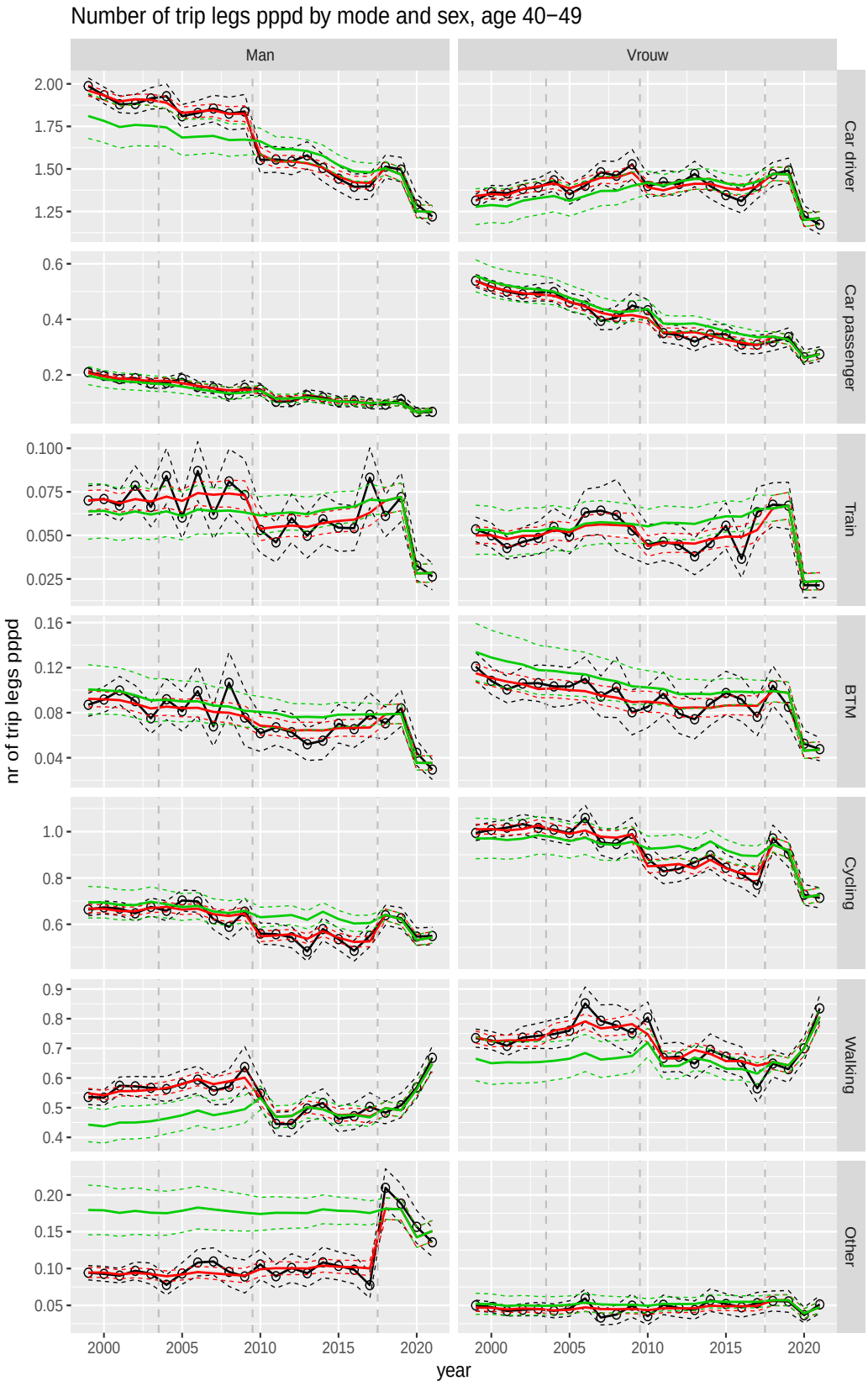

Figure A.34 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

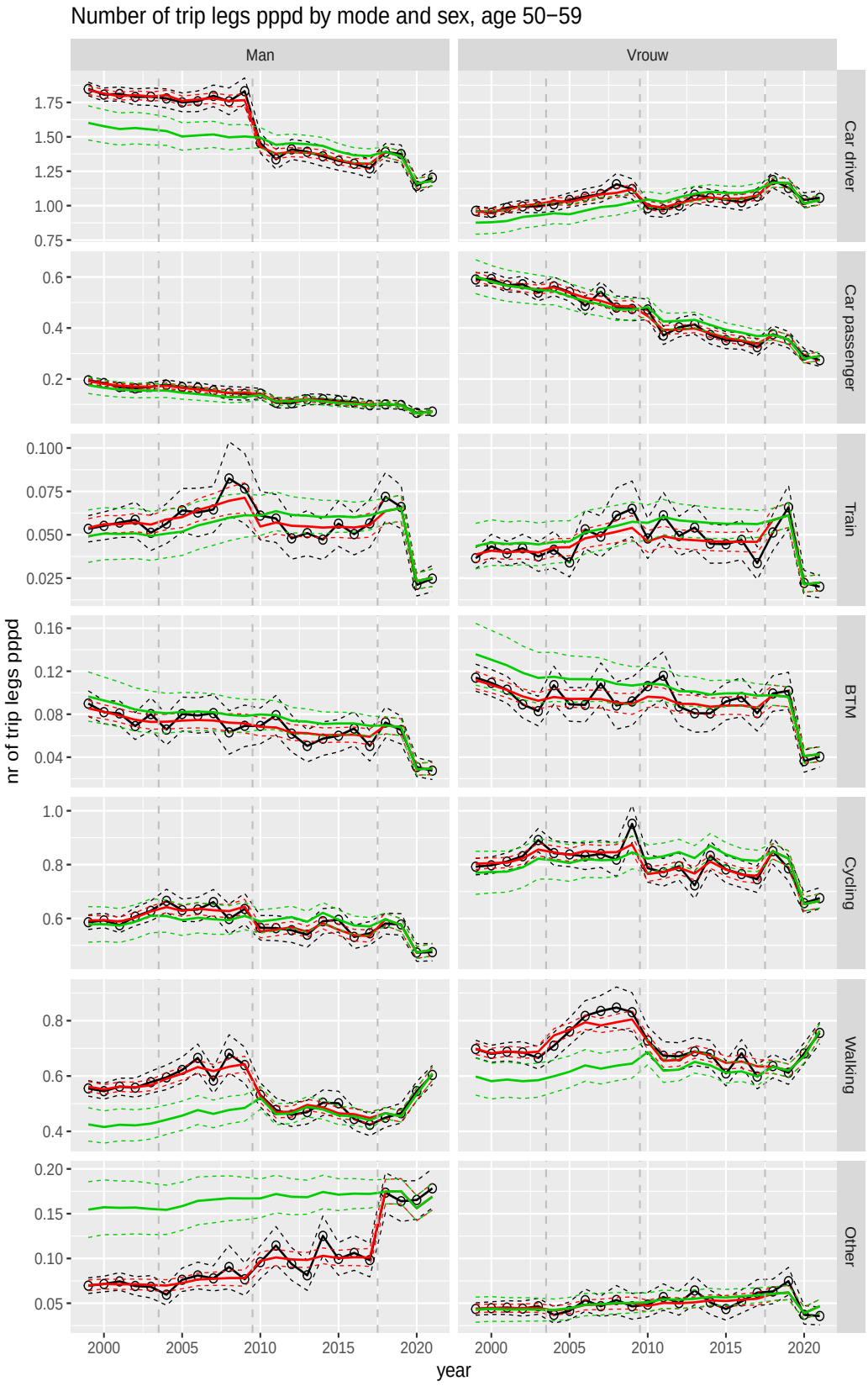

Figure A.35 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

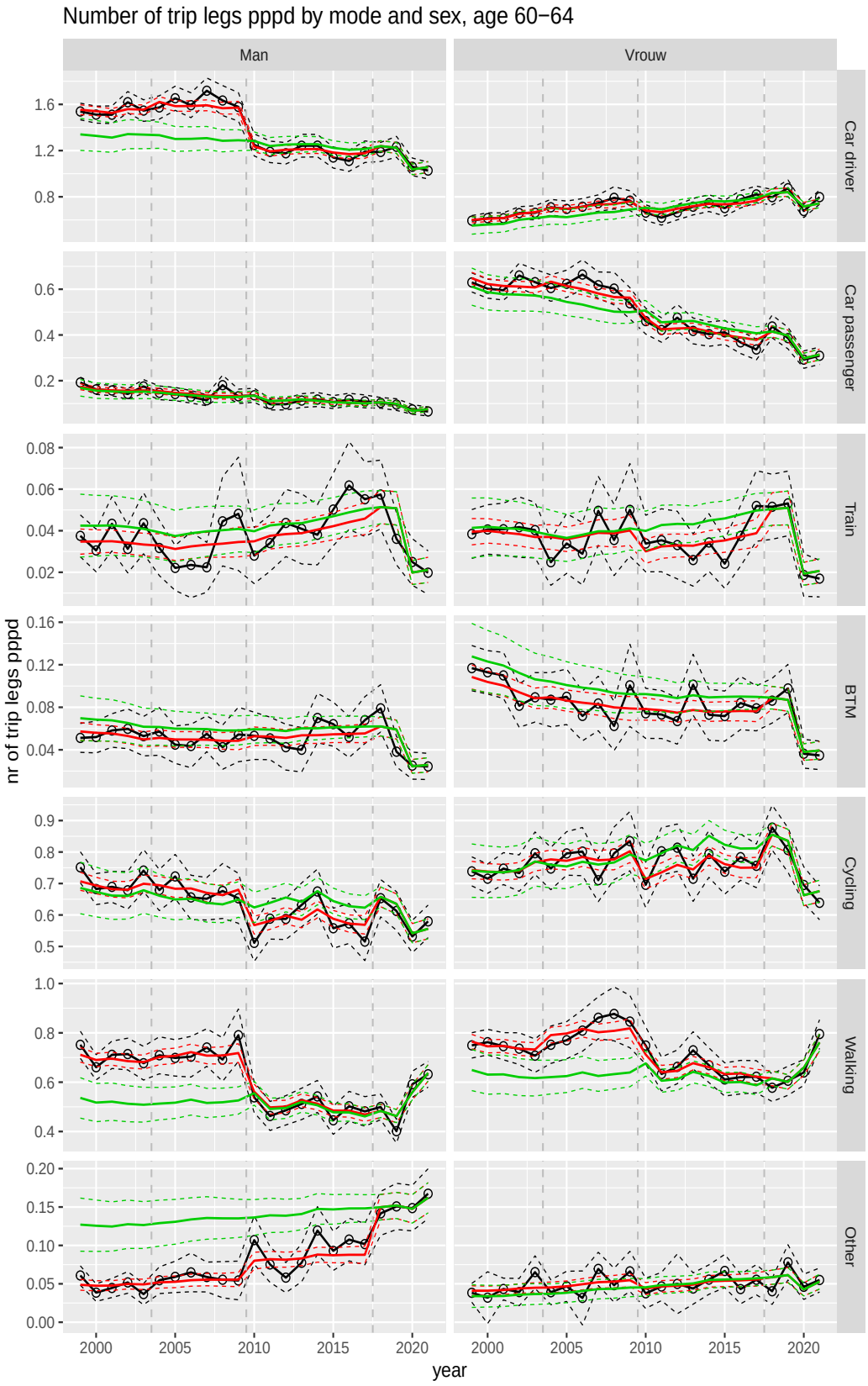

Figure A.36 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

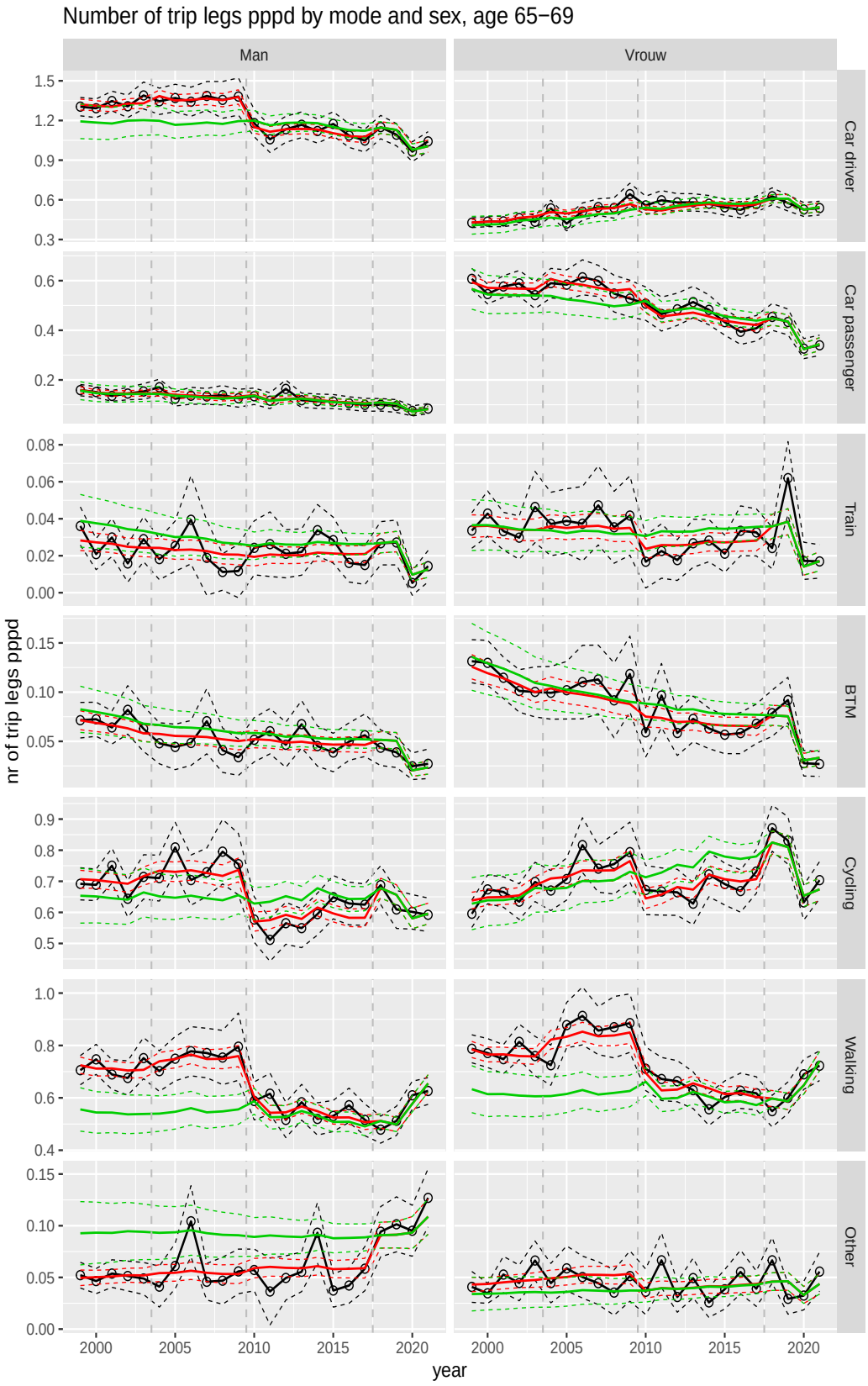

Figure A.37 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.38 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

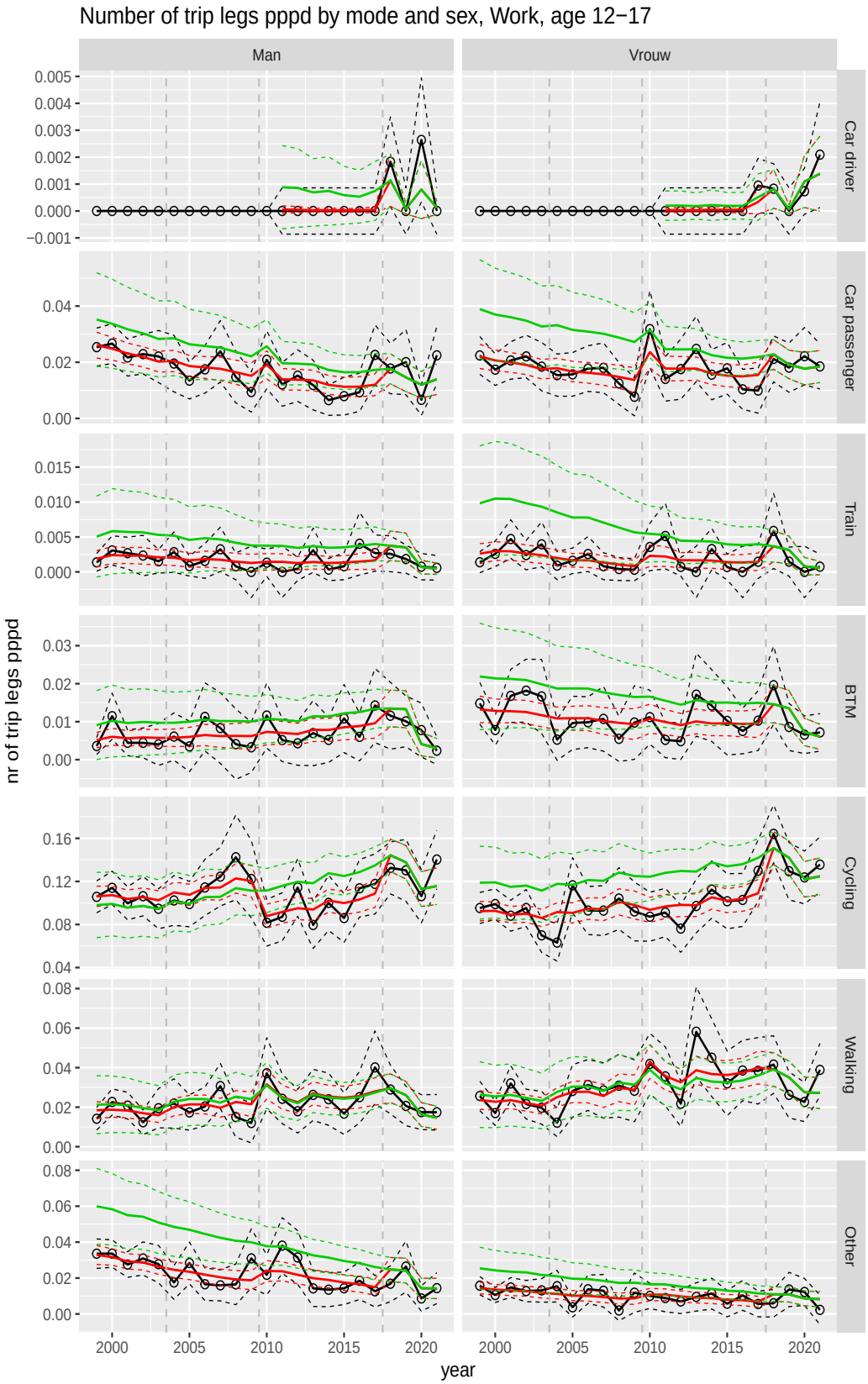

Figure A.39 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

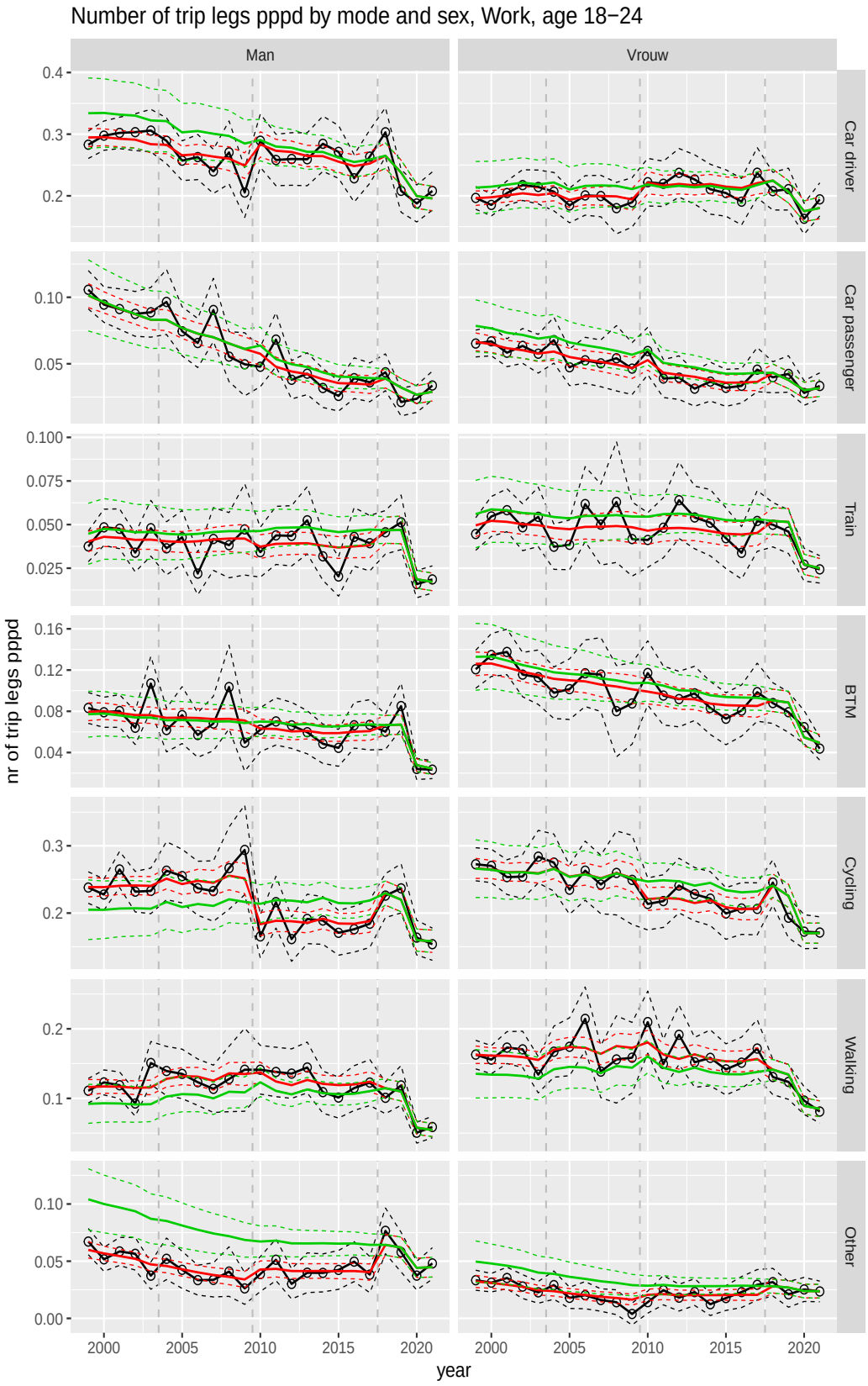

Figure A.40 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

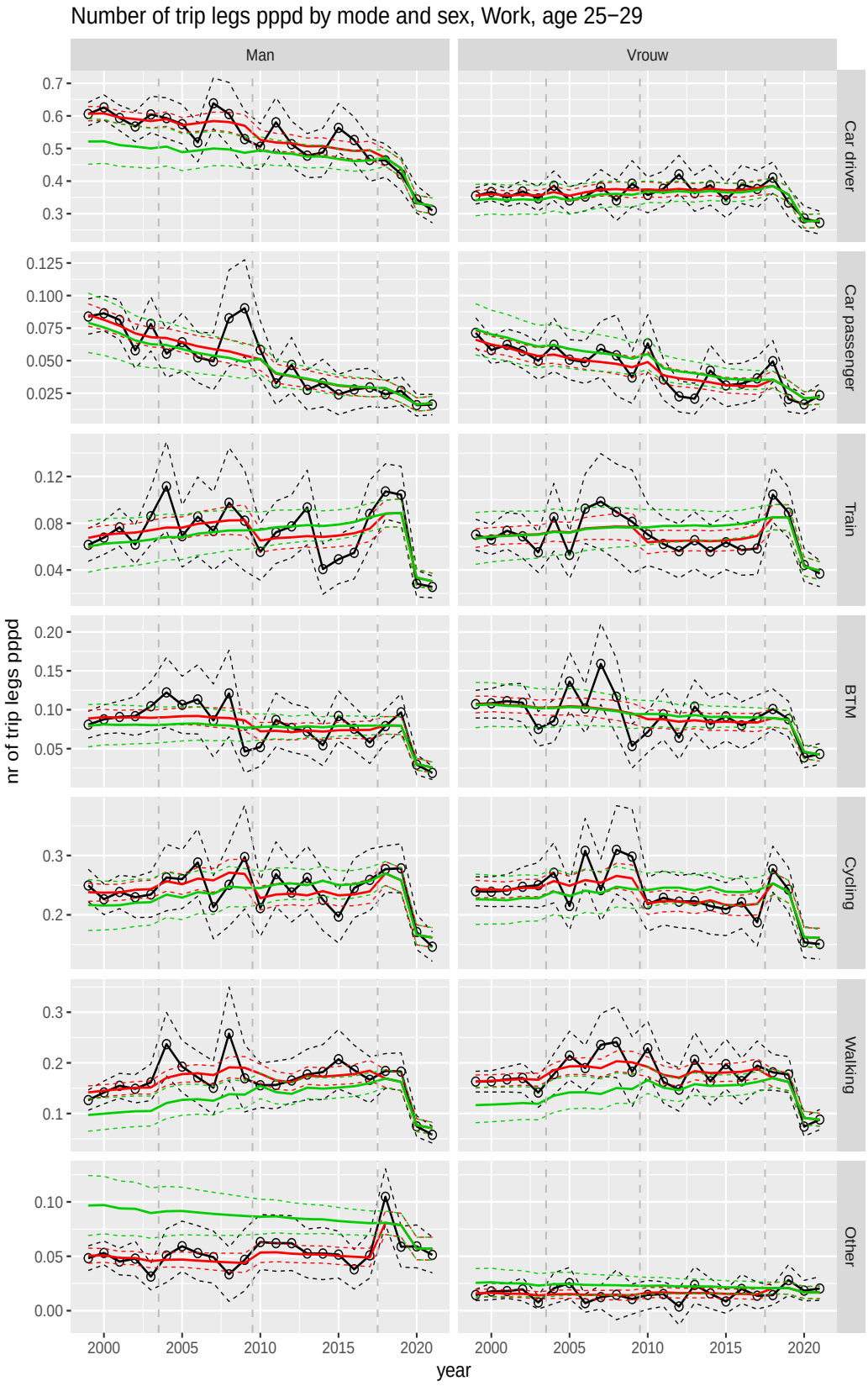

Figure A.41 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

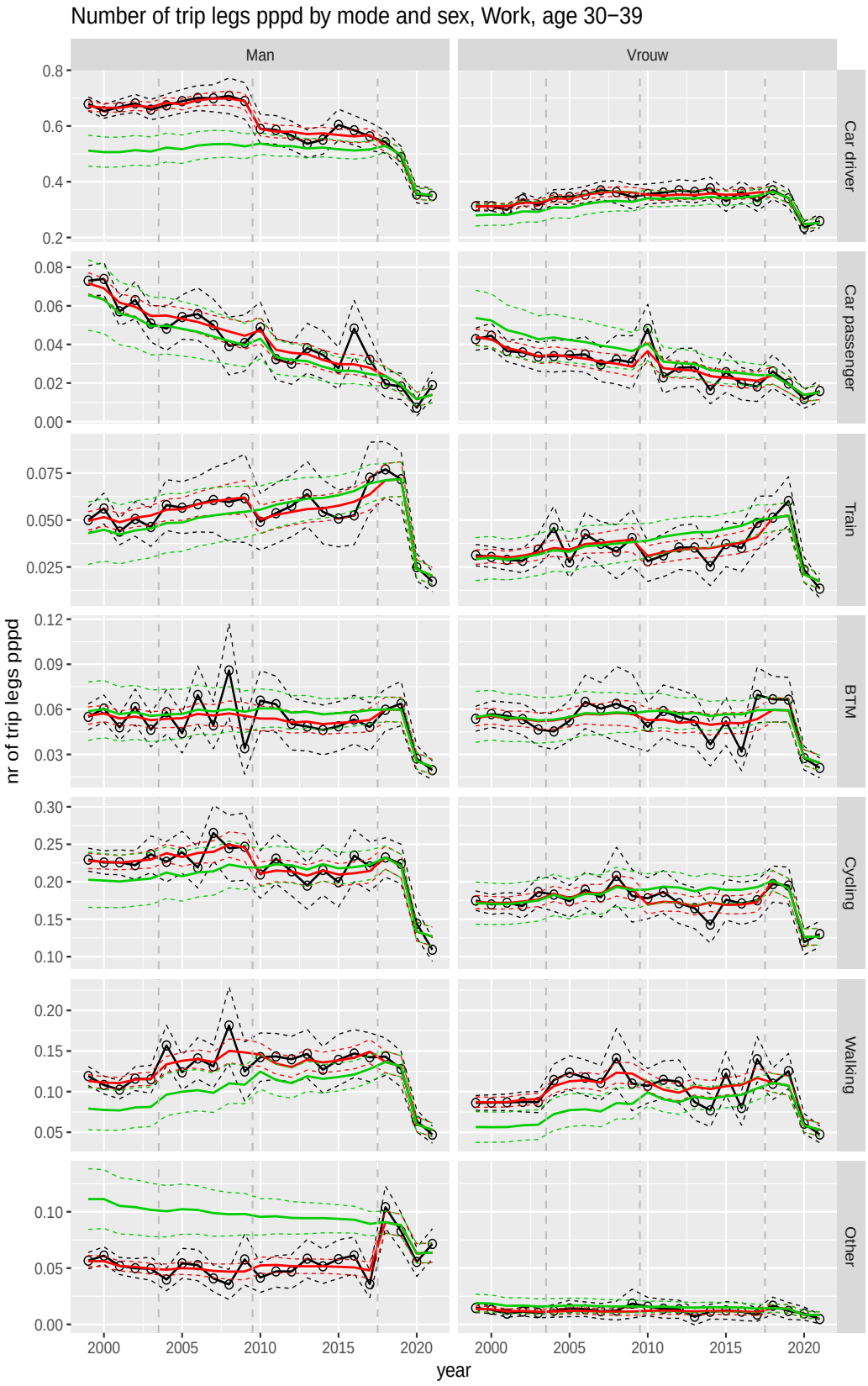

Figure A.42 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

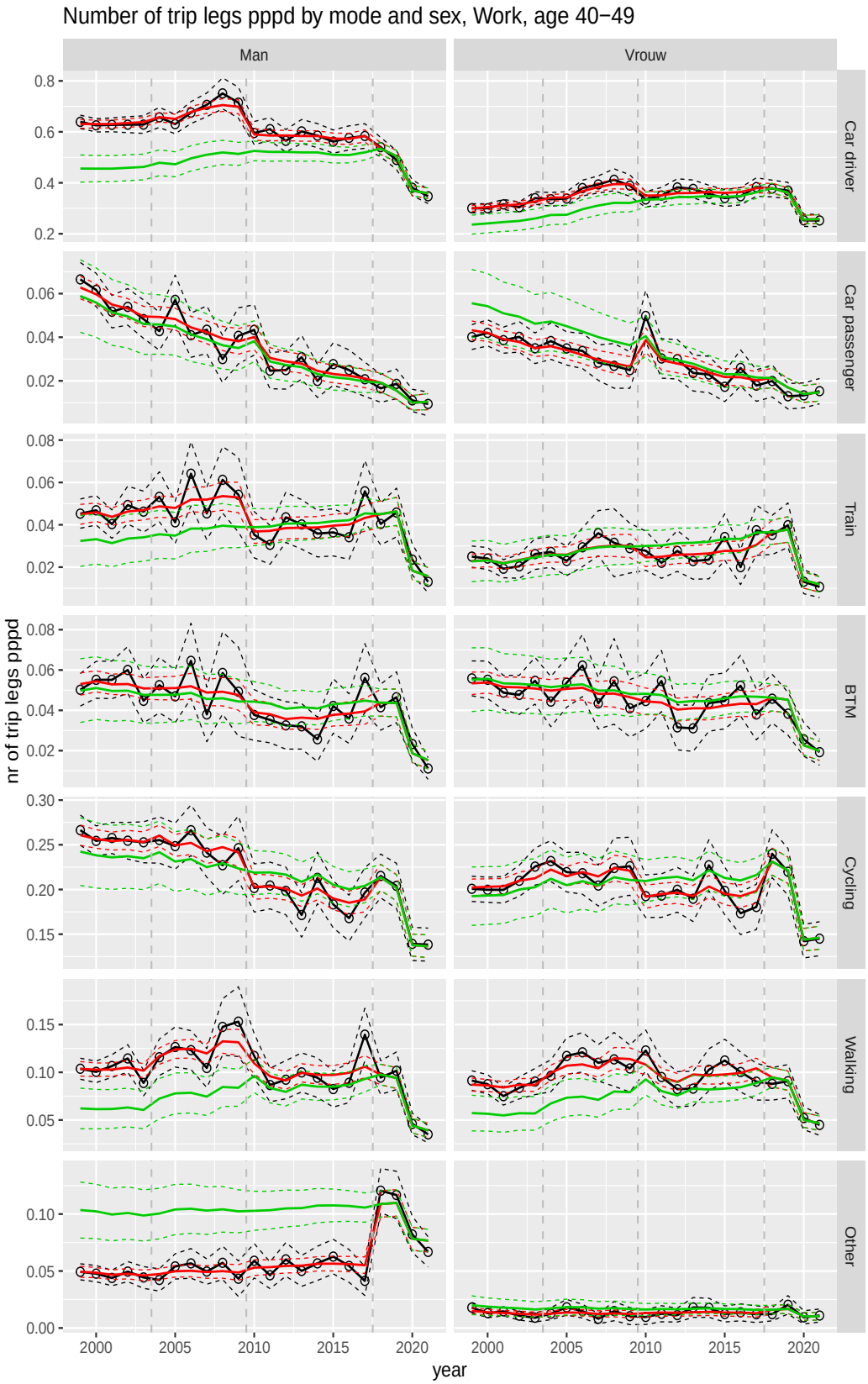

Figure A.43 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

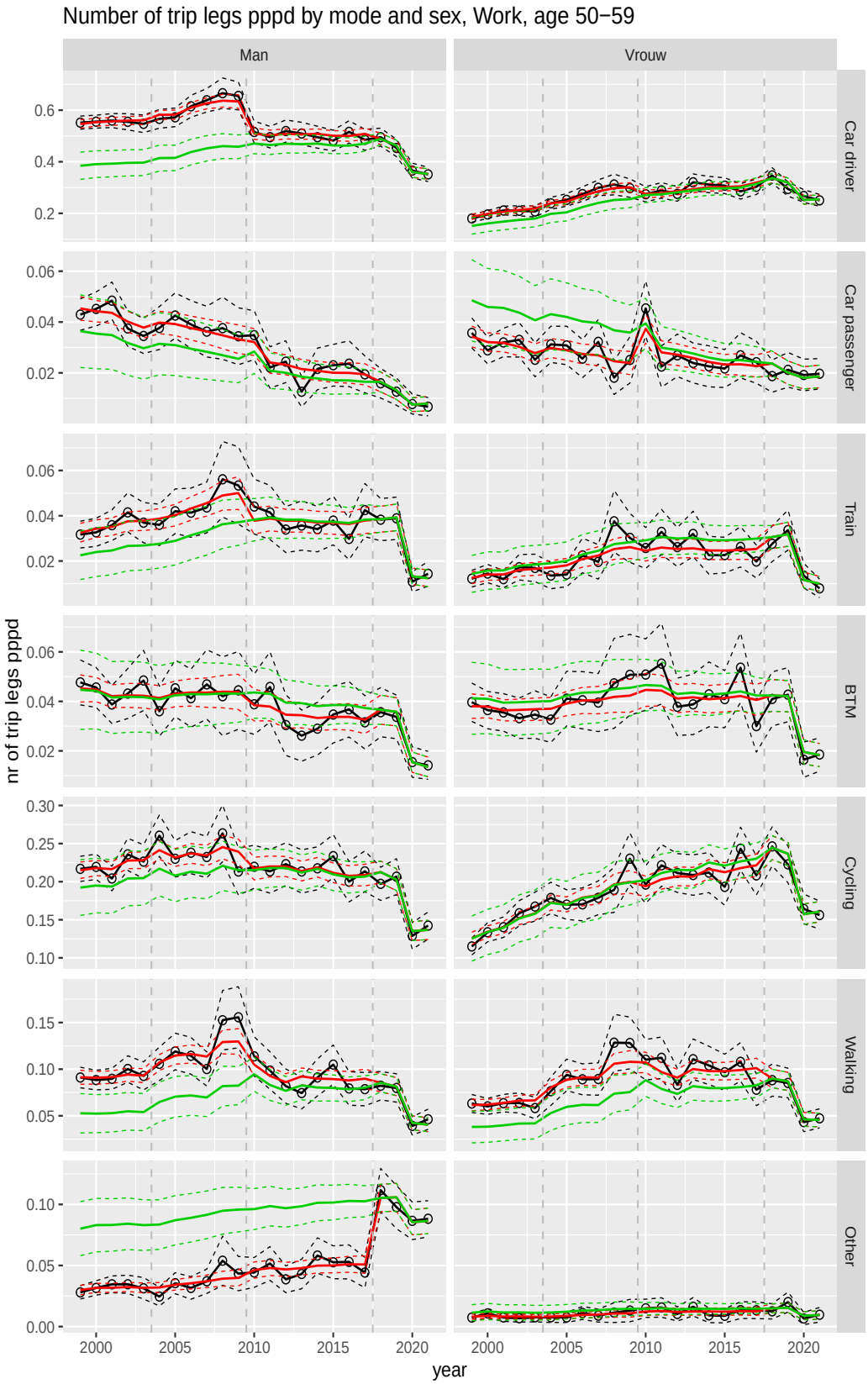

Figure A.44 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

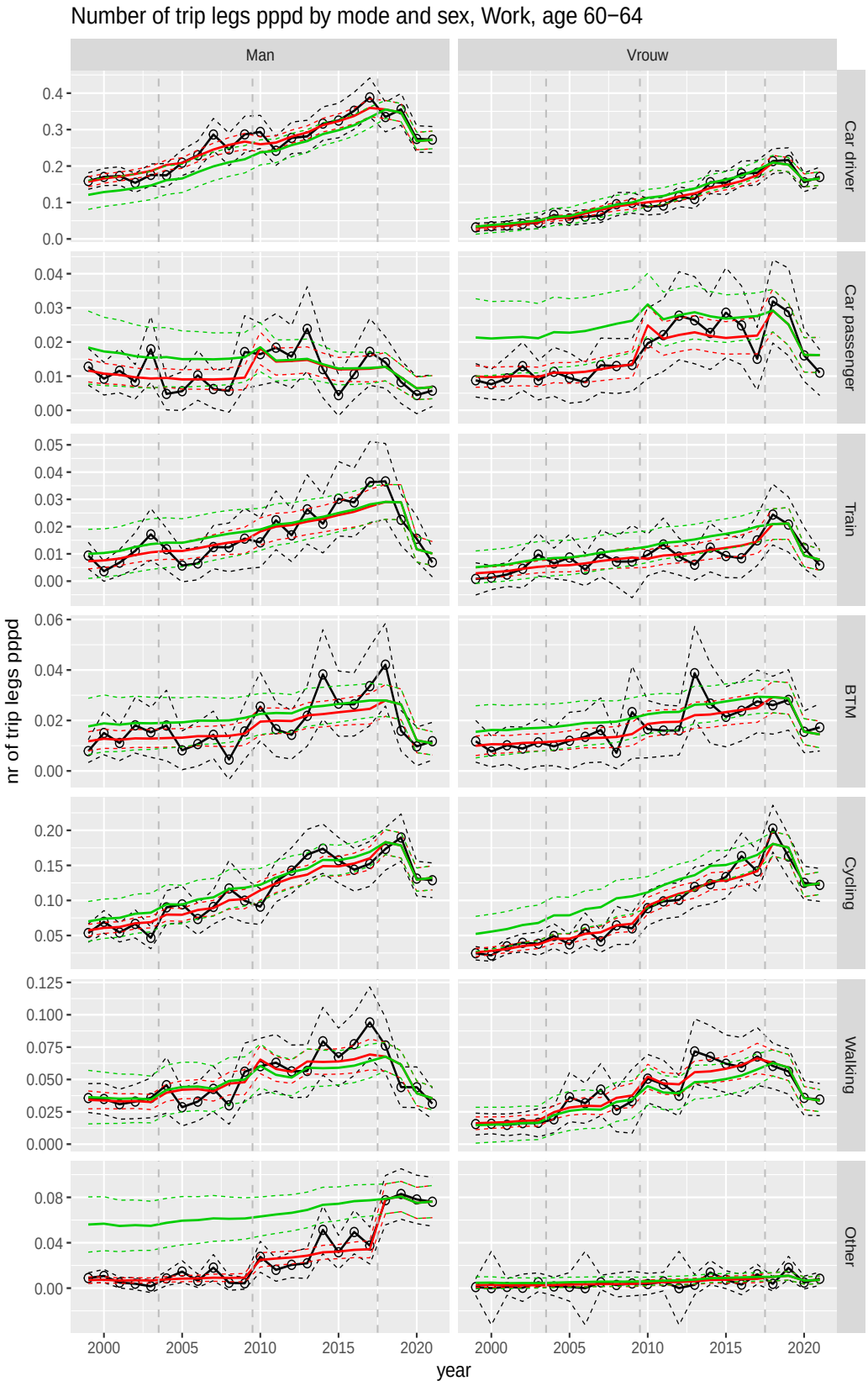

Figure A.45 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

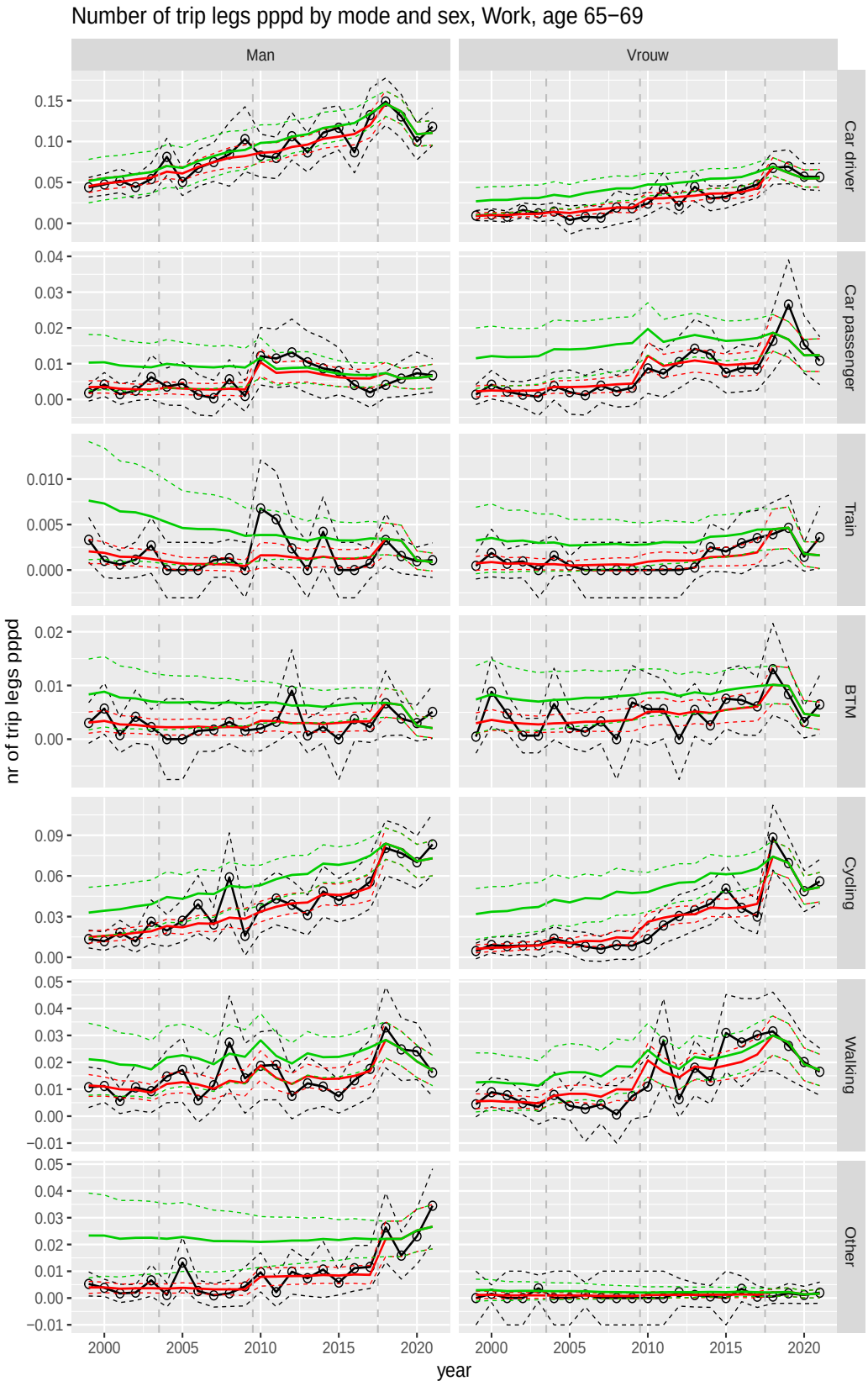

Figure A.46 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

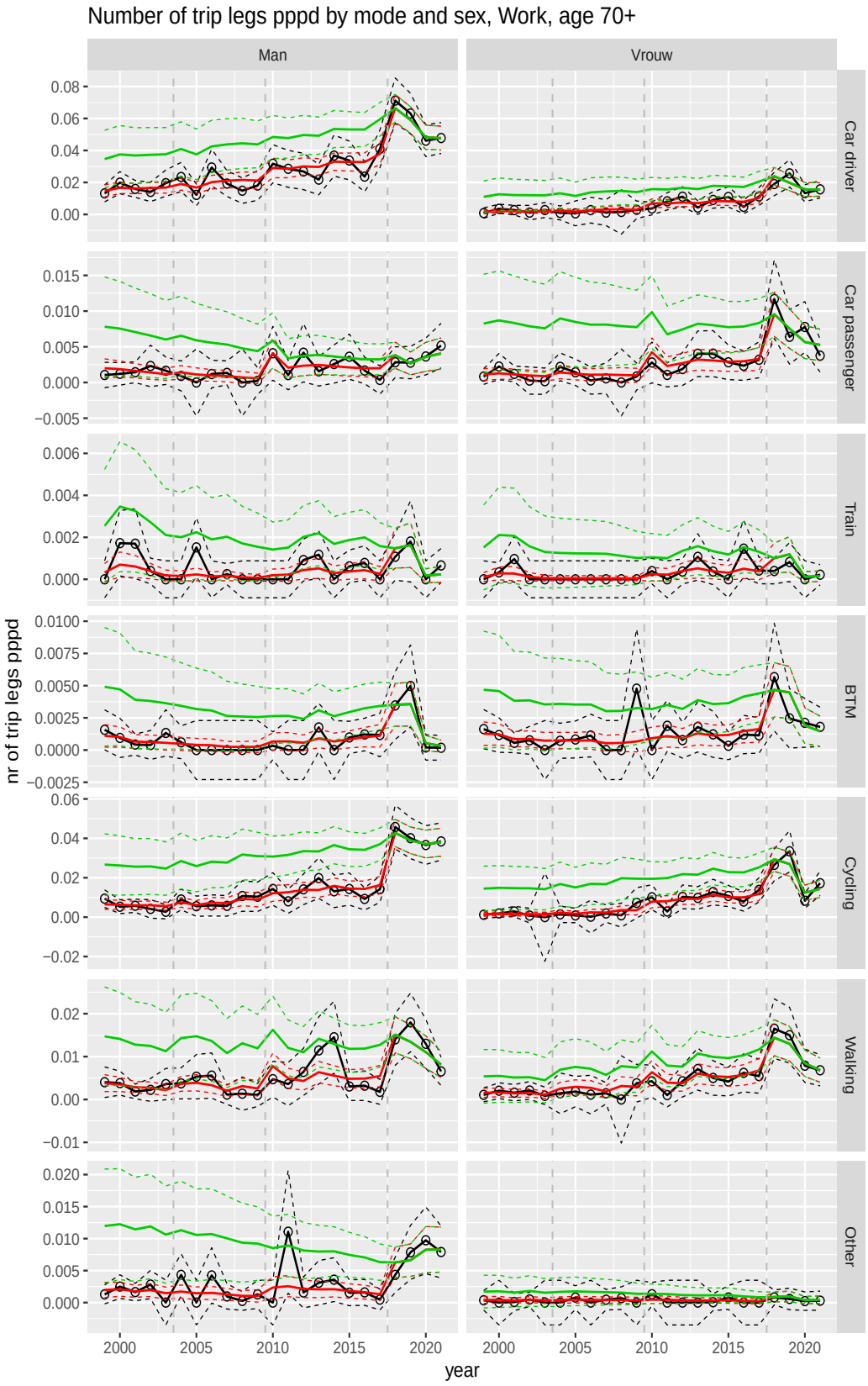

Figure A.47 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

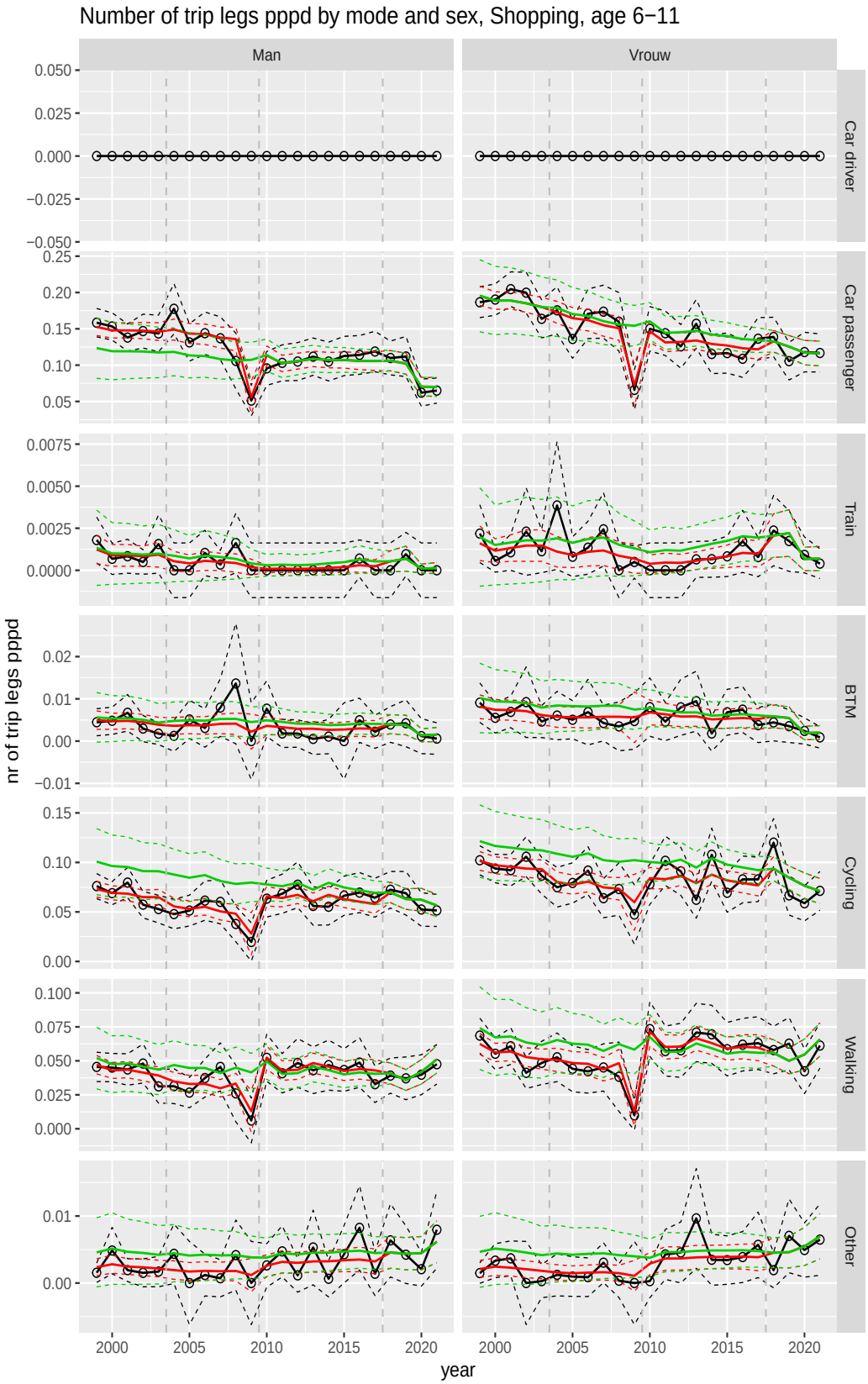

Figure A.48 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

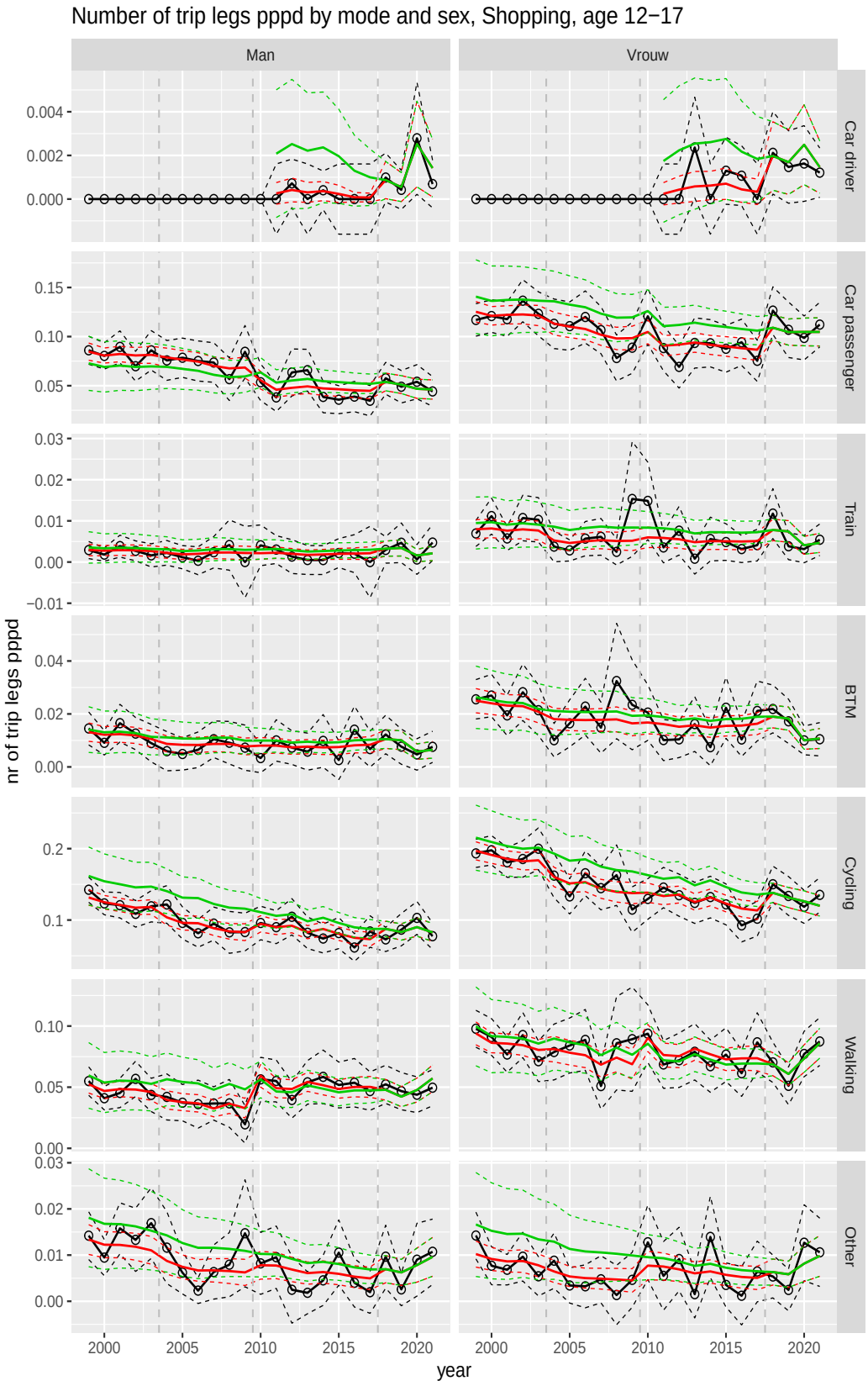

Figure A.49 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

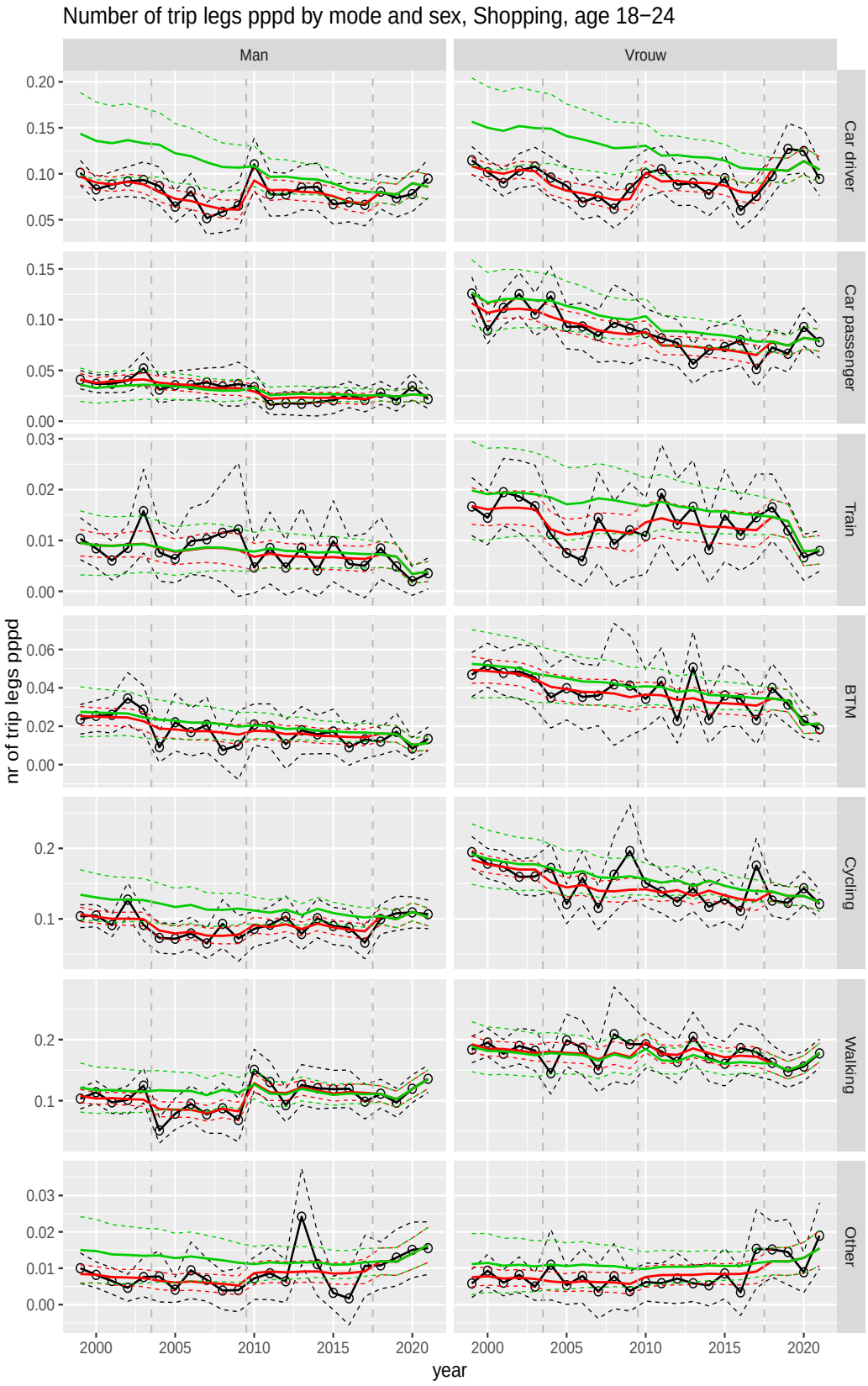

Figure A.50 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

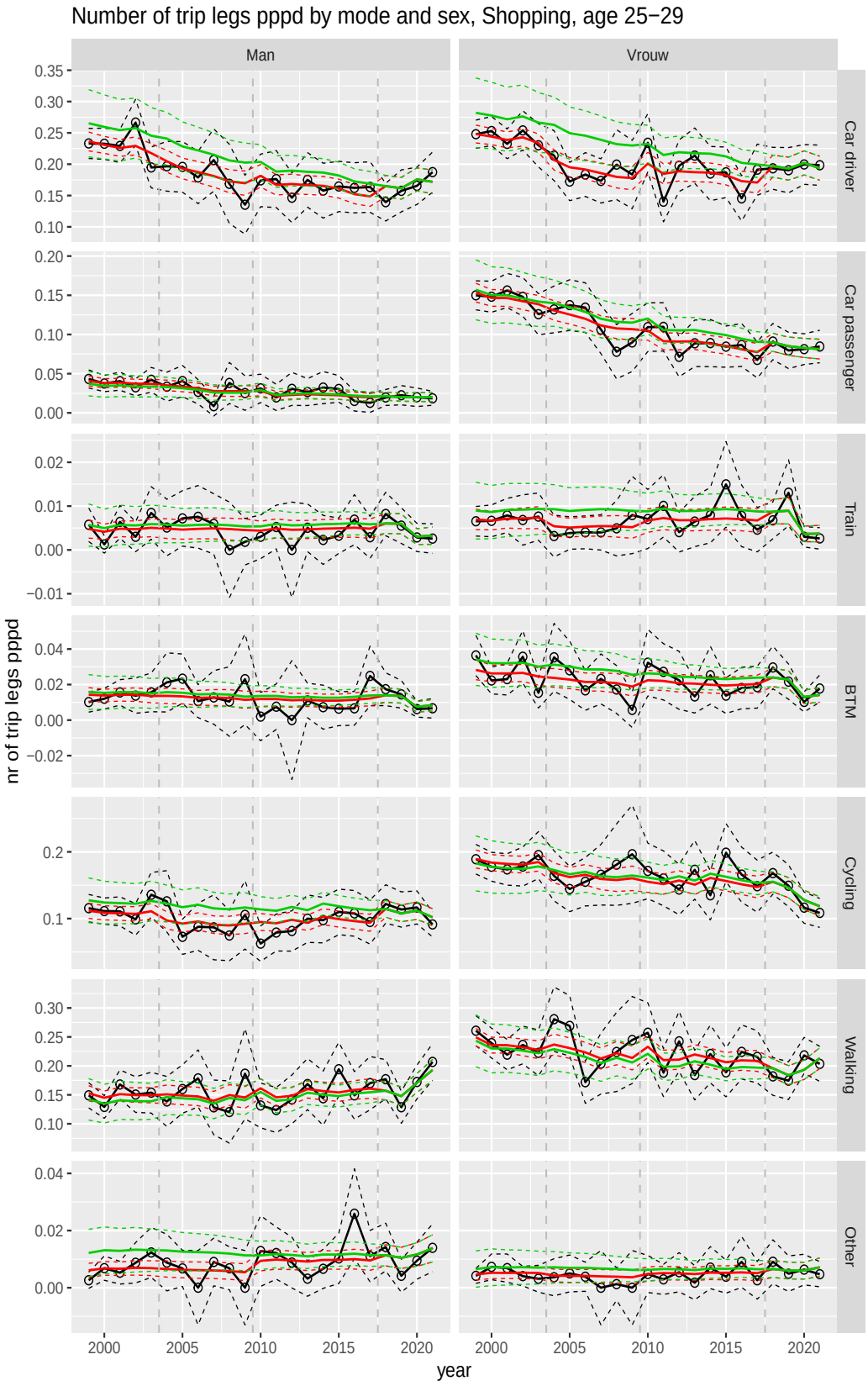

Figure A.51 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

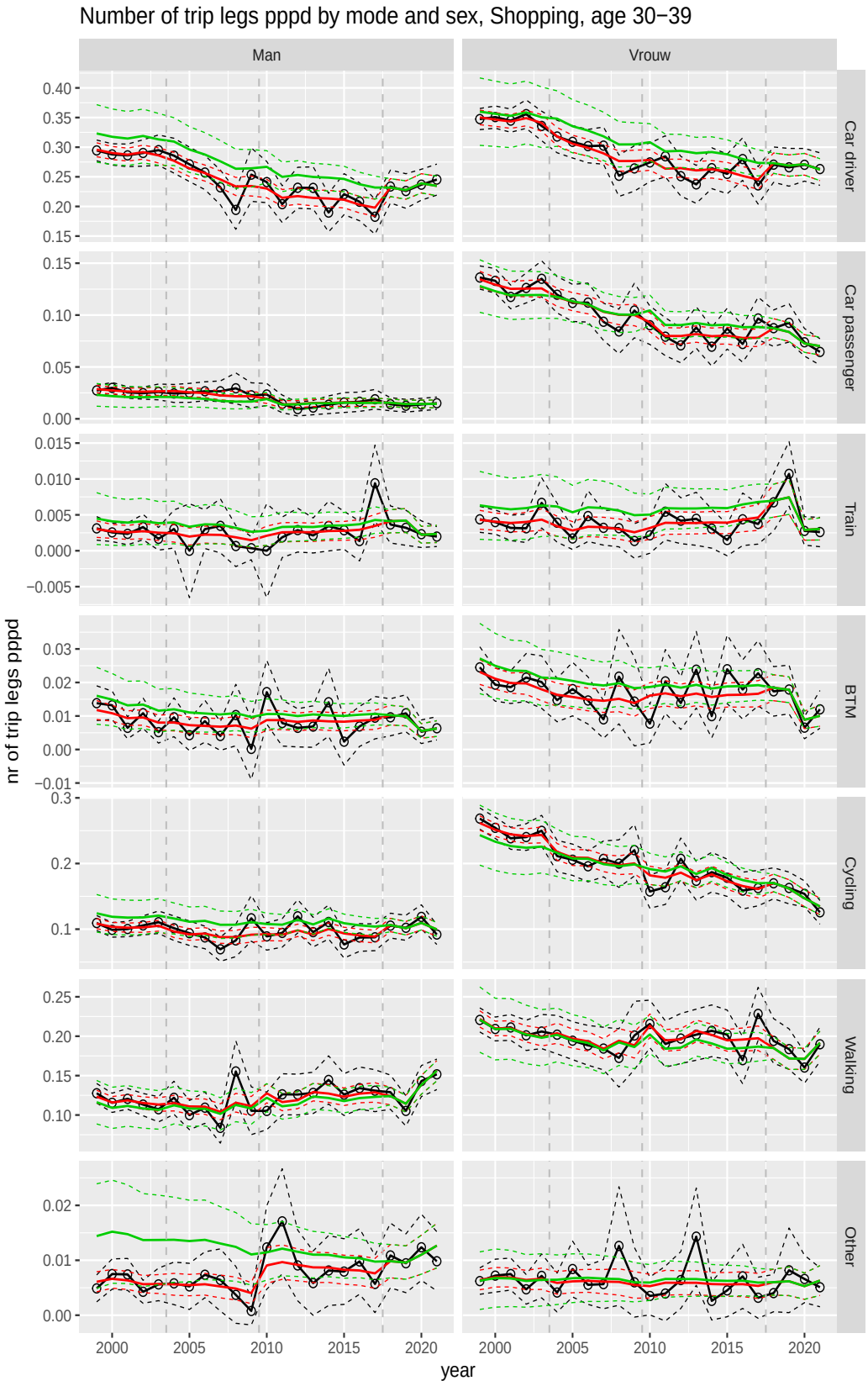

Figure A.52 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

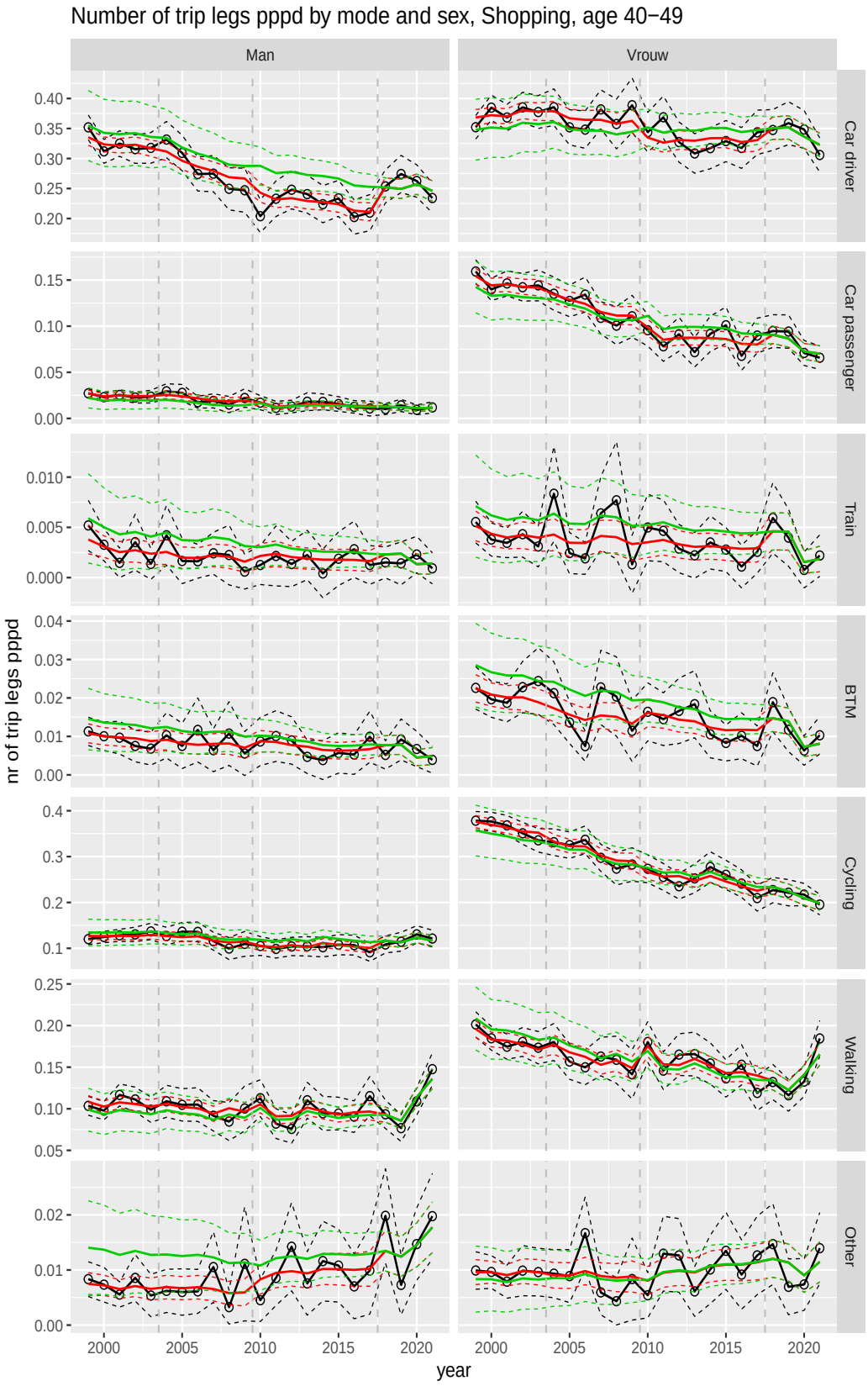

Figure A.53 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

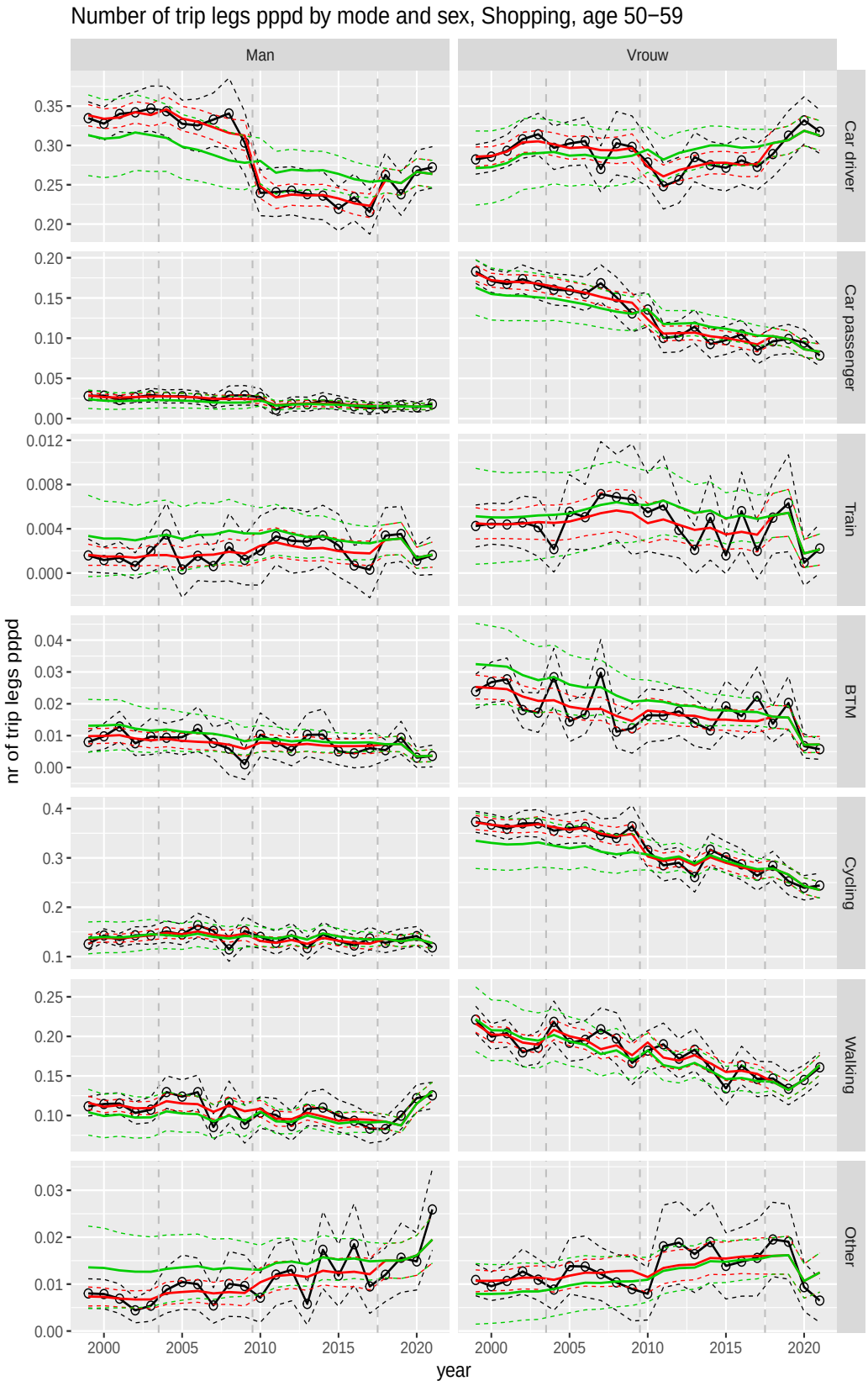

Figure A.54 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

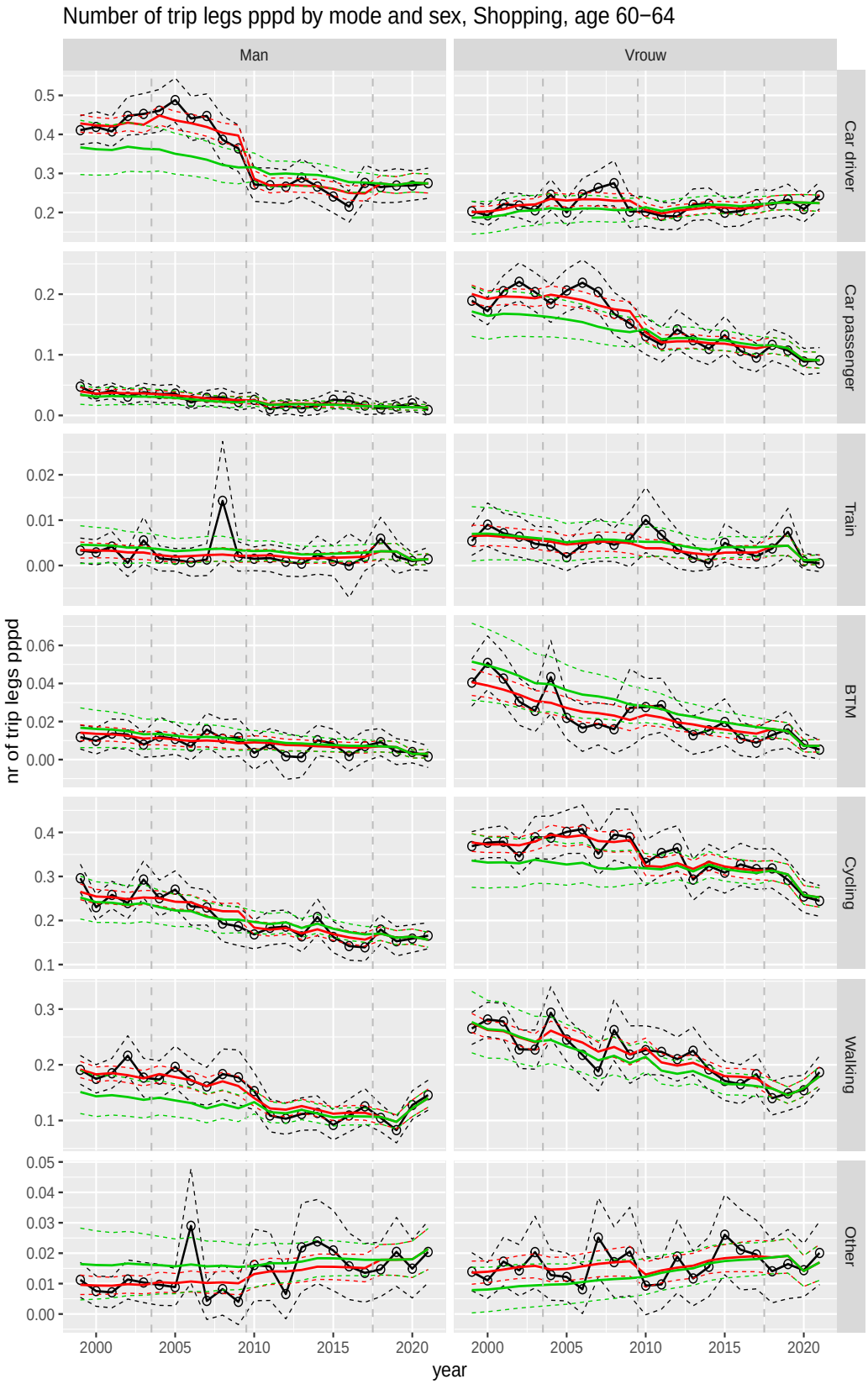

Figure A.55 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

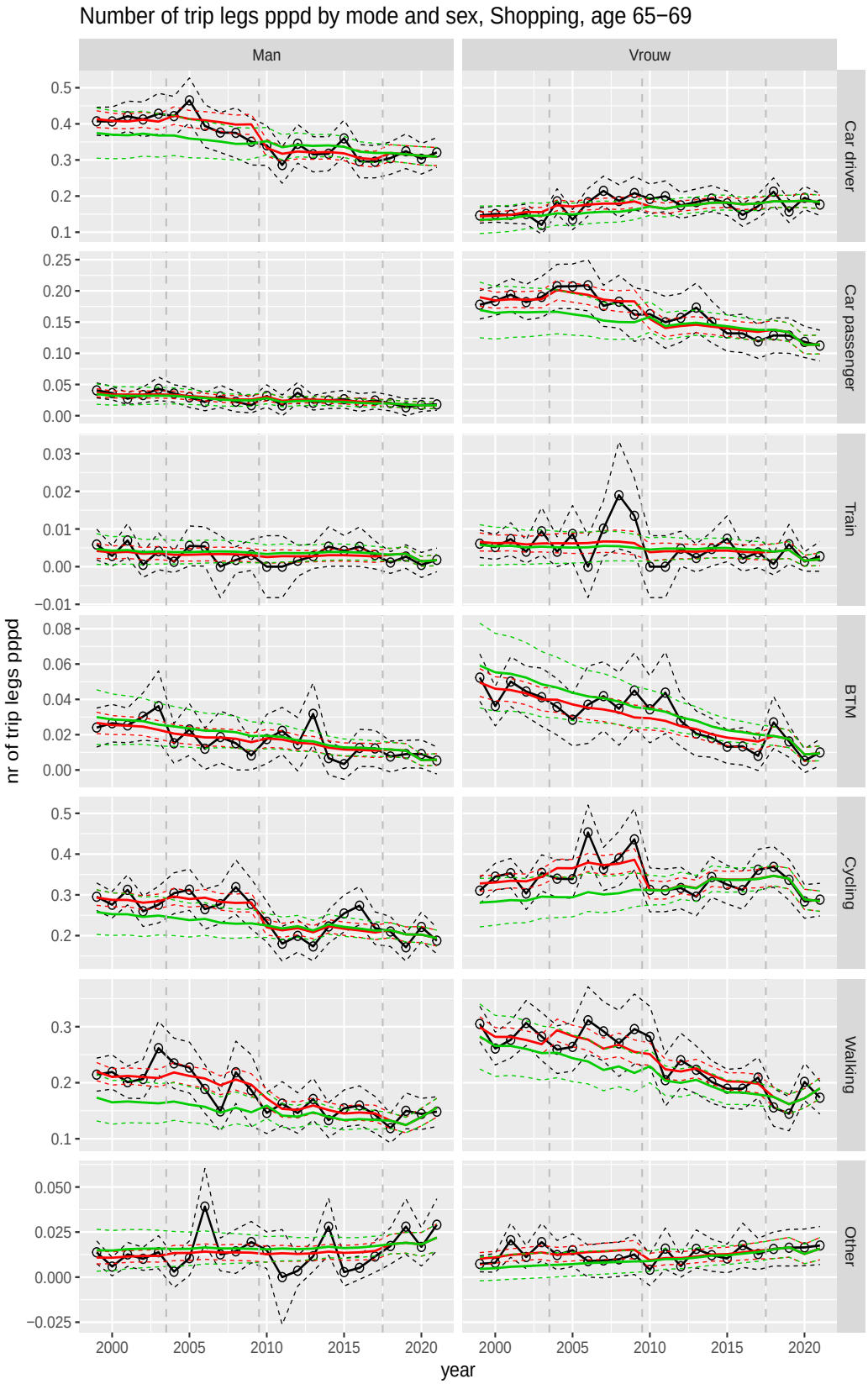

Figure A.56 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

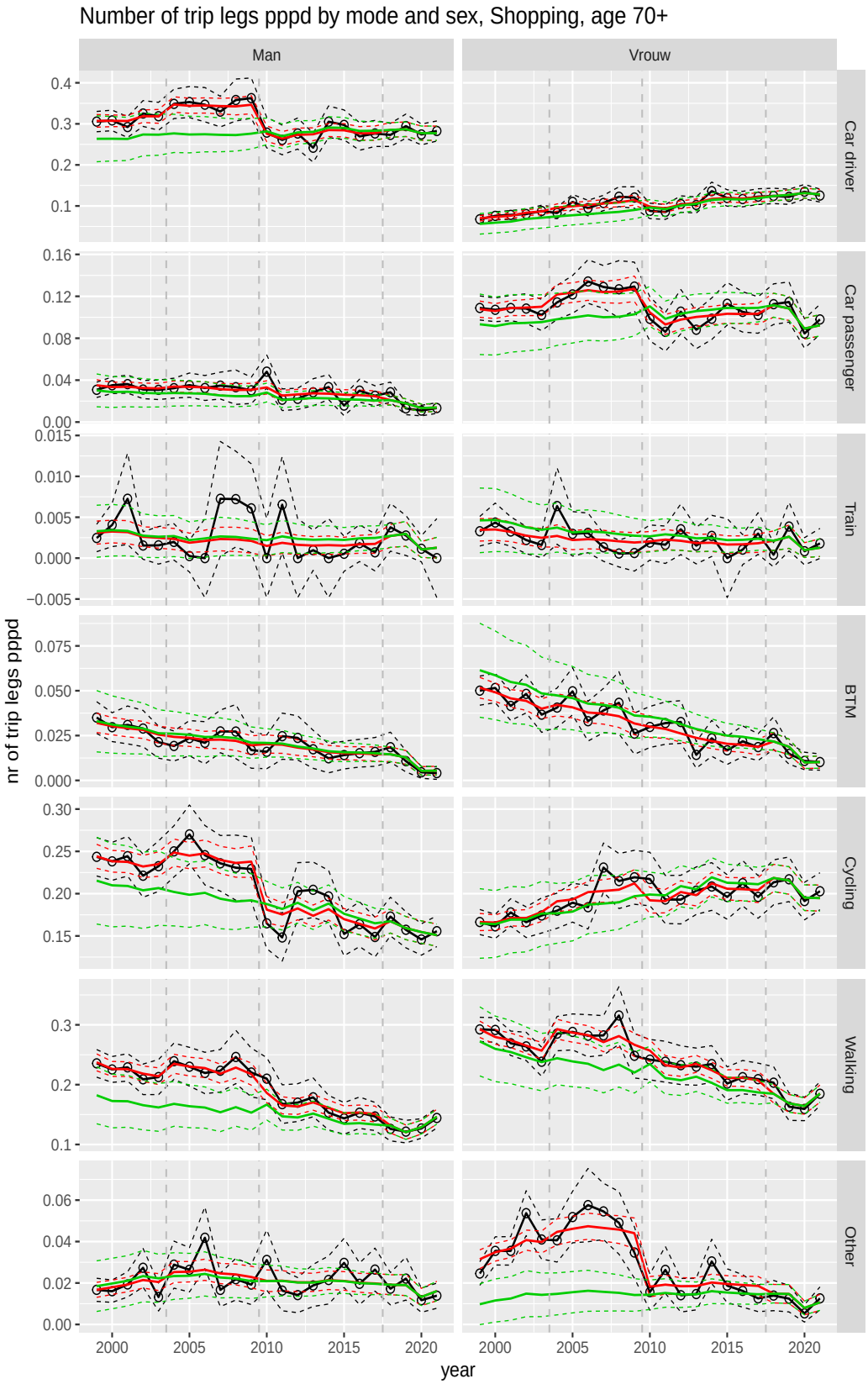

Figure A.57 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

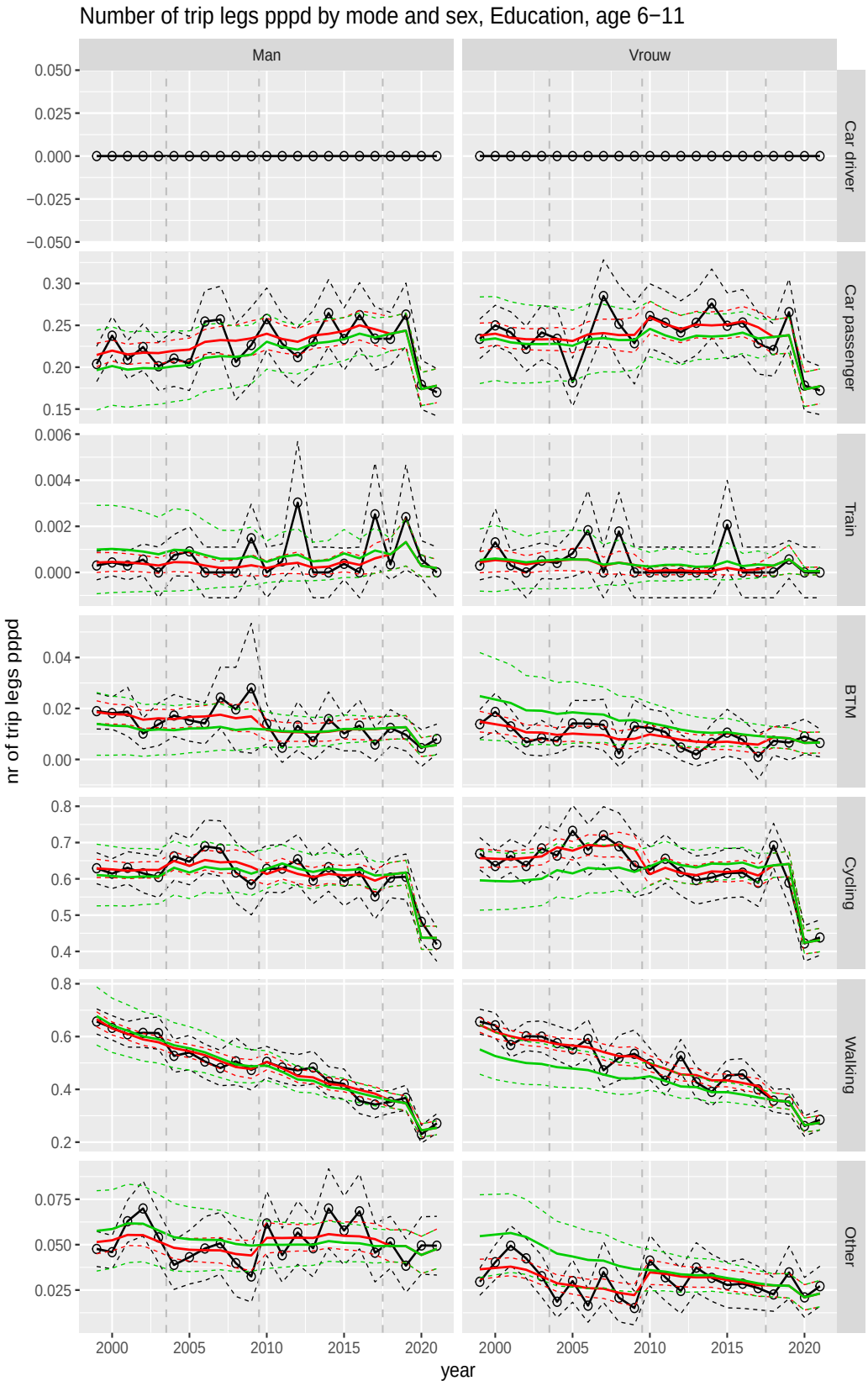

Figure A.58 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

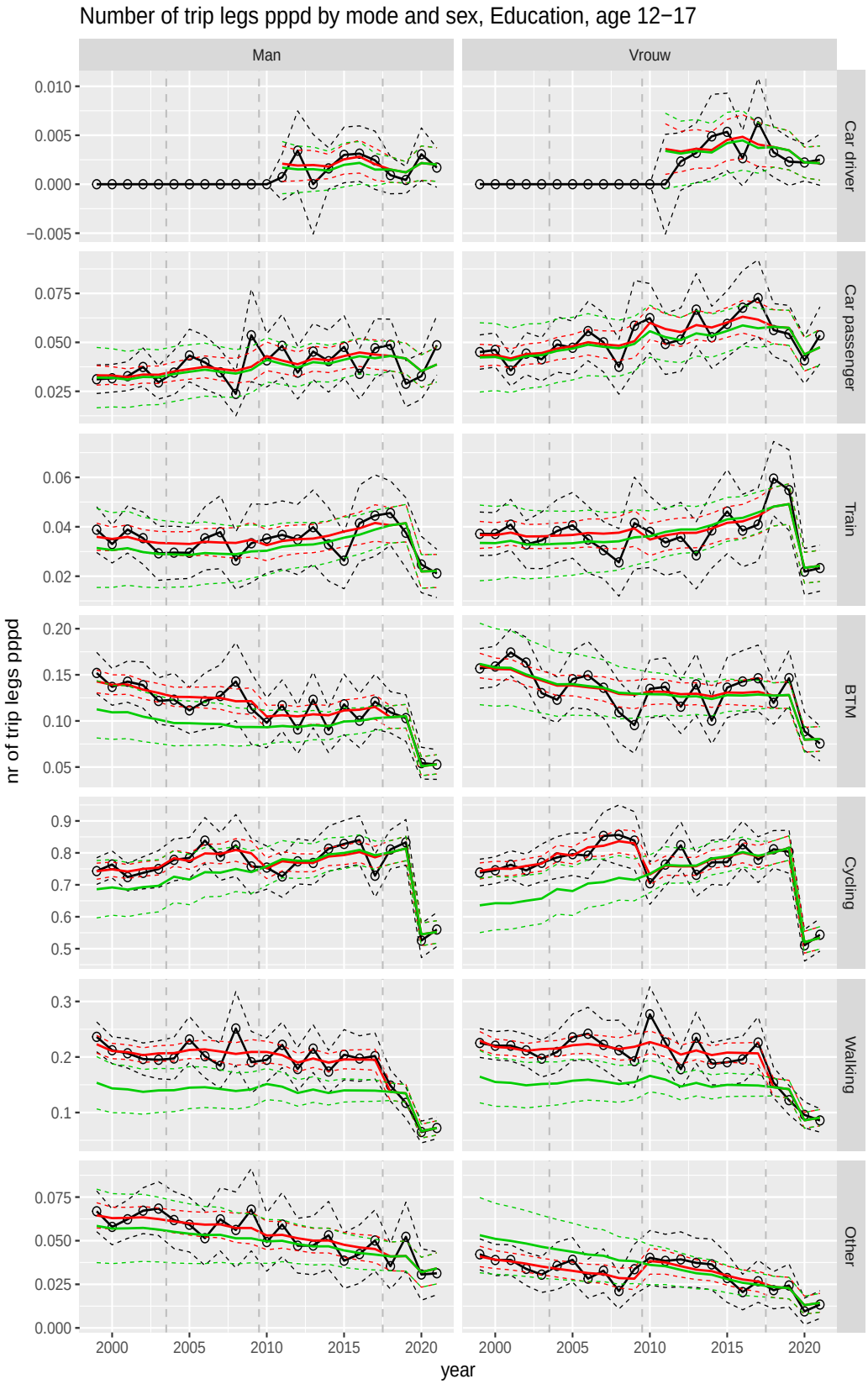

Figure A.59 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

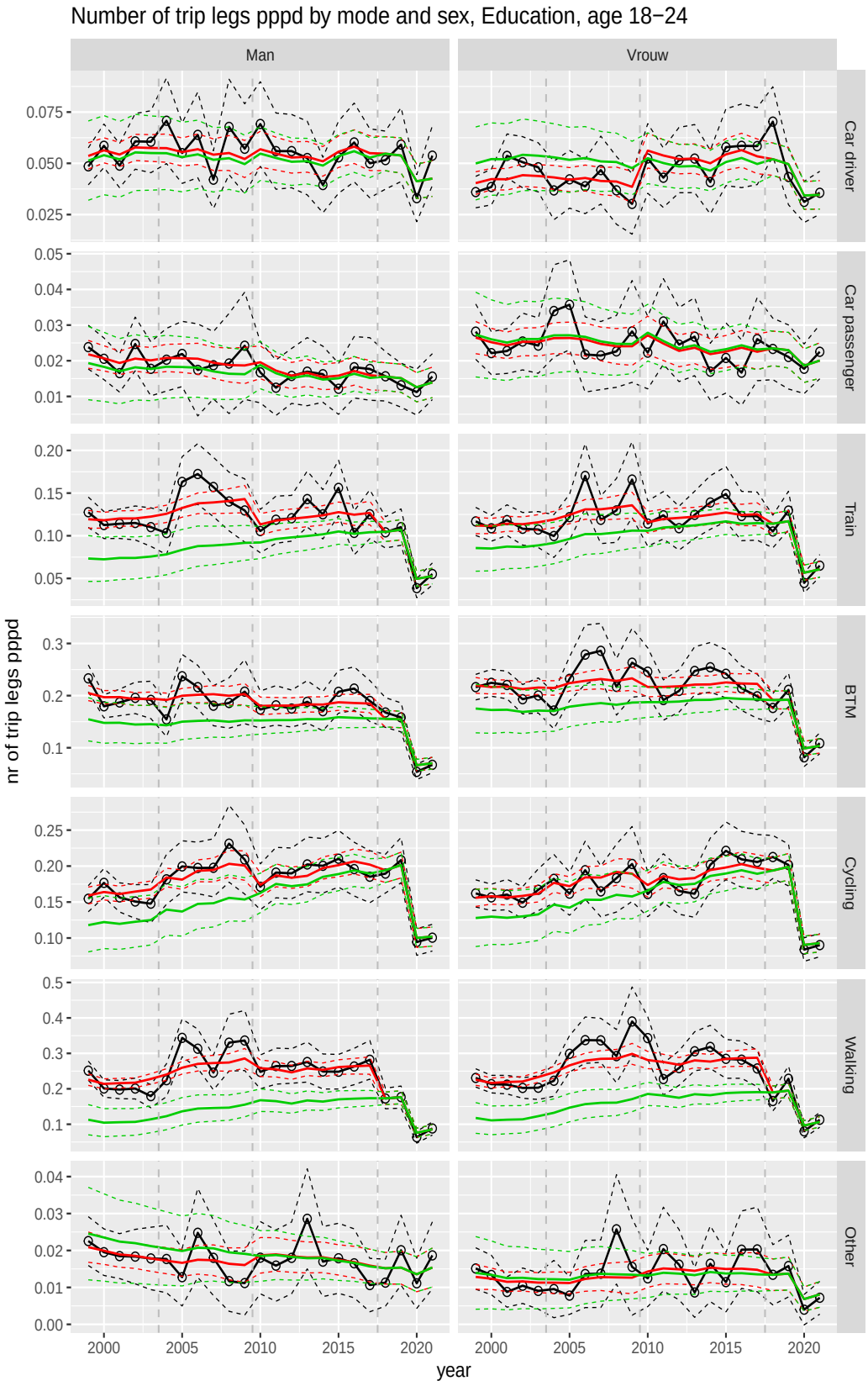

Figure A.60 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

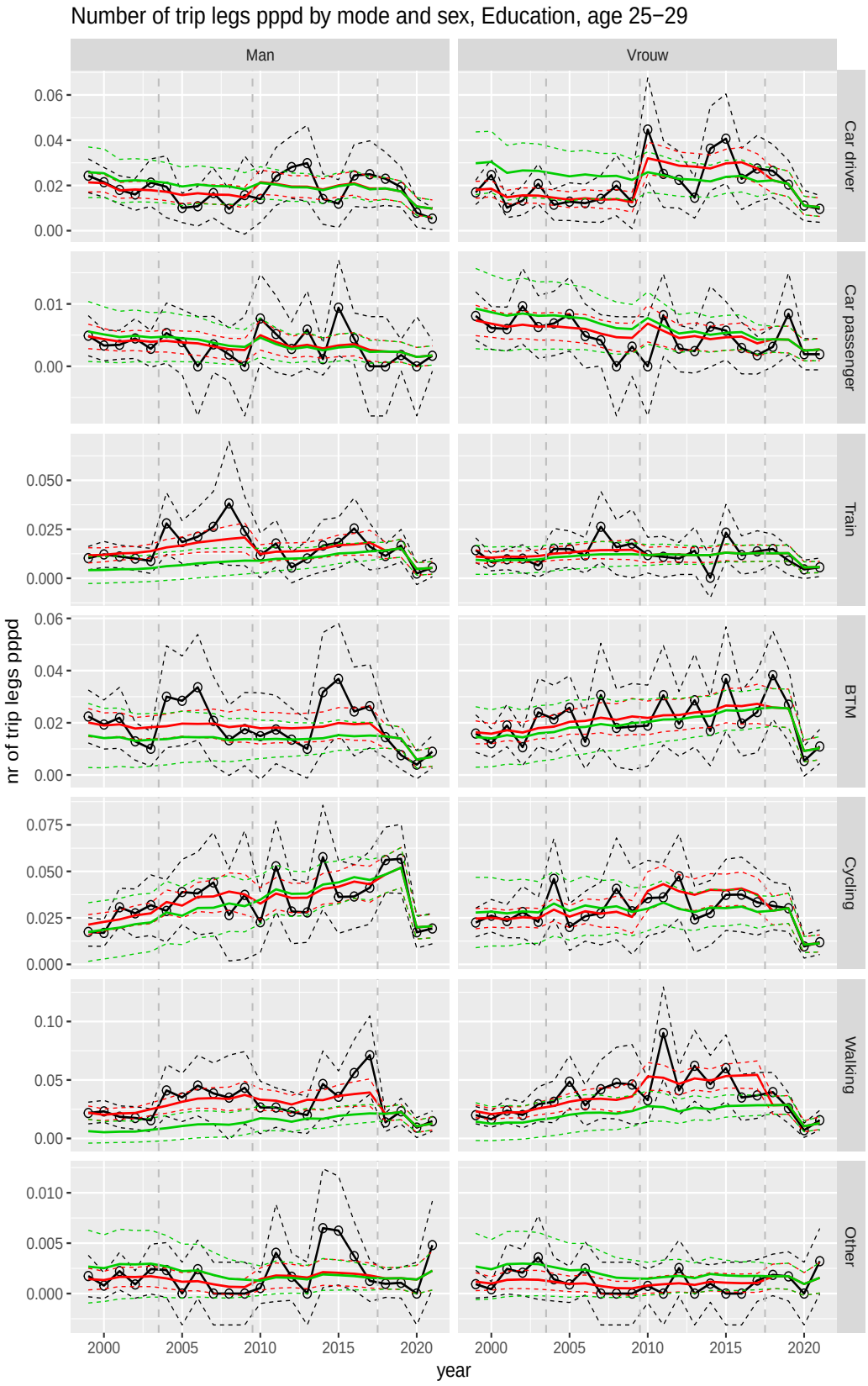

Figure A.61 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

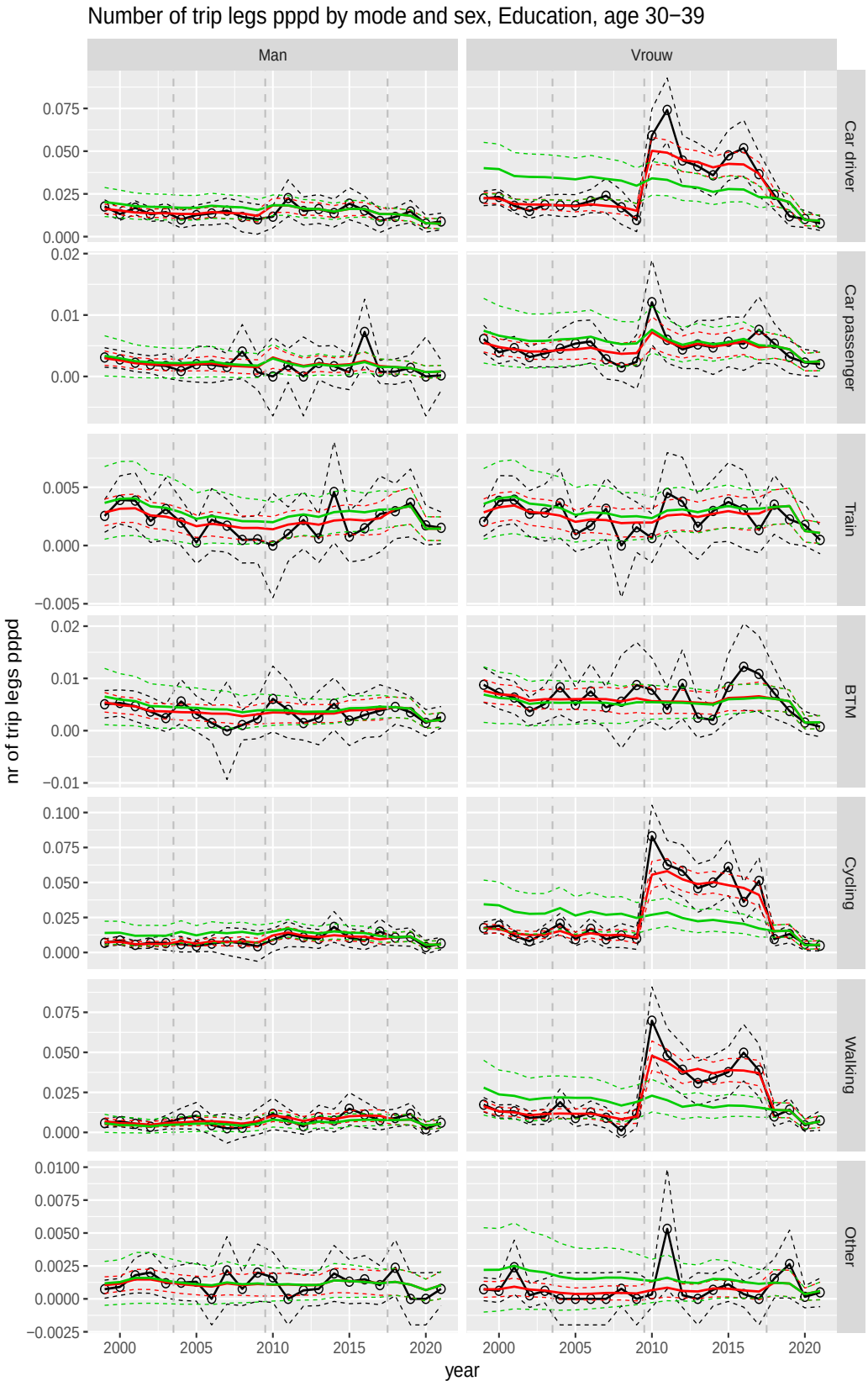

Figure A.62 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

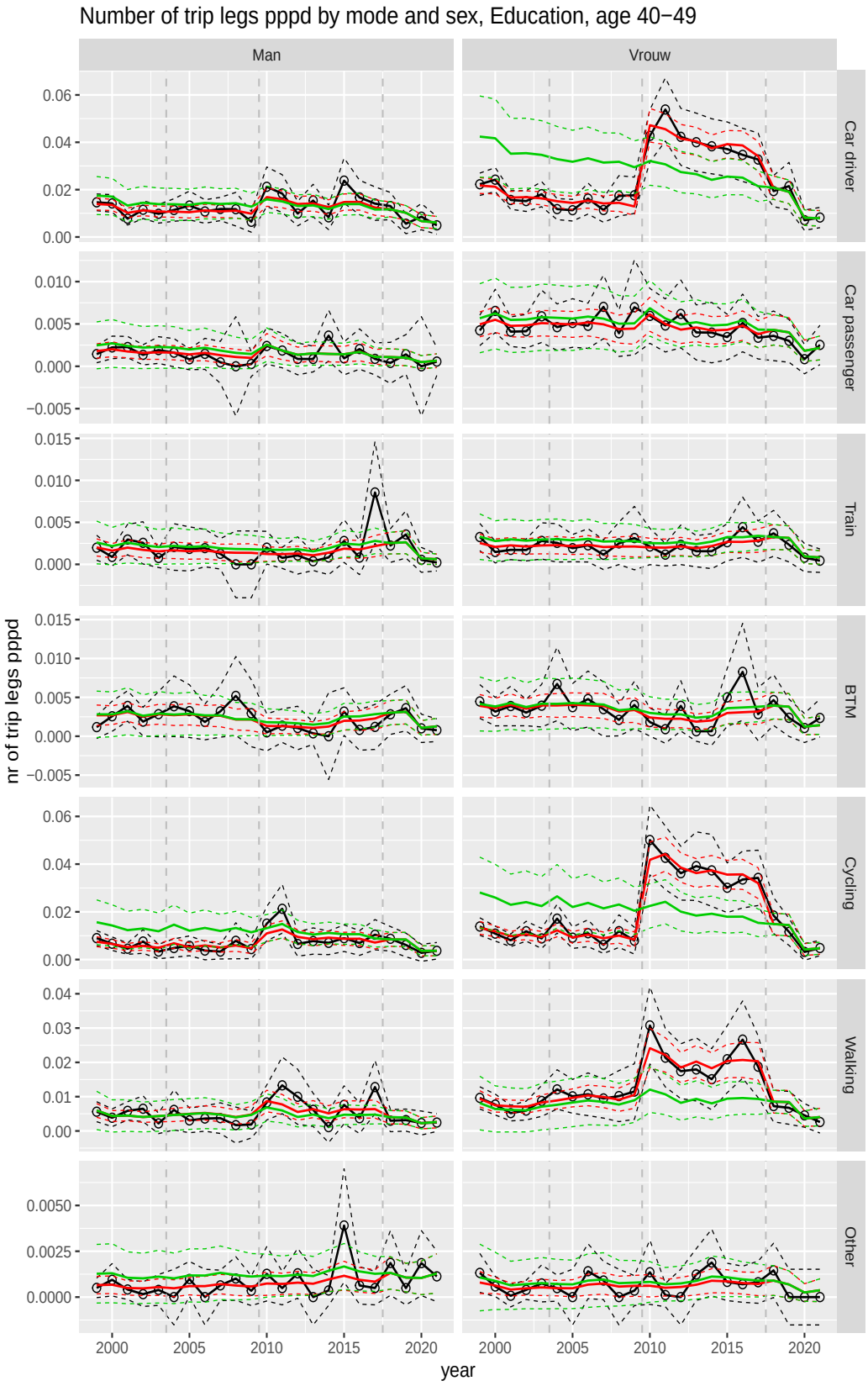

Figure A.63 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

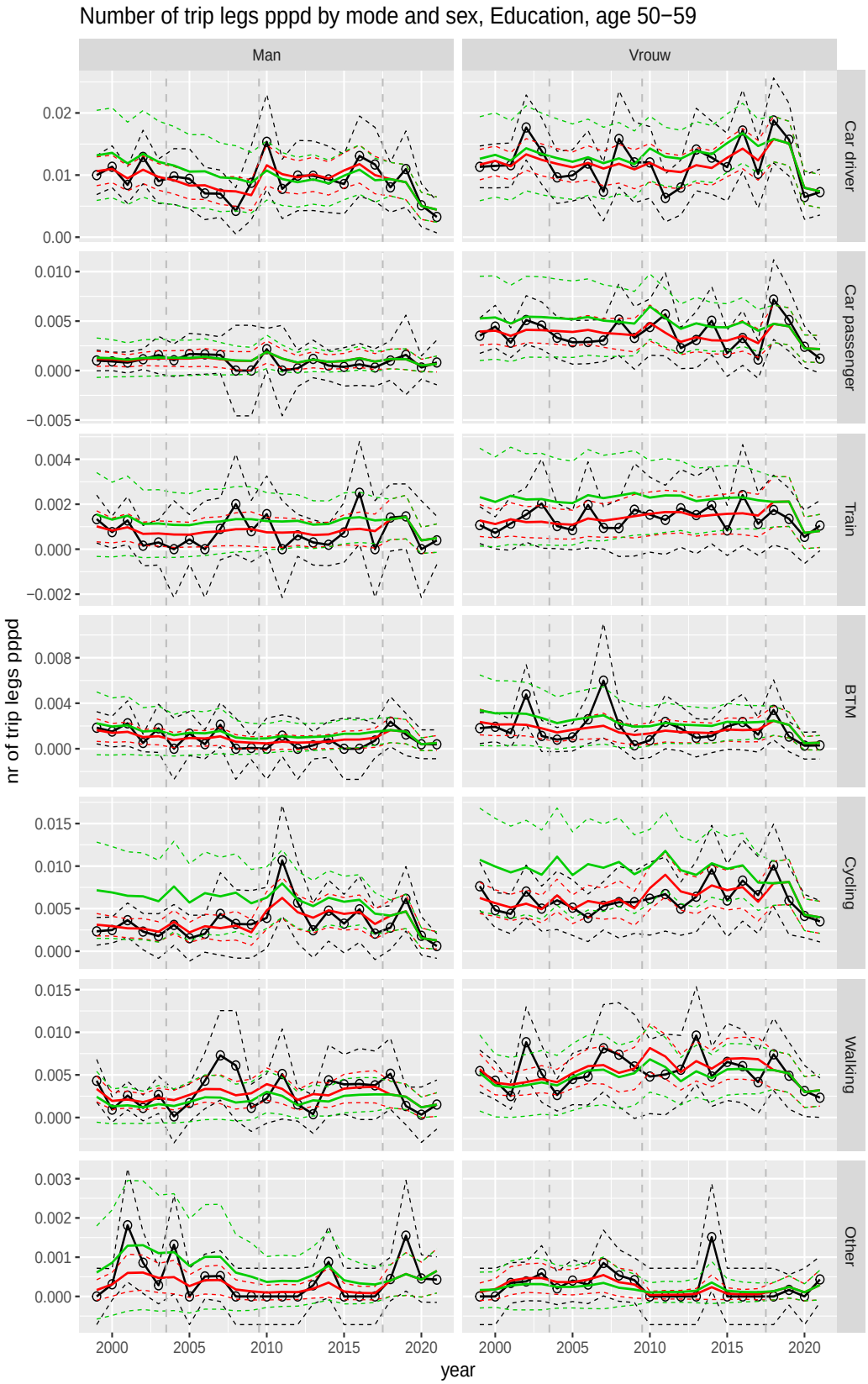

Figure A.64 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

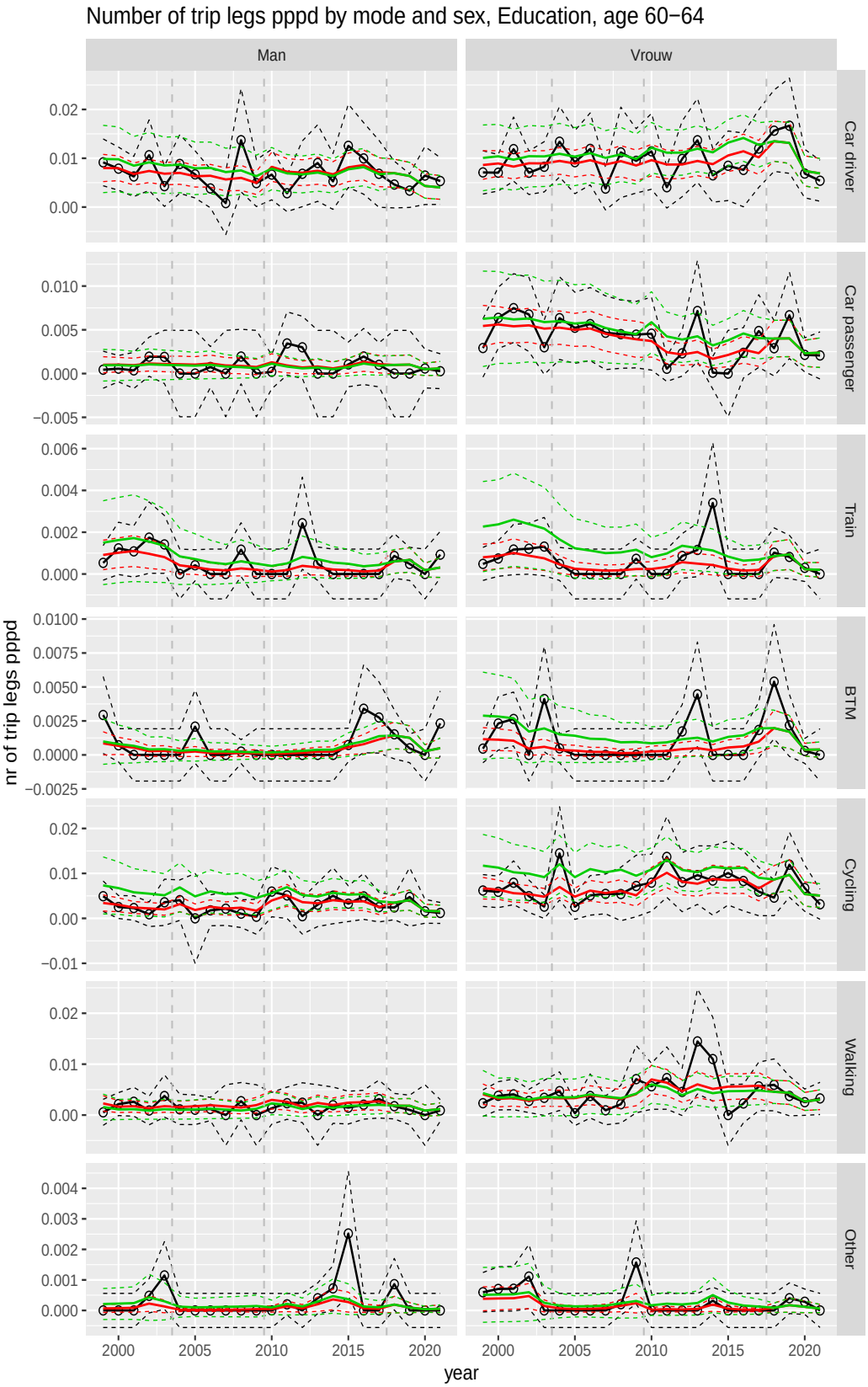

Figure A.65 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

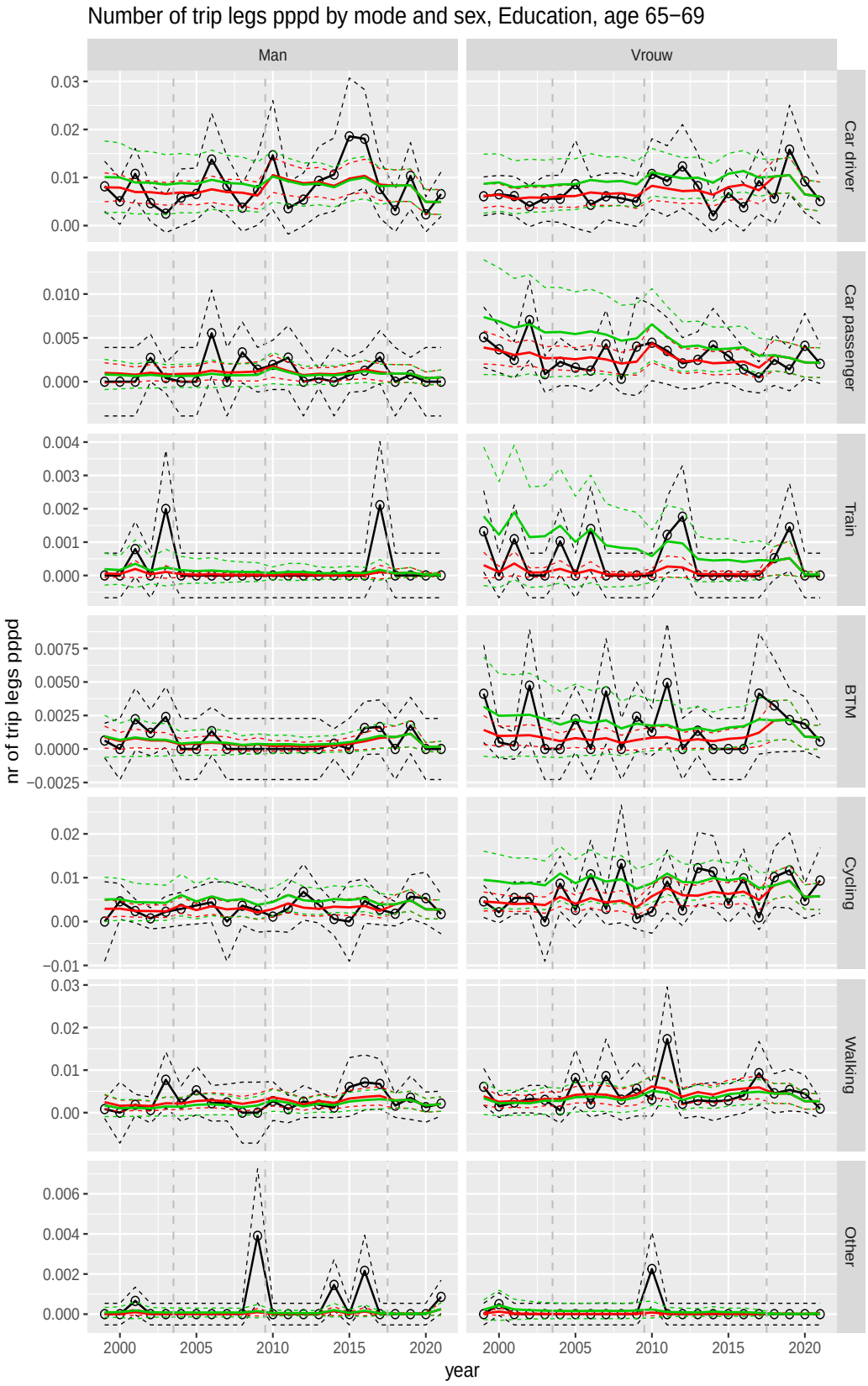

Figure A.66 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

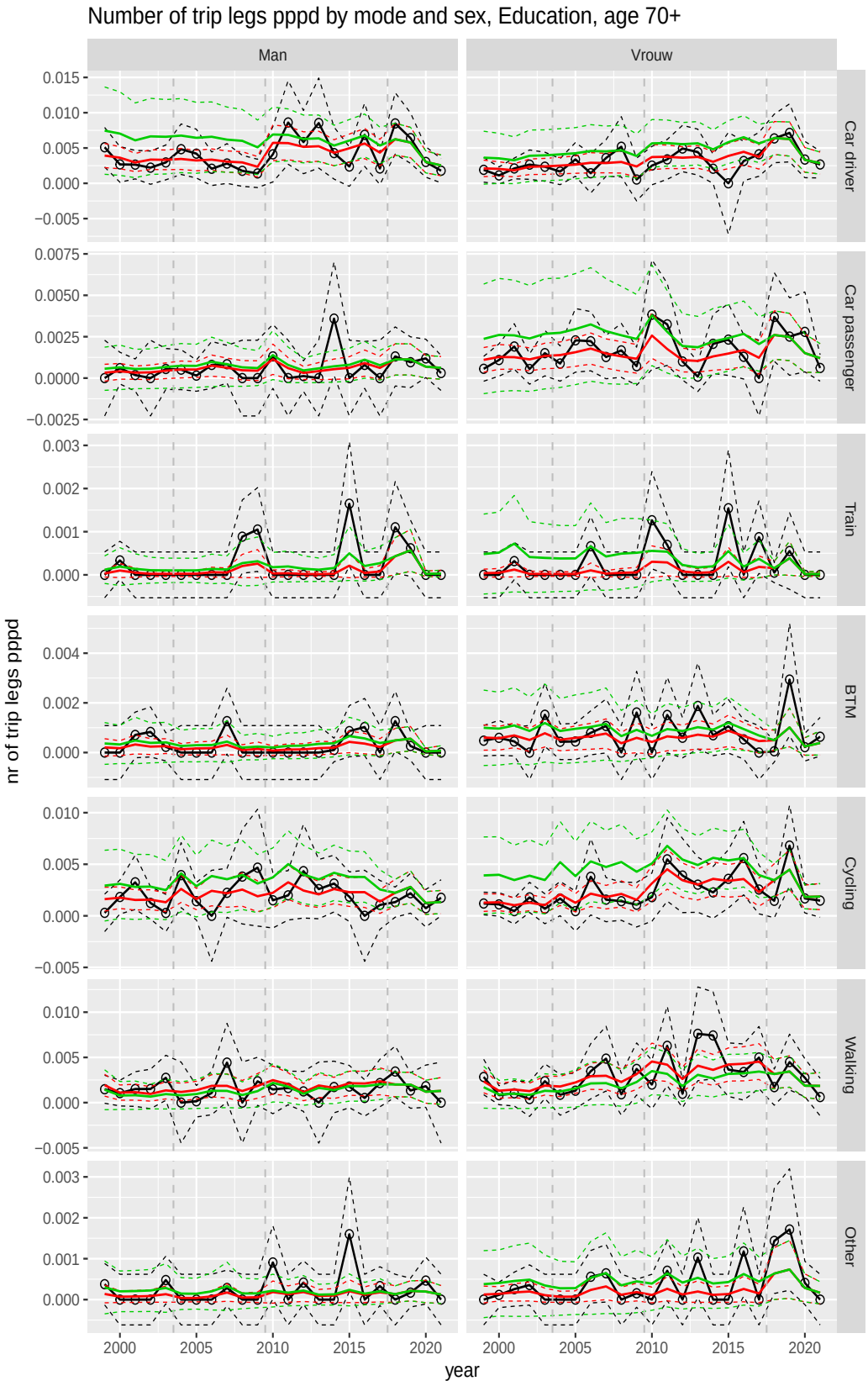

Figure A.67 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

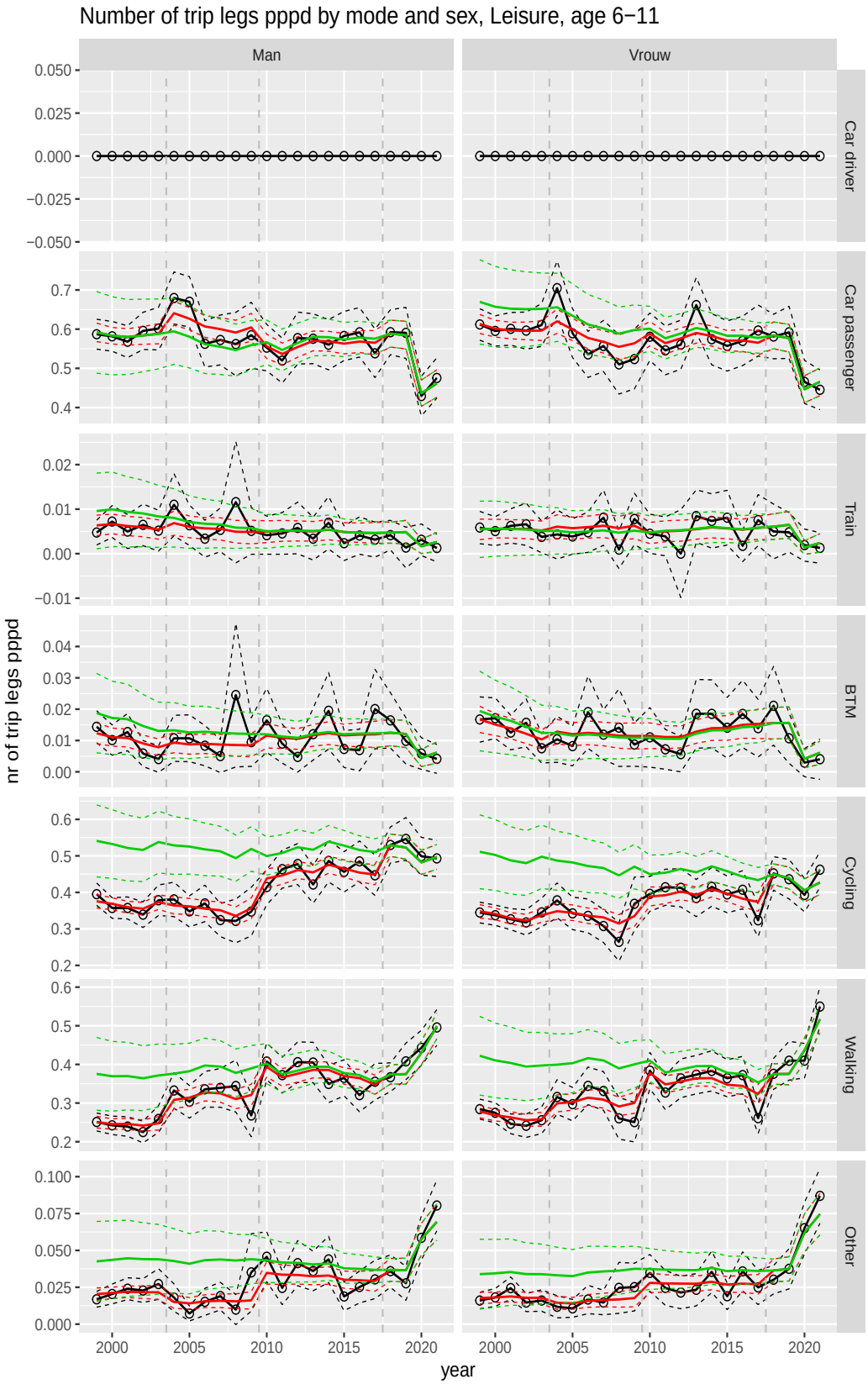

Figure A.68 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

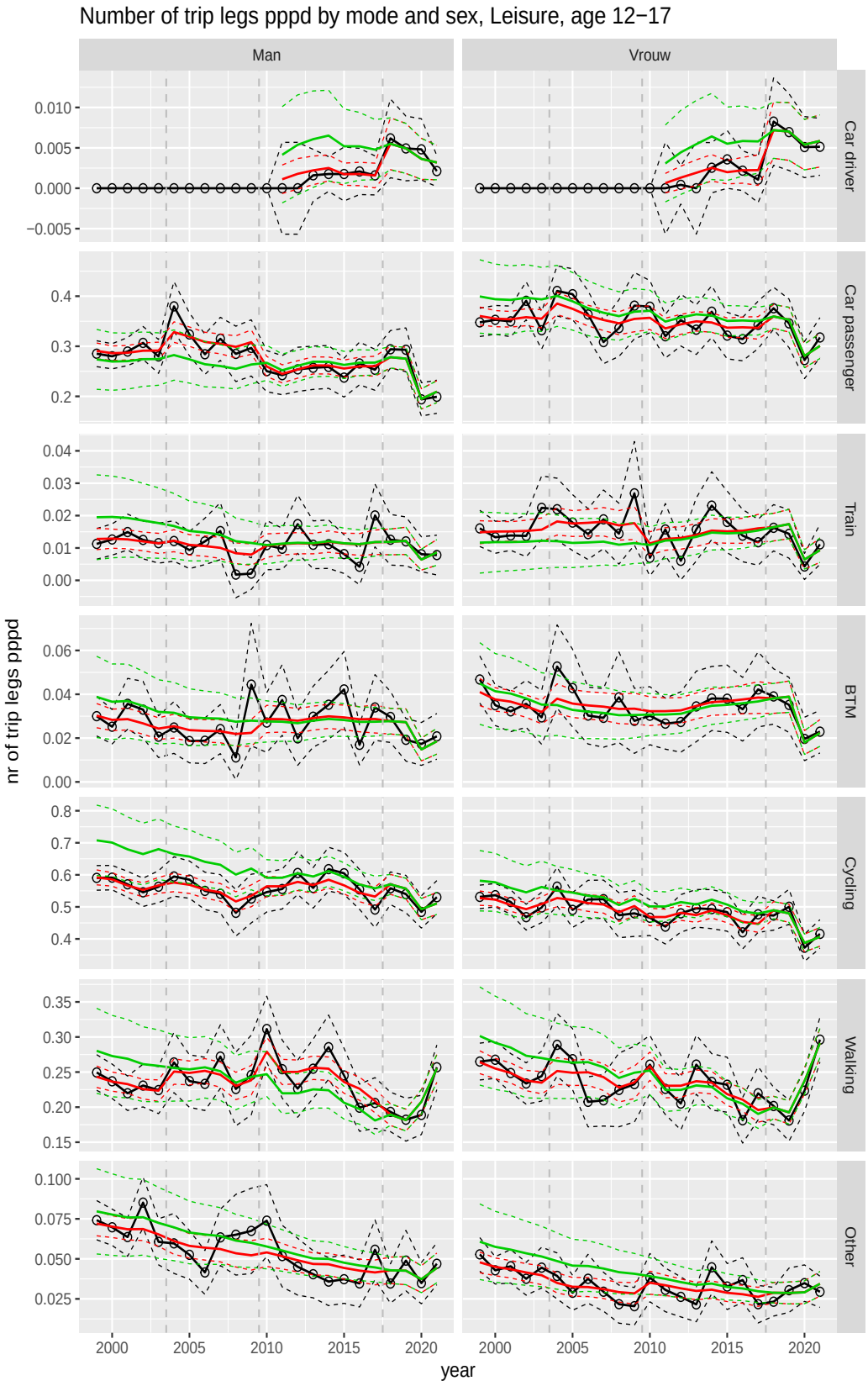

Figure A.69 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

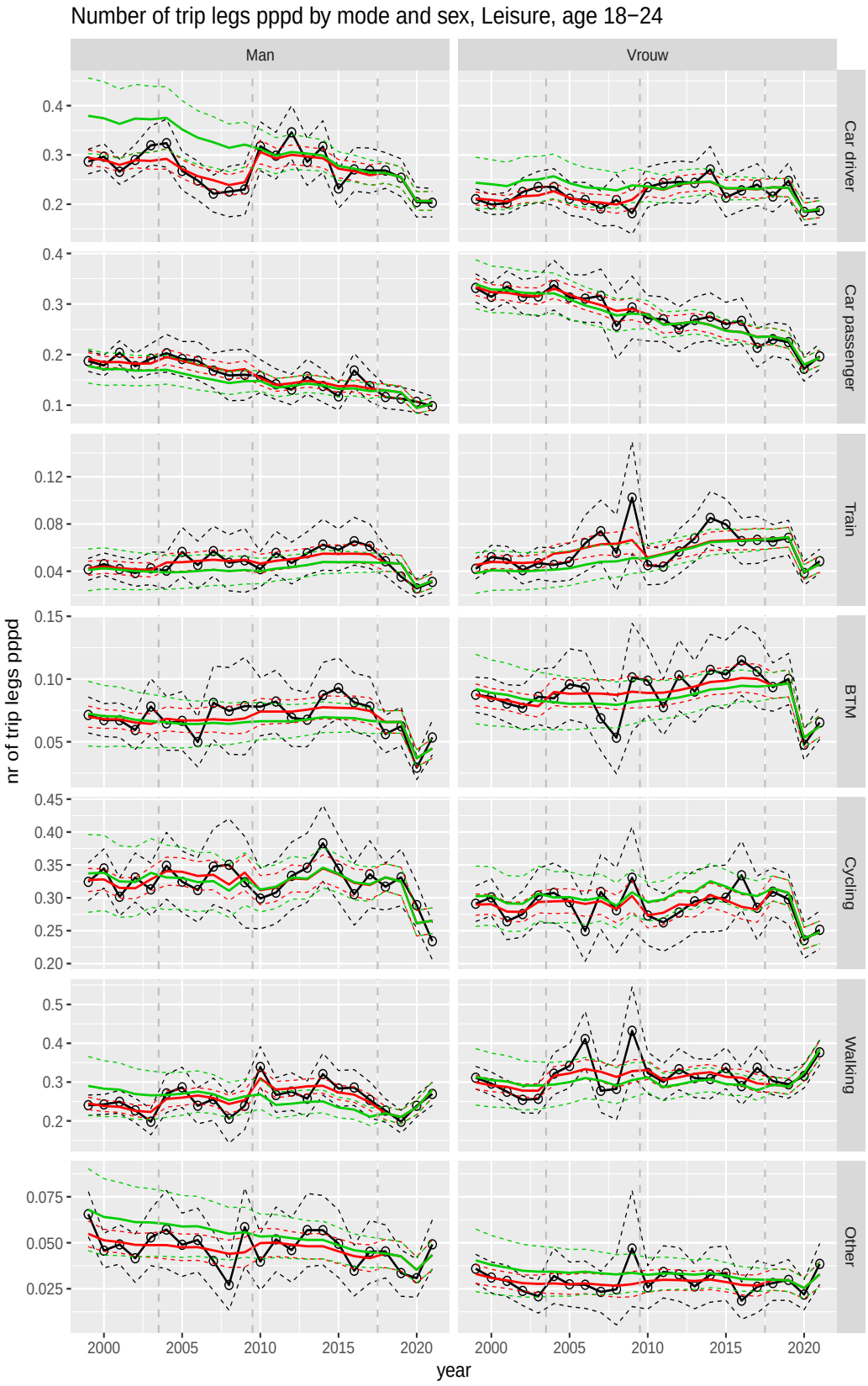

Figure A.70 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

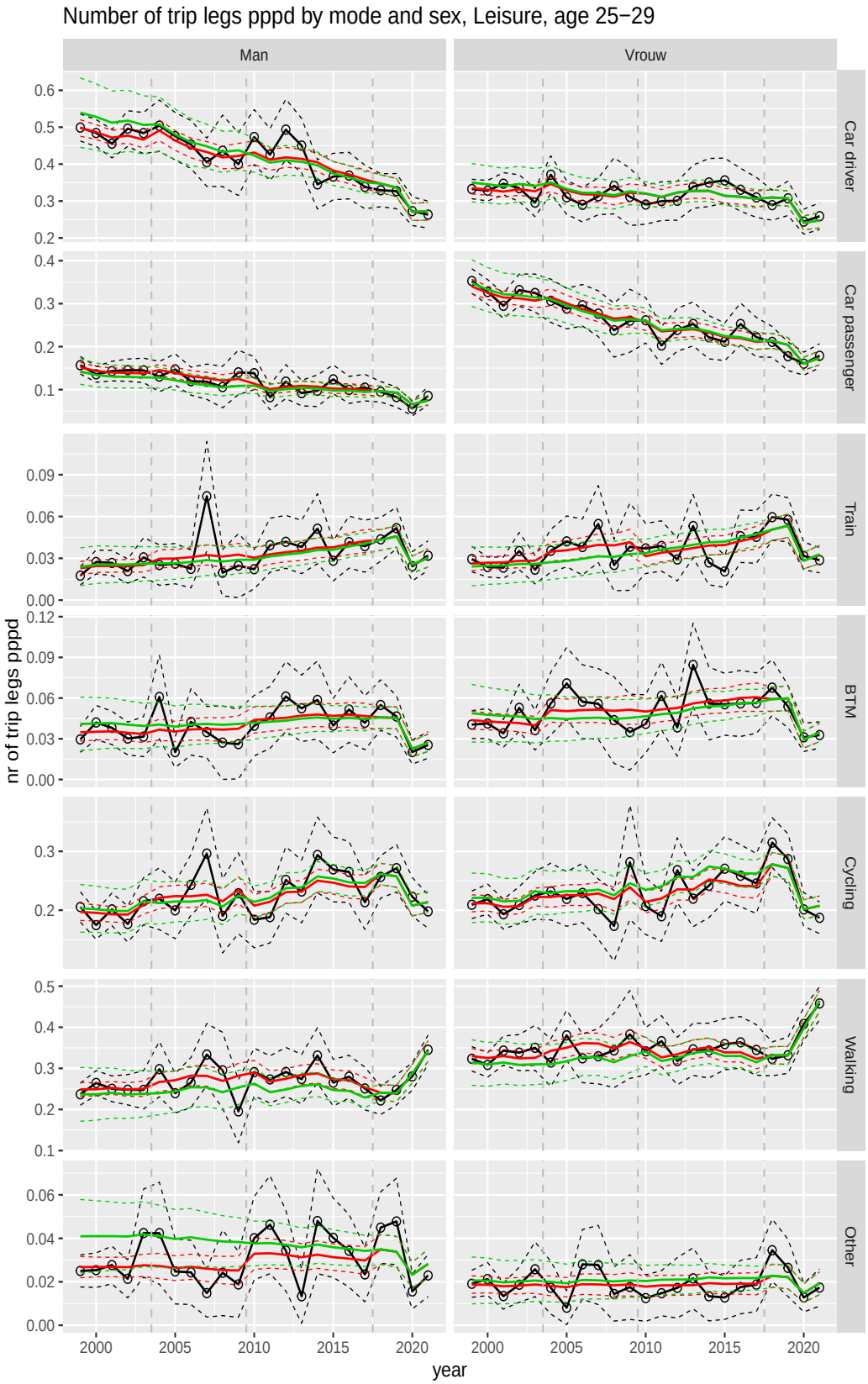

Figure A.71 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

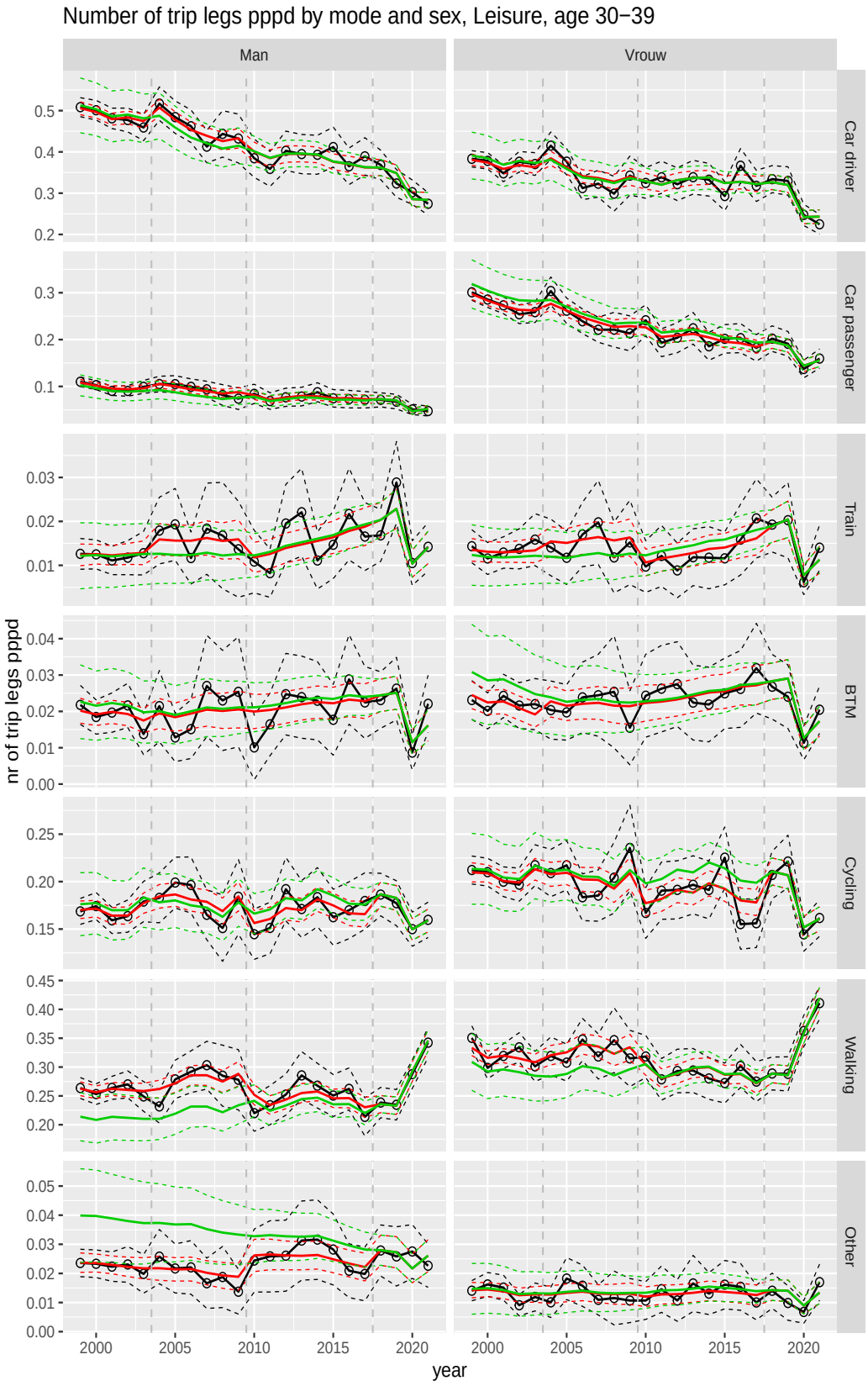

Figure A.72 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

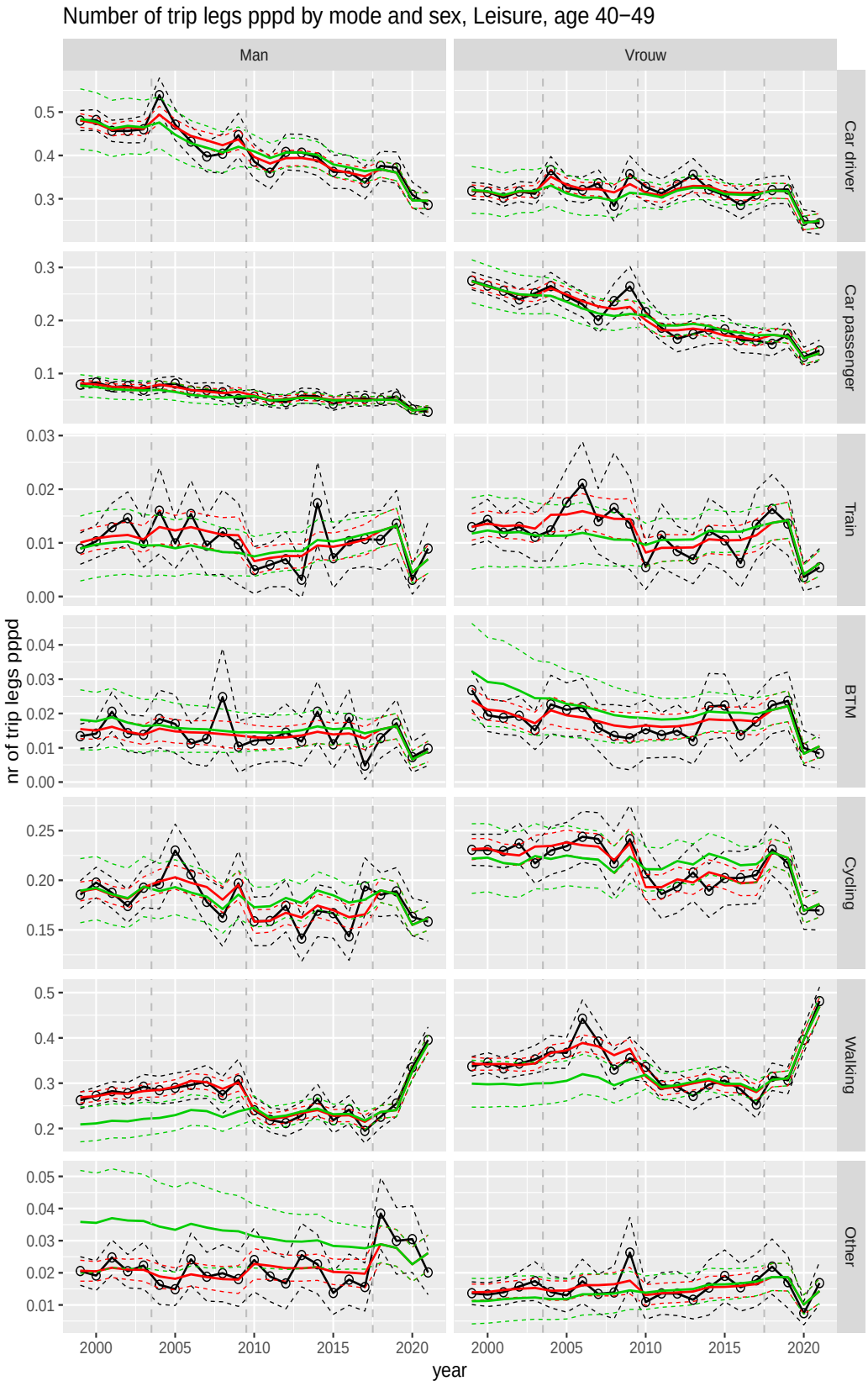

Figure A.73 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

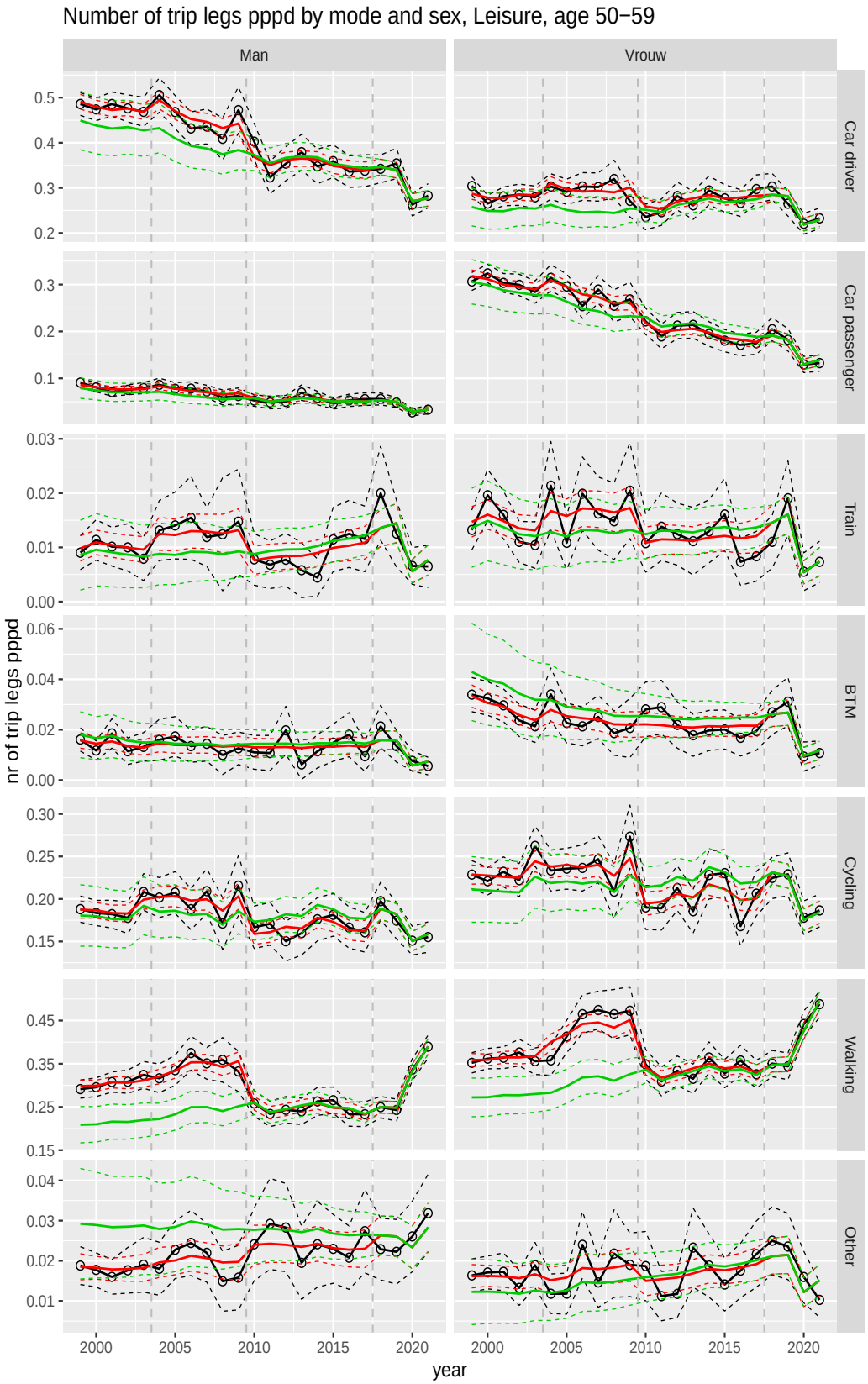

Figure A.74 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

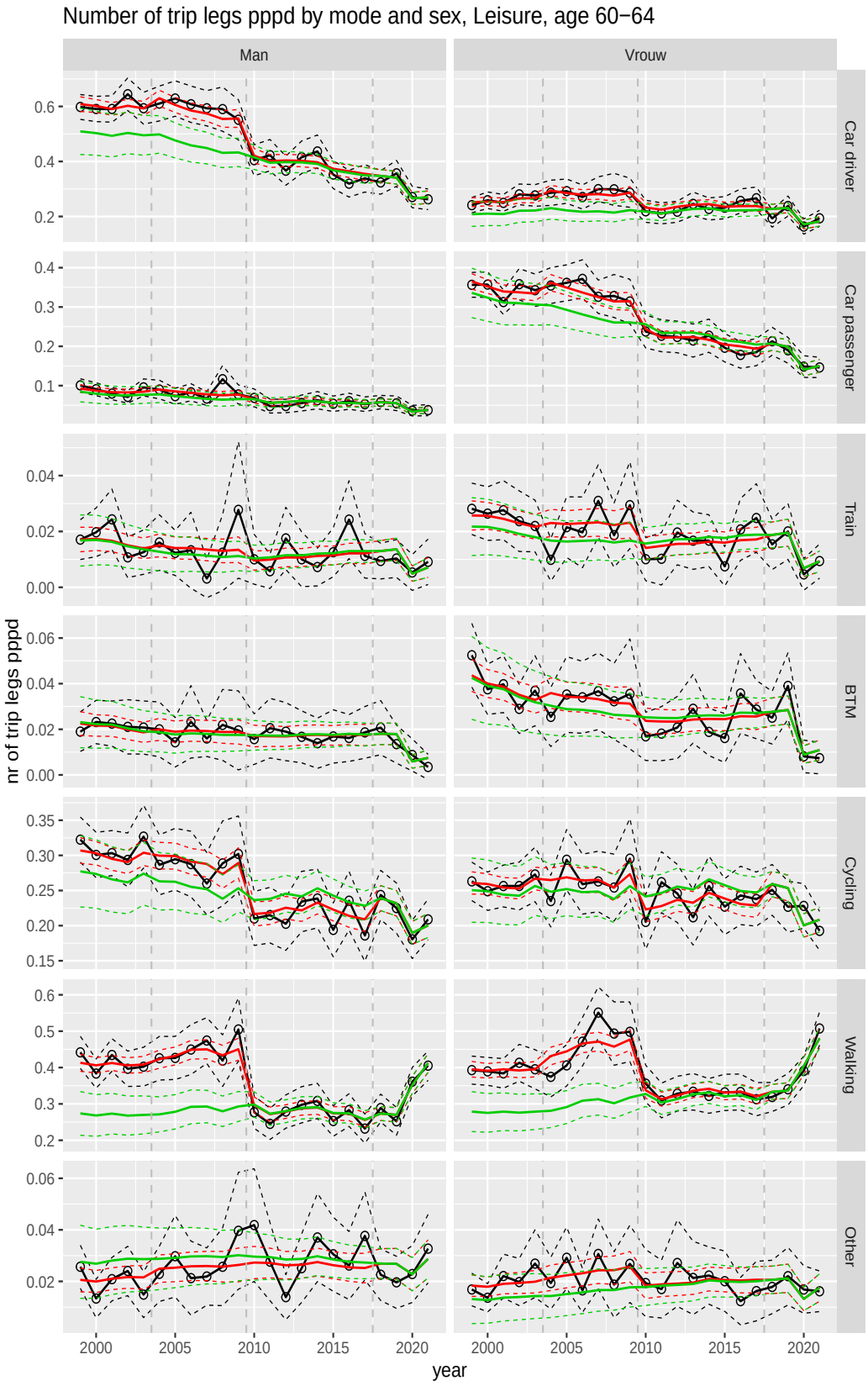

Figure A.75 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

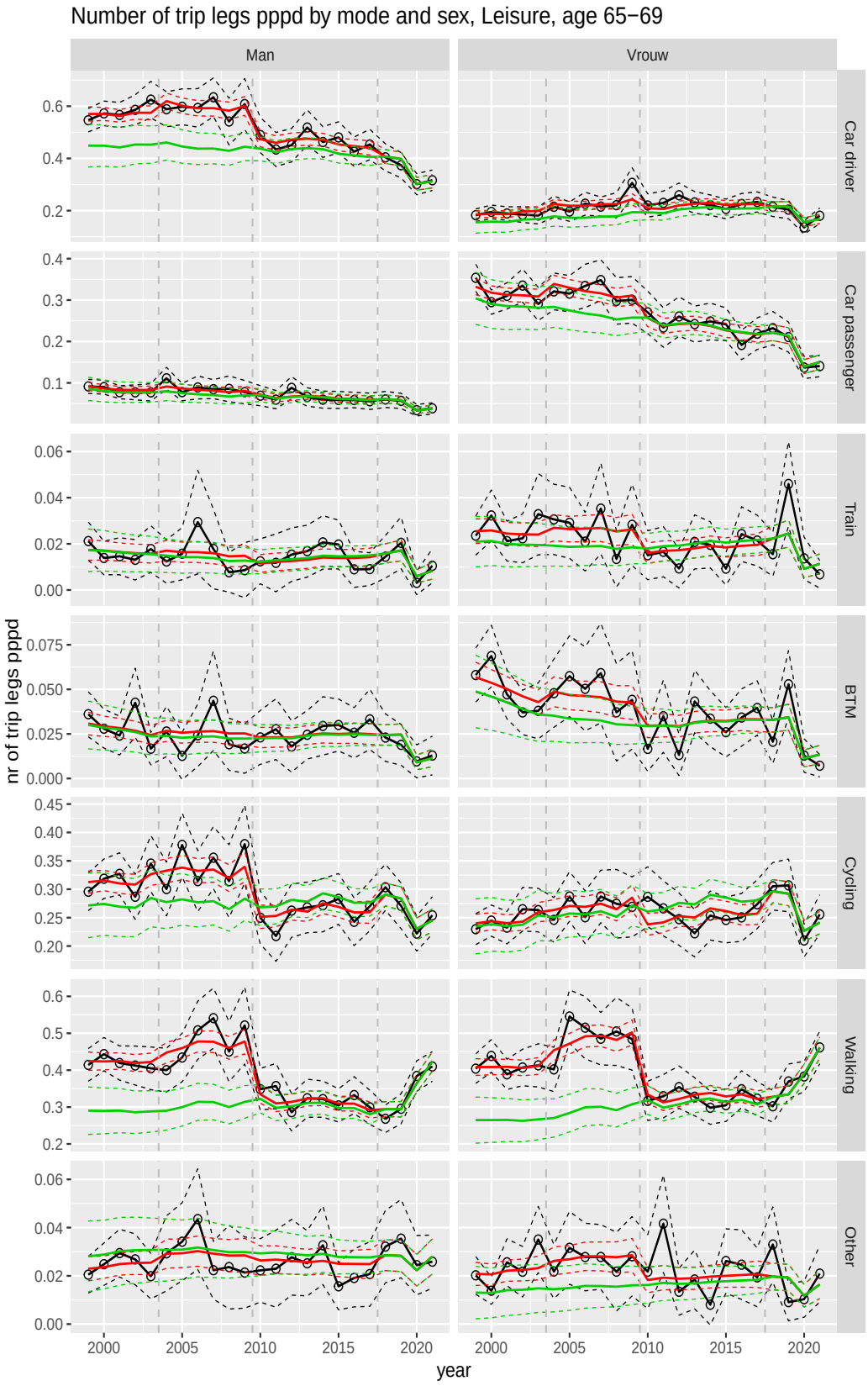

Figure A.76 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

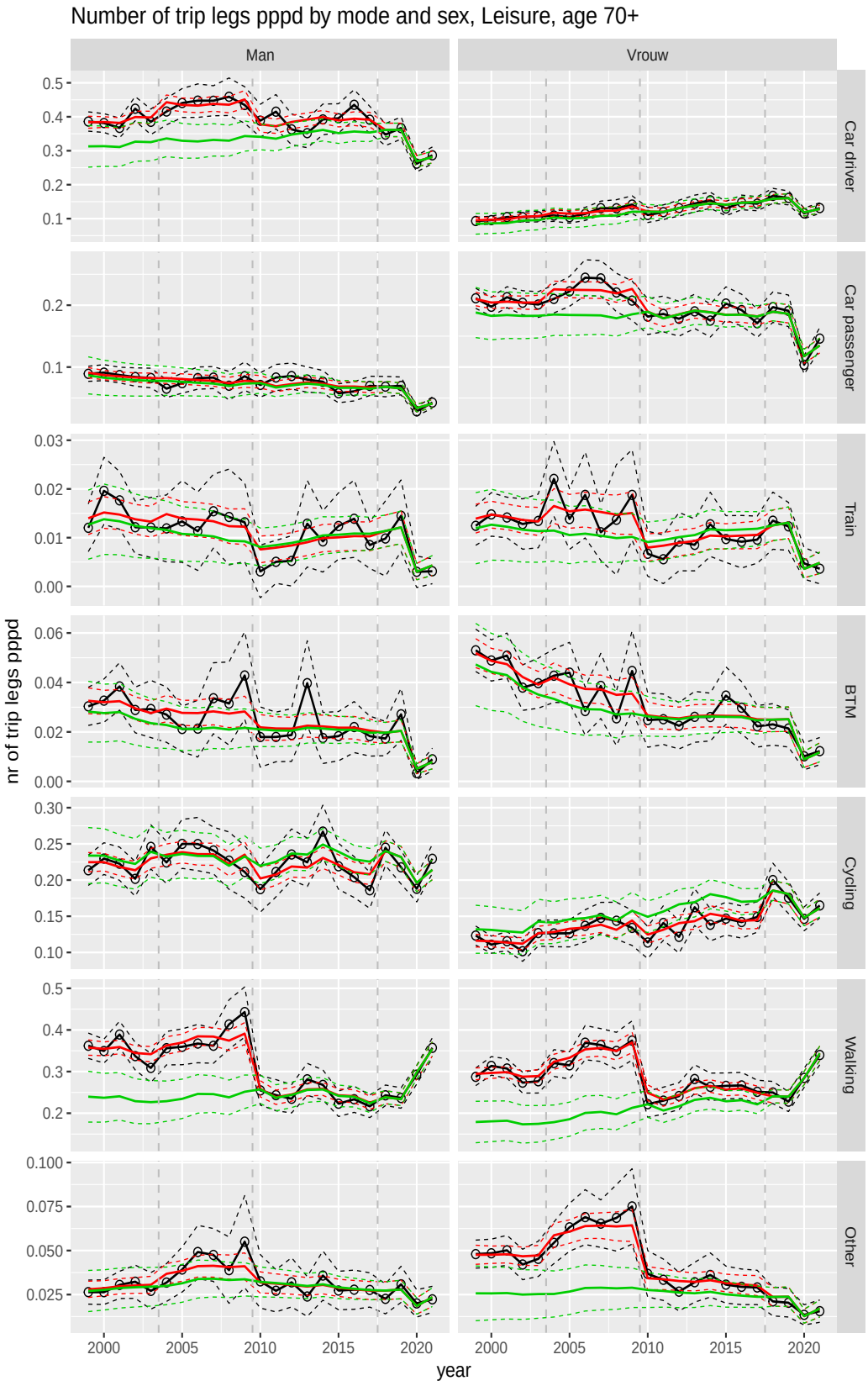

Figure A.77 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.78 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

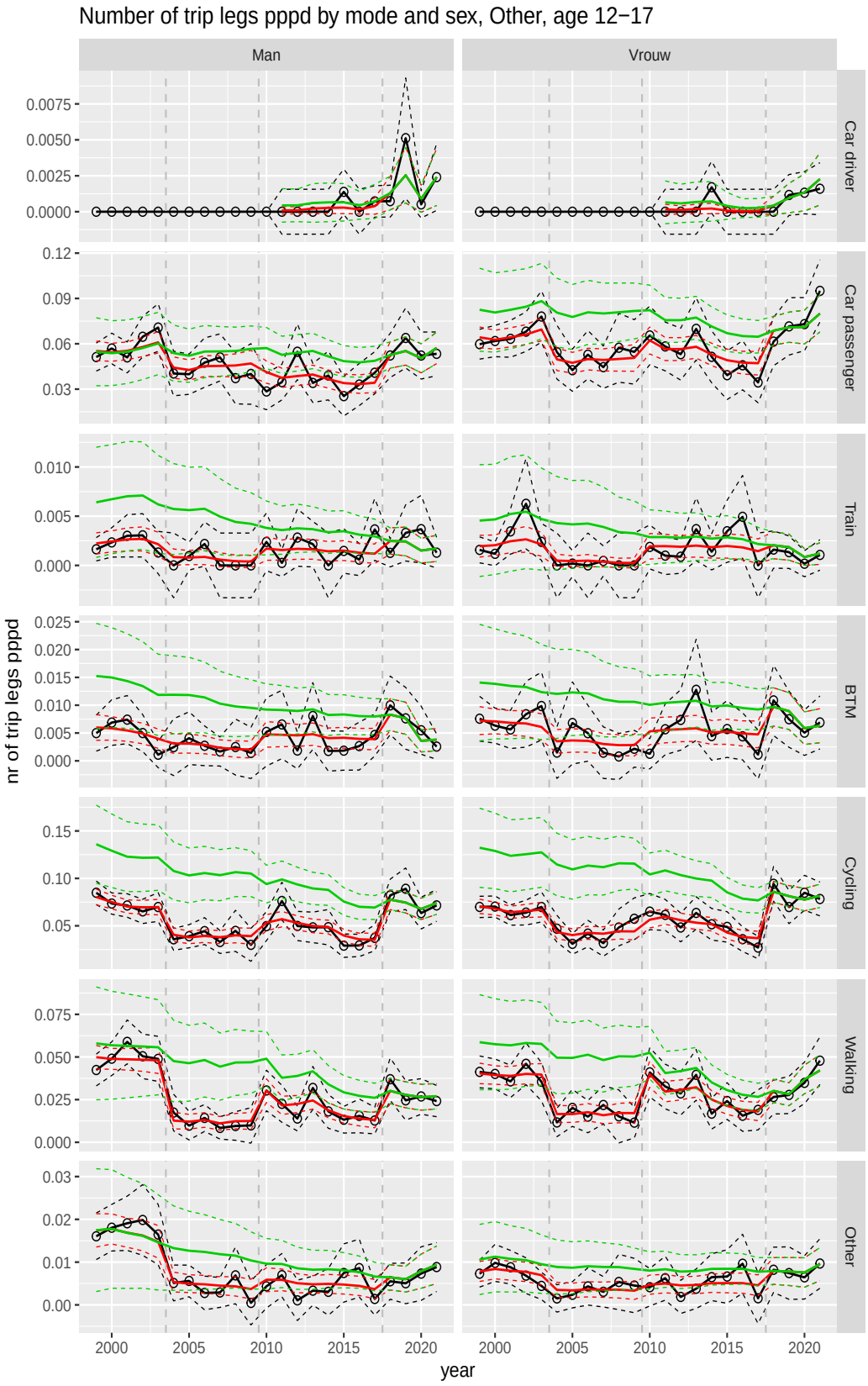

Figure A.79 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

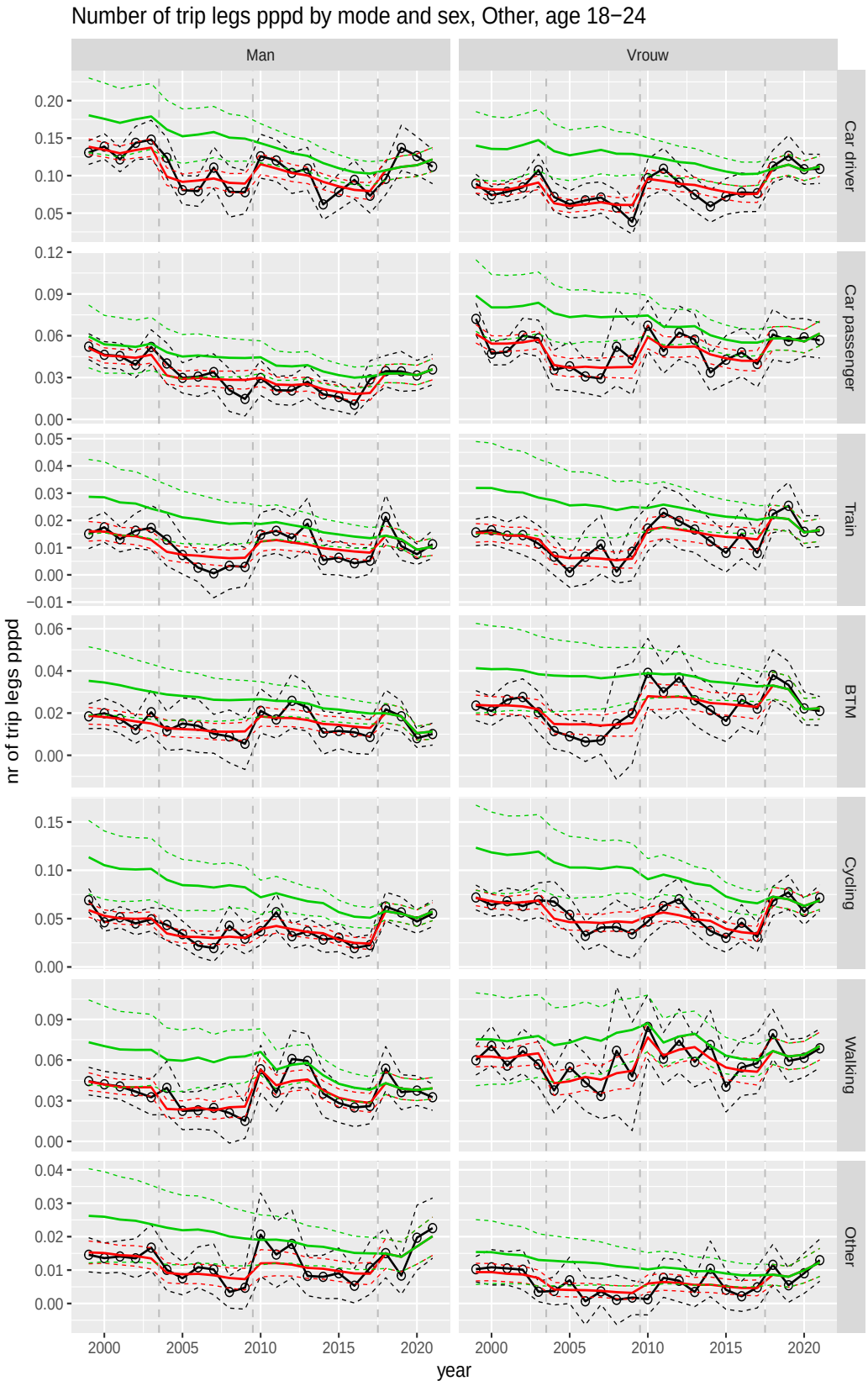

Figure A.80 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

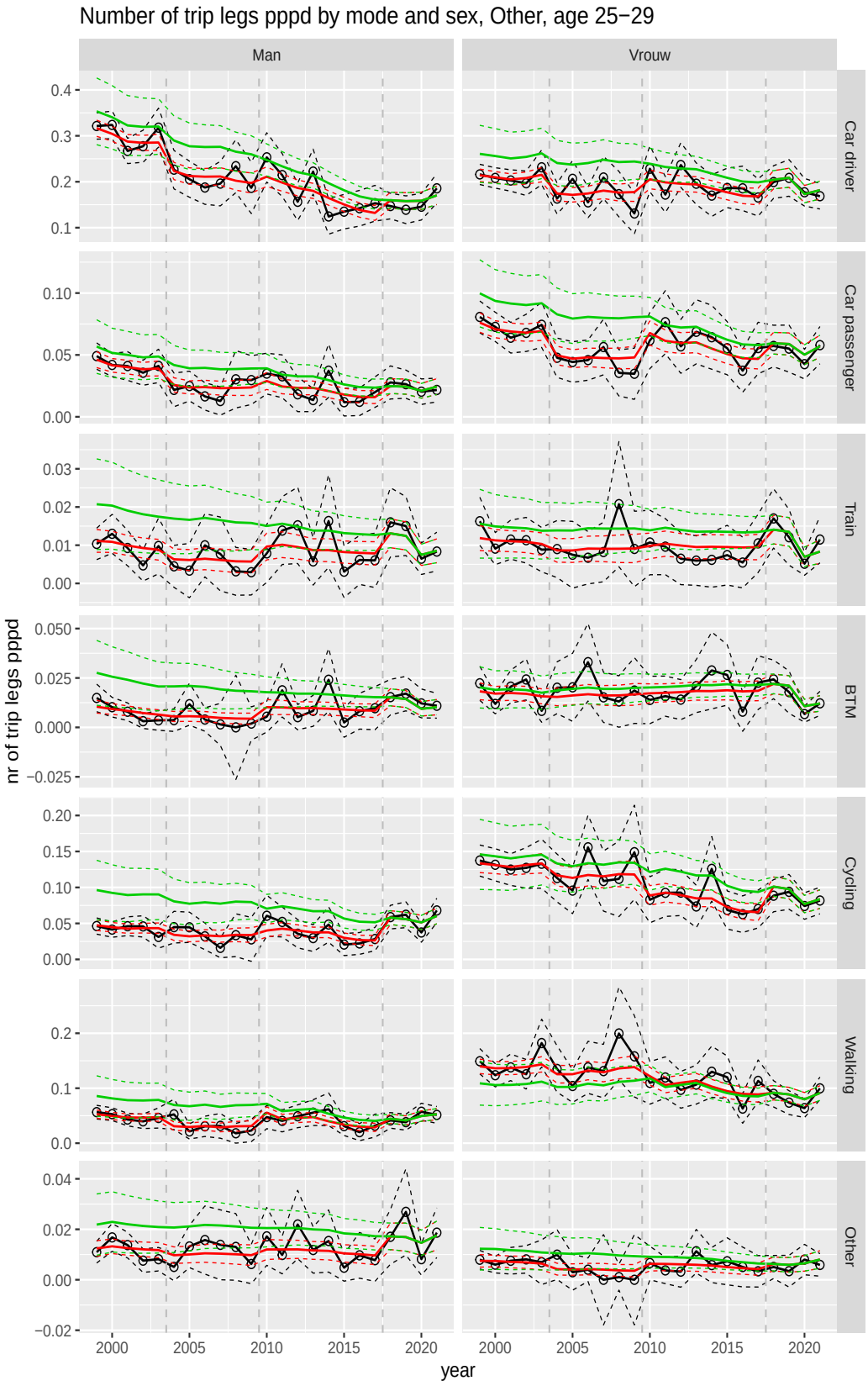

Figure A.81 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

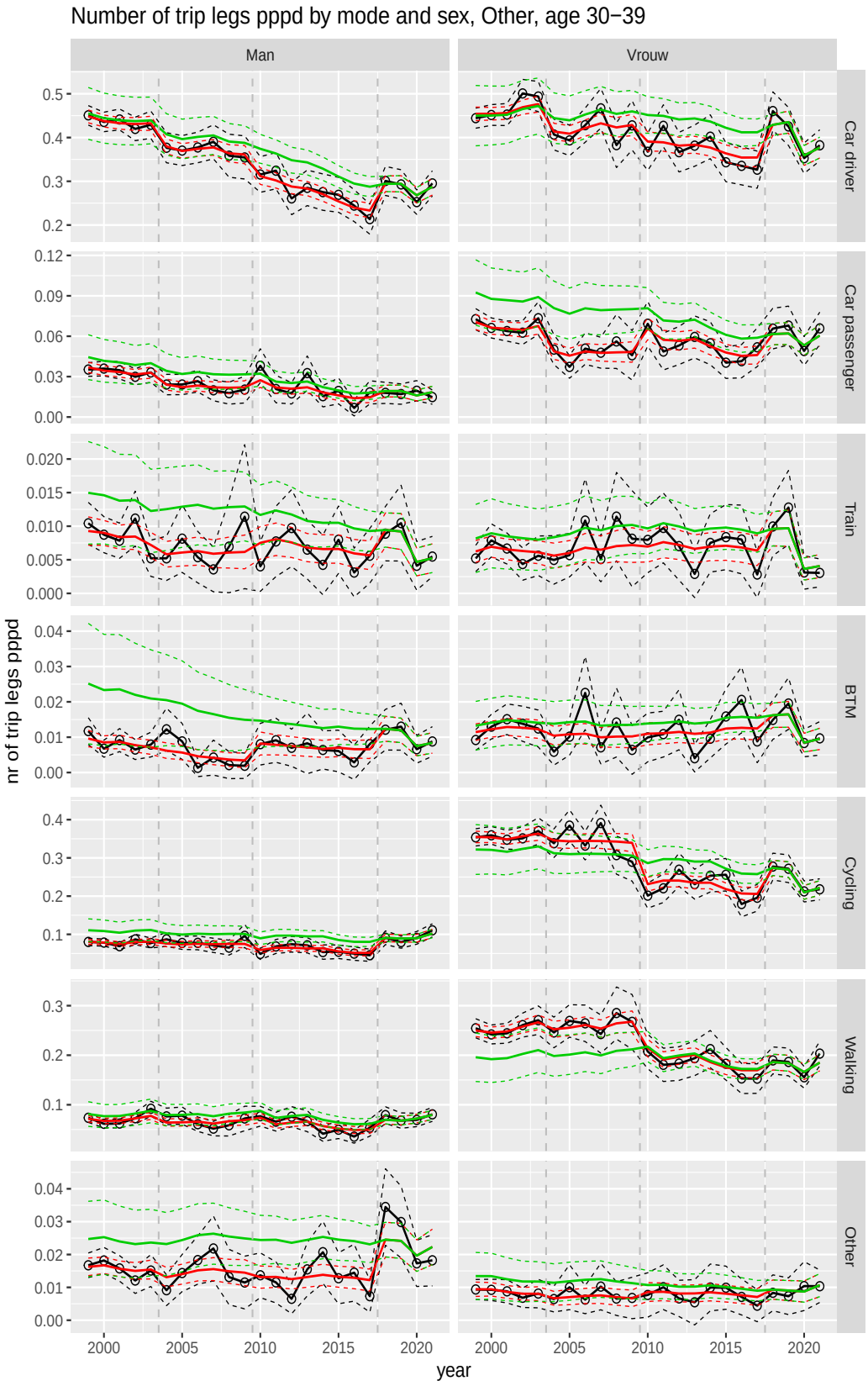

Figure A.82 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

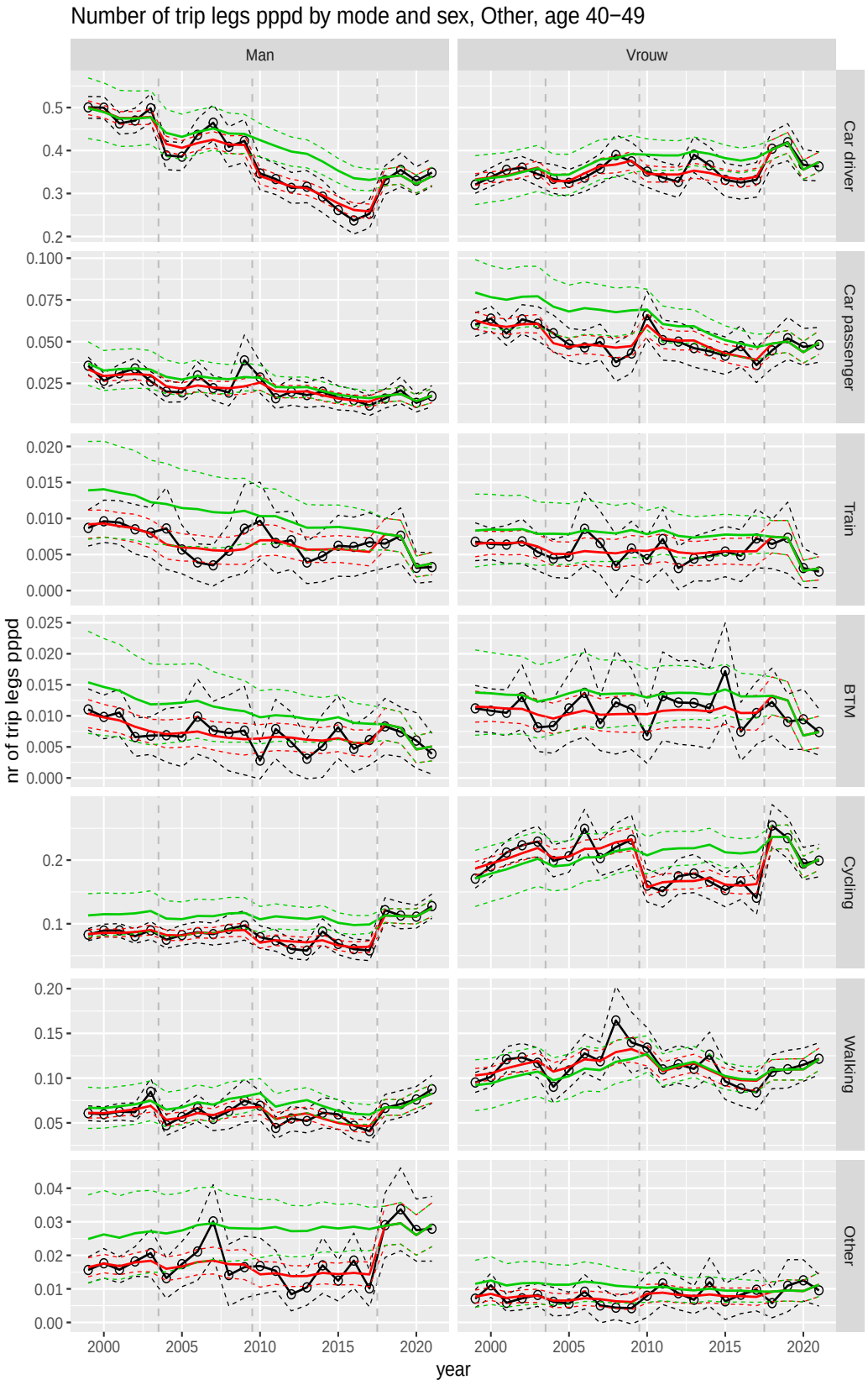

Figure A.83 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

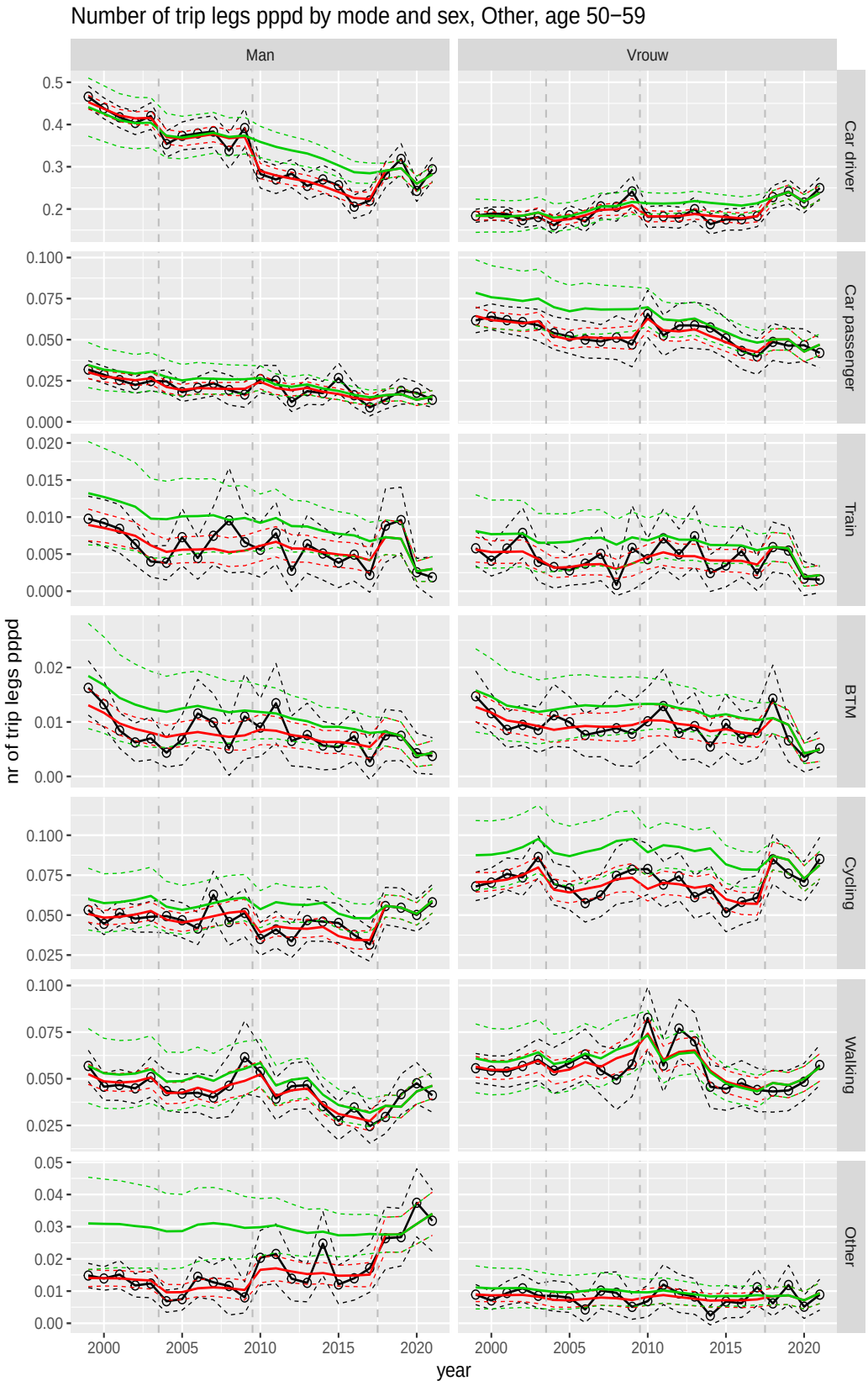

Figure A.84 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

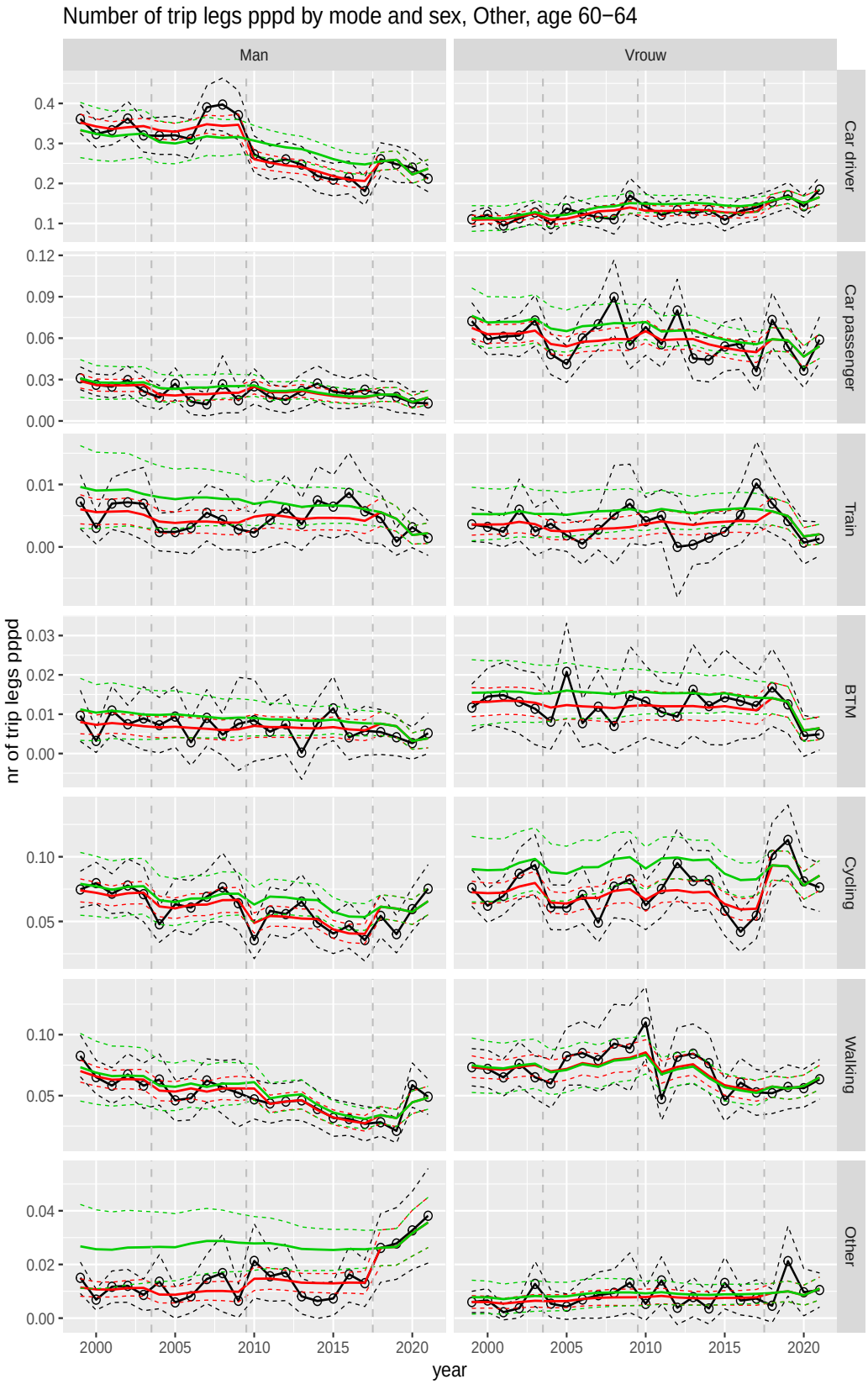

Figure A.85 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

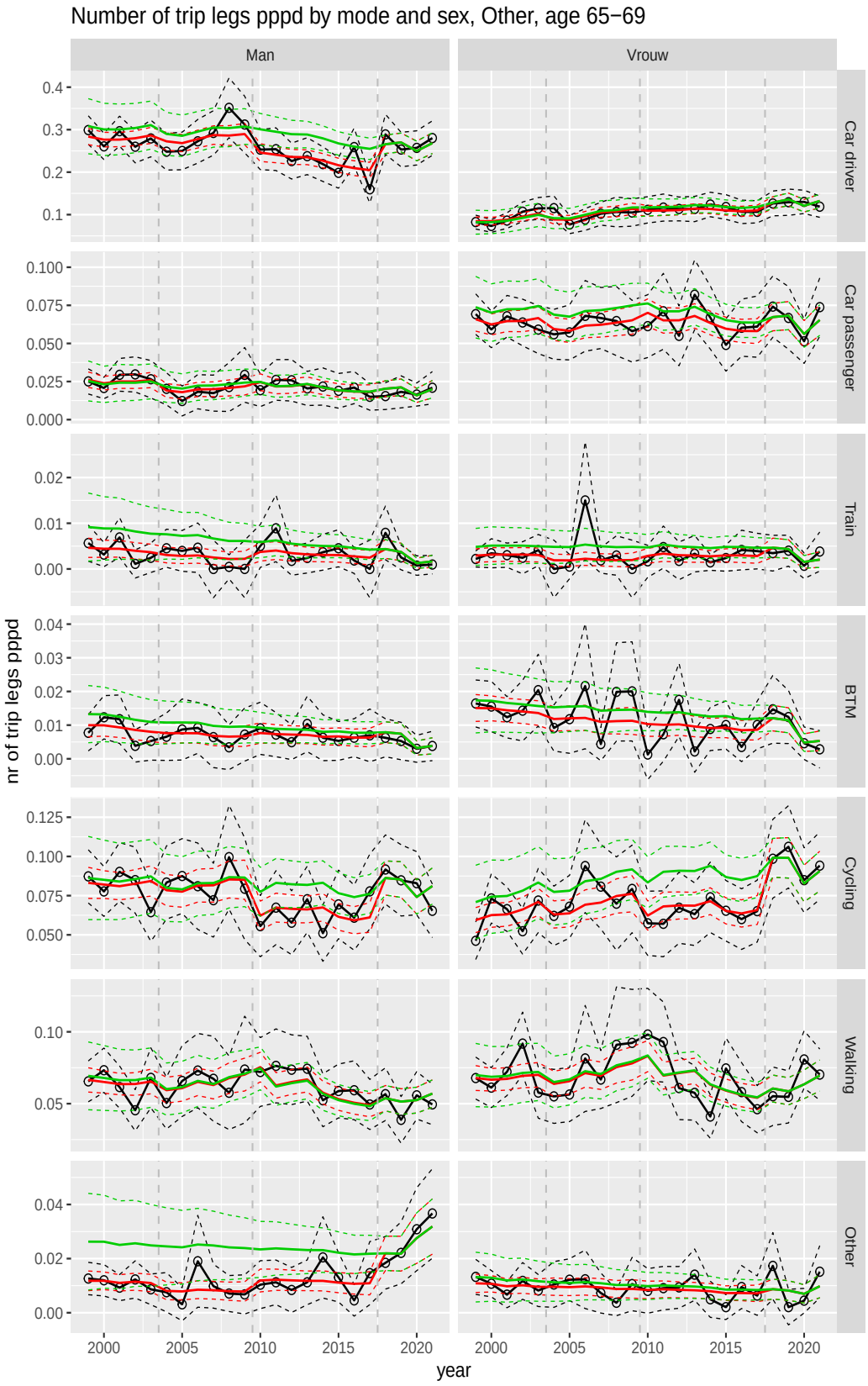

Figure A.86 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

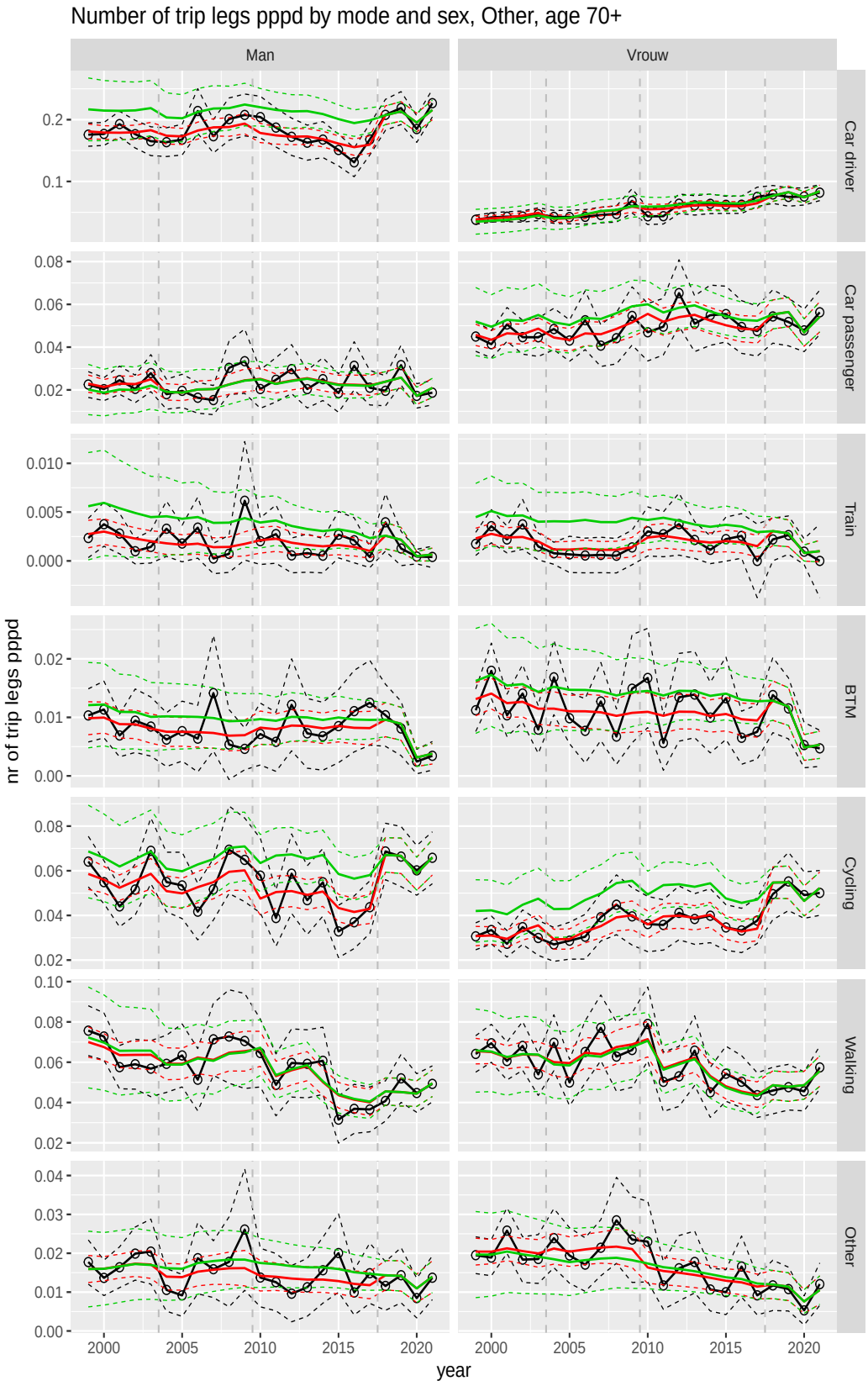

Figure A.87 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

A.4 Average distance per trip leg

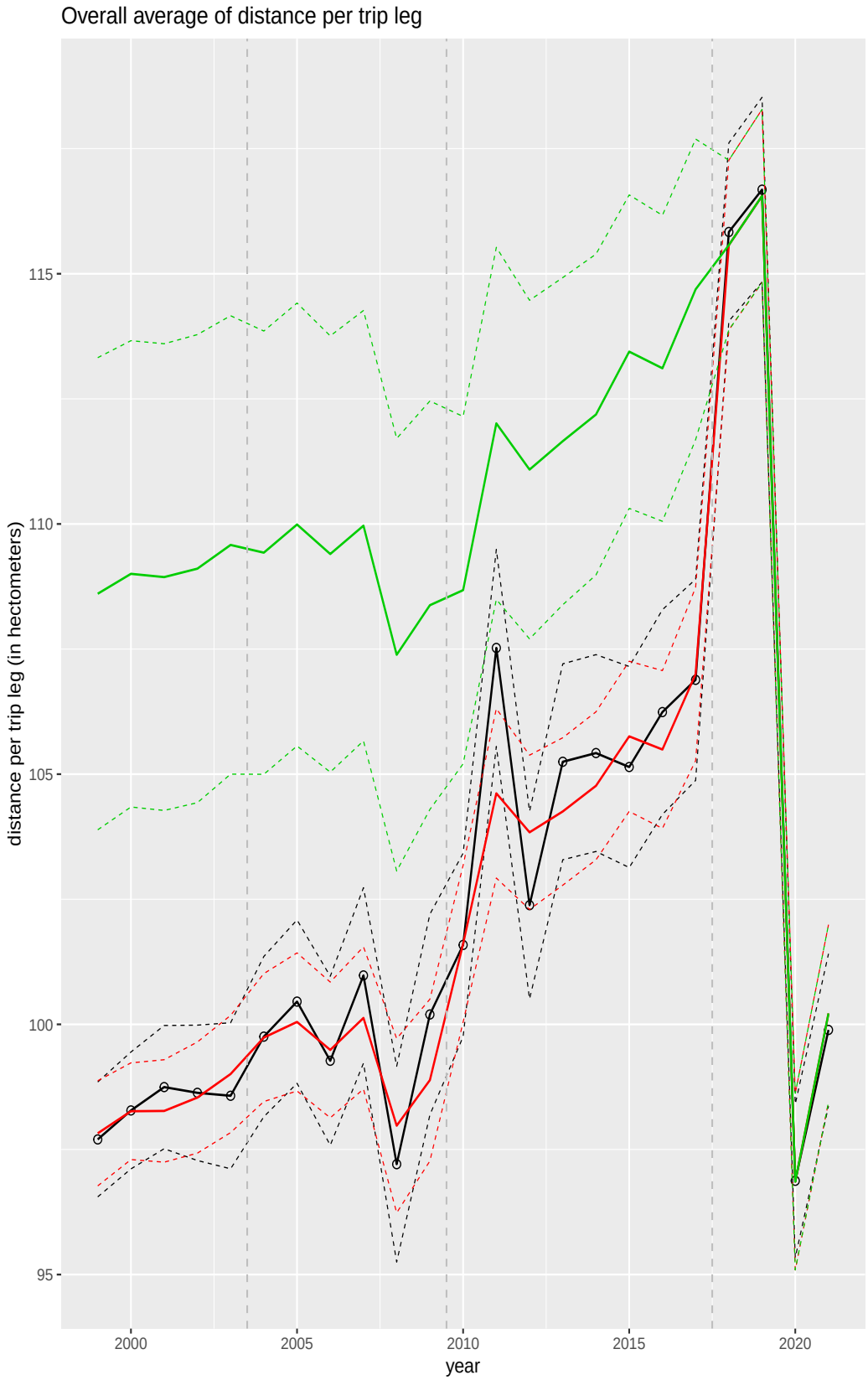

Figure A.88 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

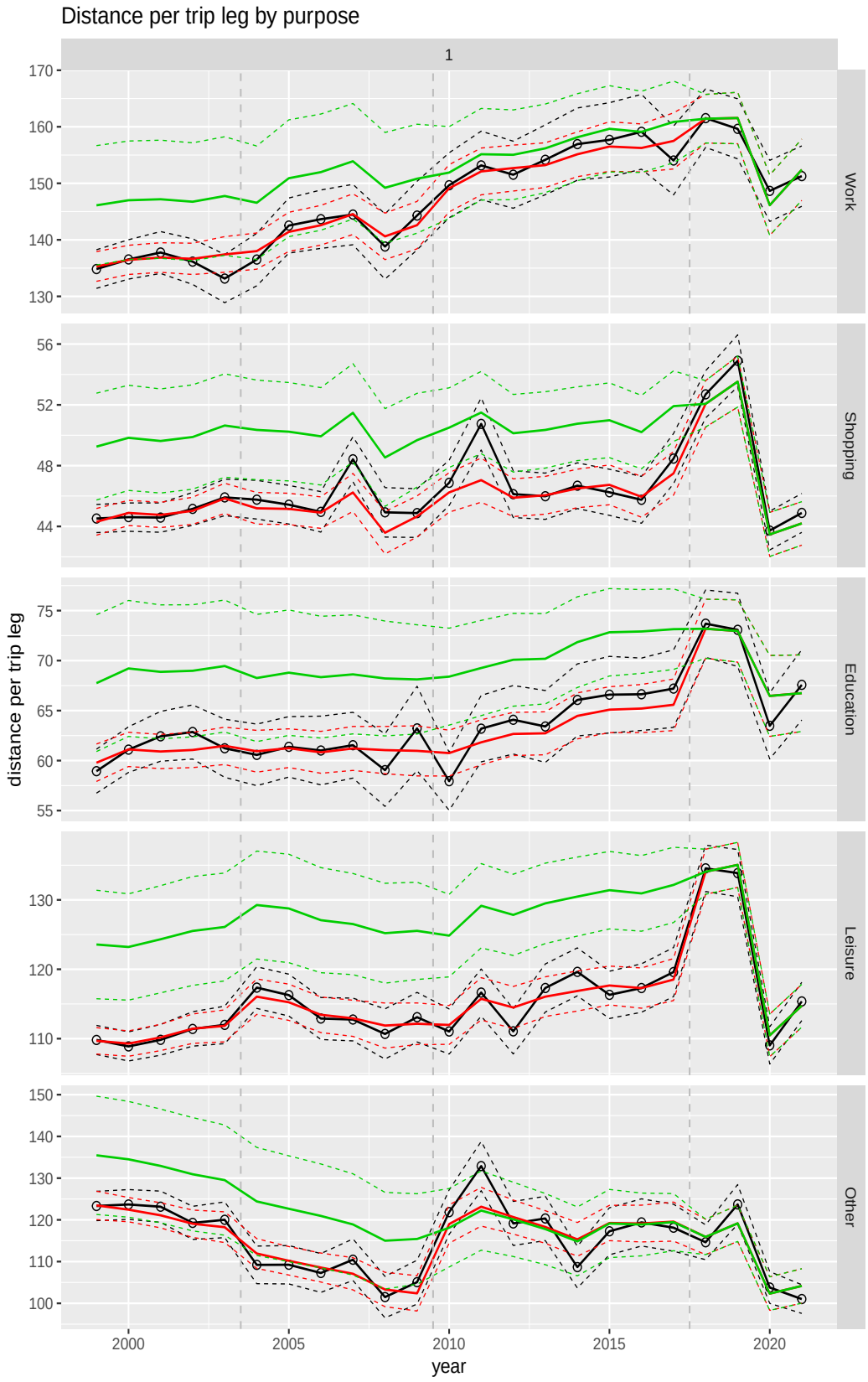

Figure A.89 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

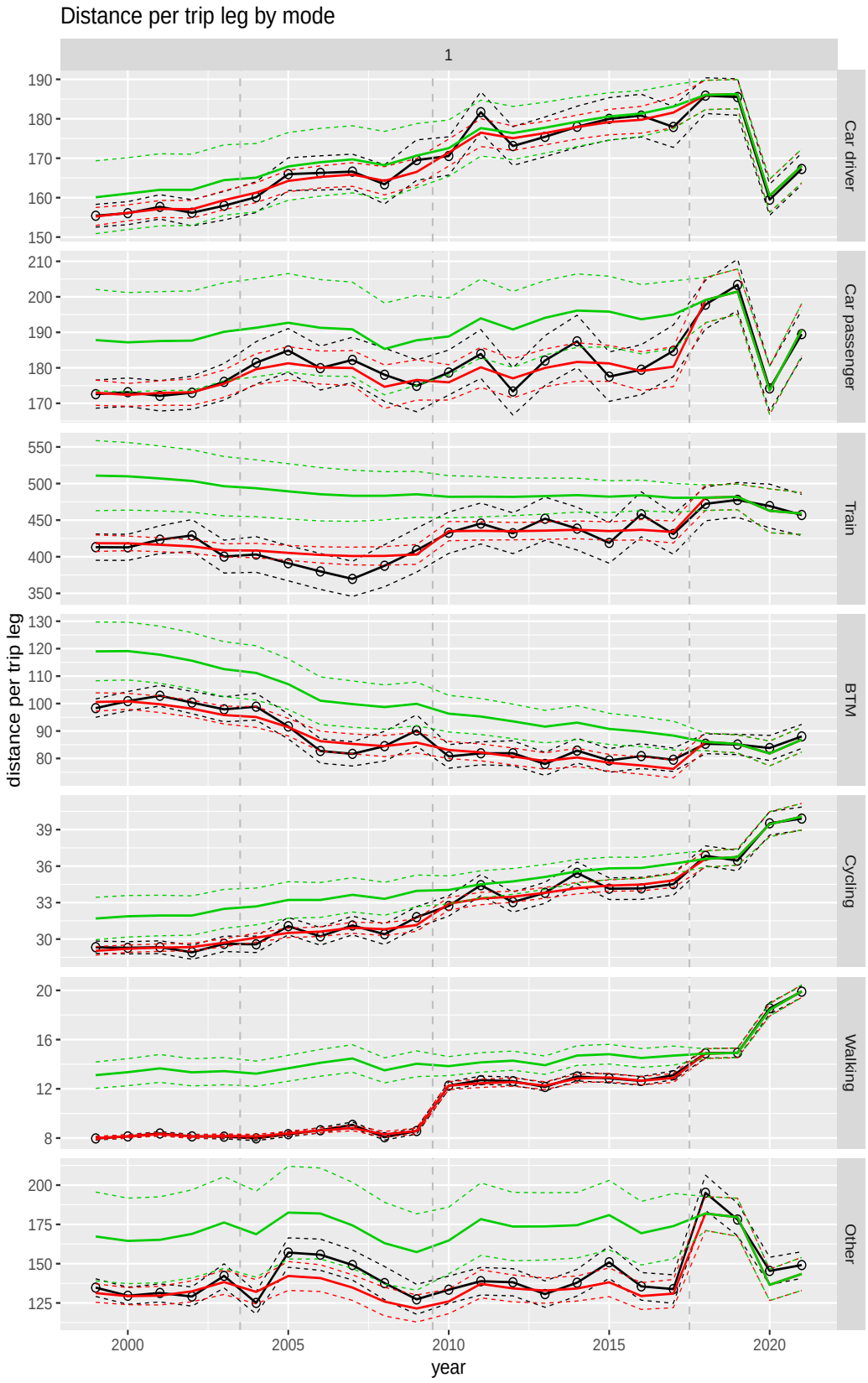

Figure A.90 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

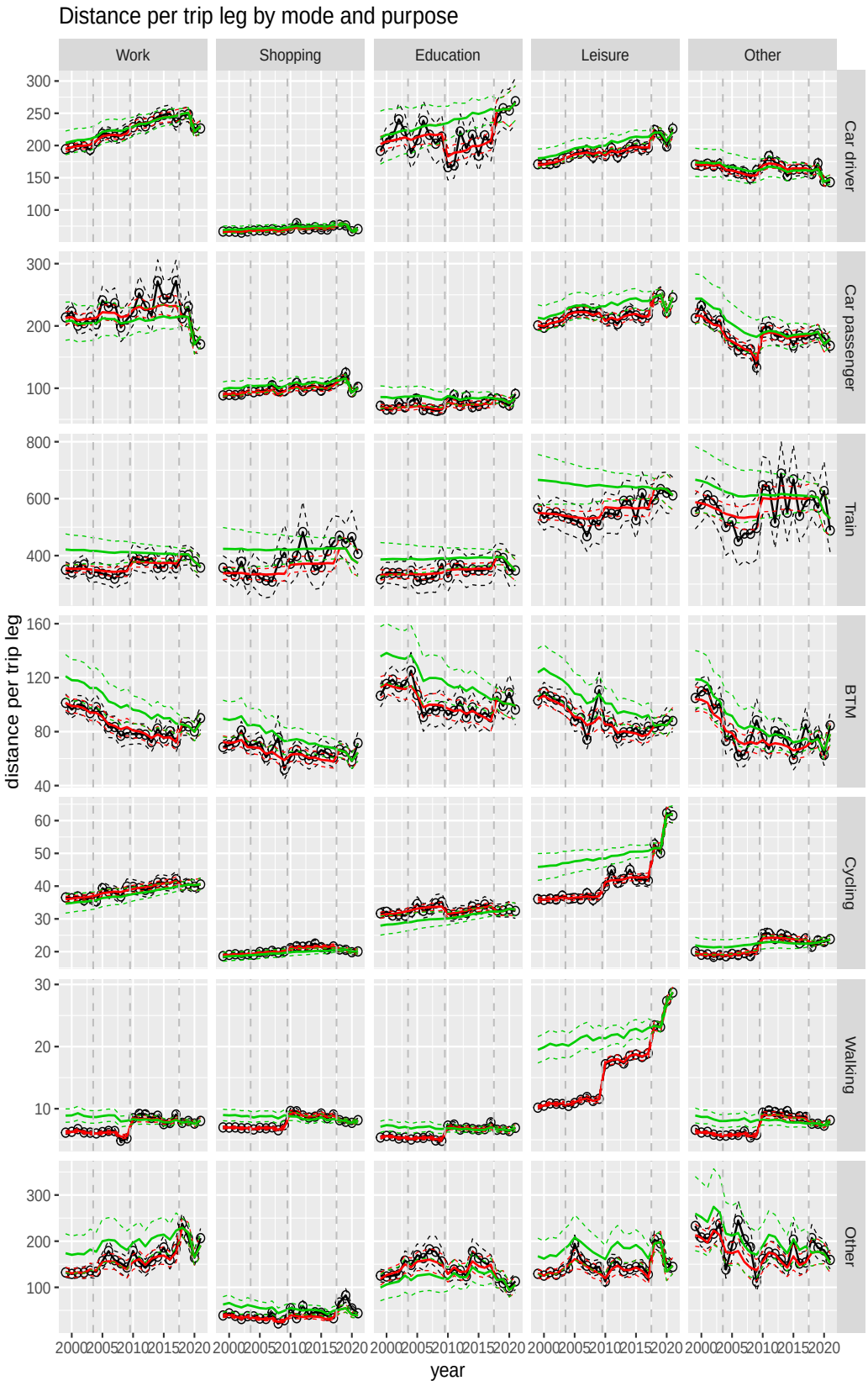

Figure A.91 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

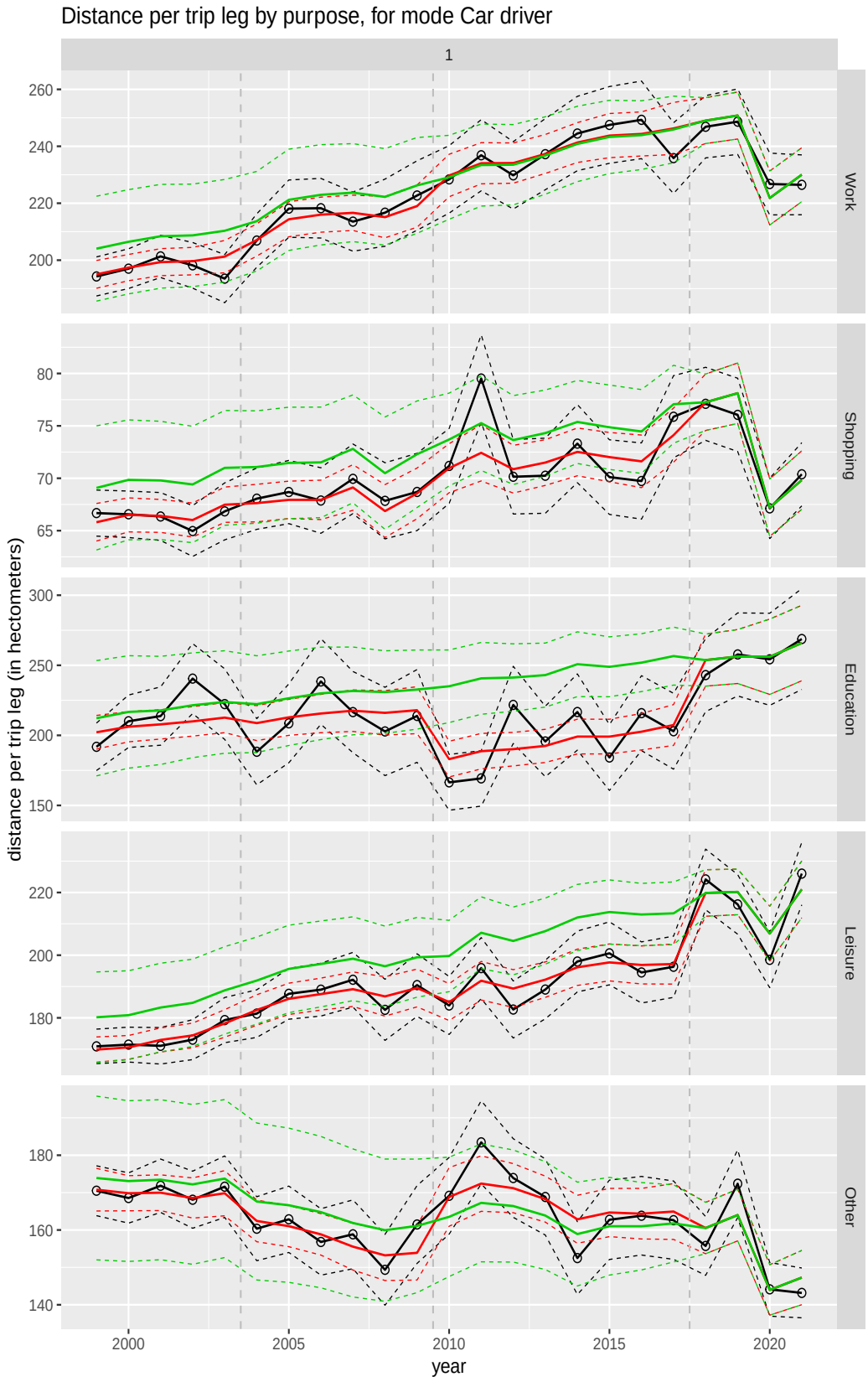

Figure A.92 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

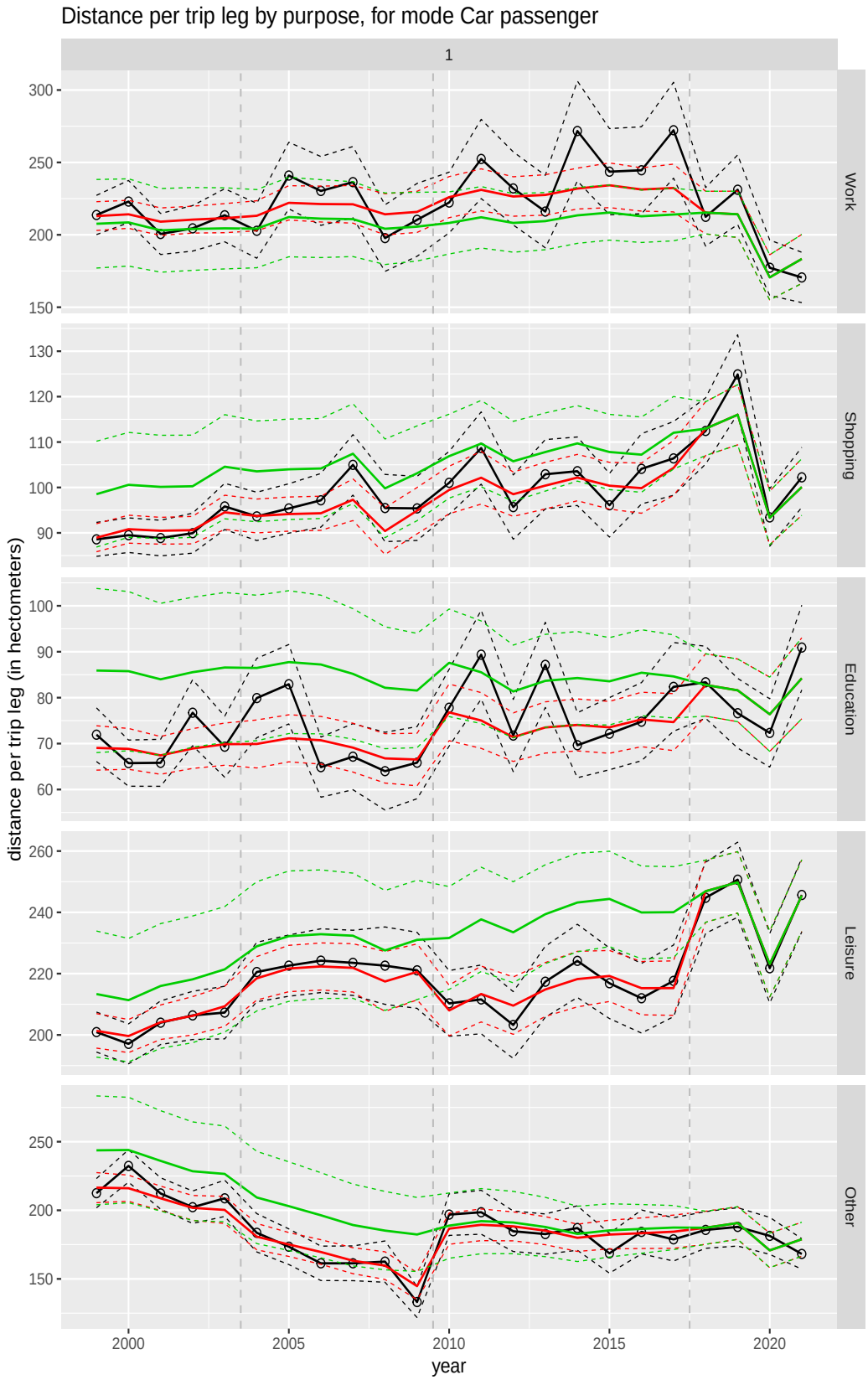

Figure A.93 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

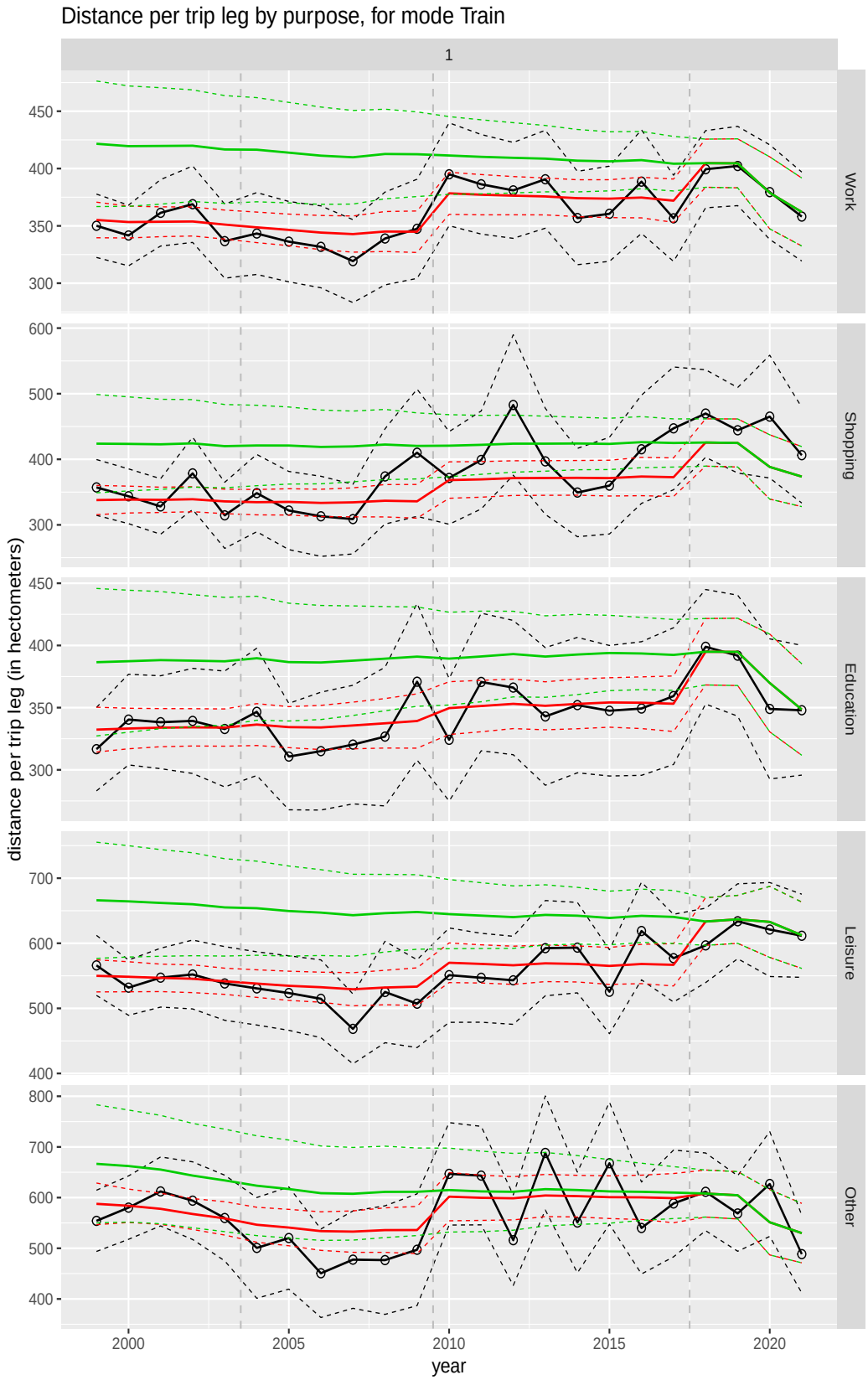

Figure A.94 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.95 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

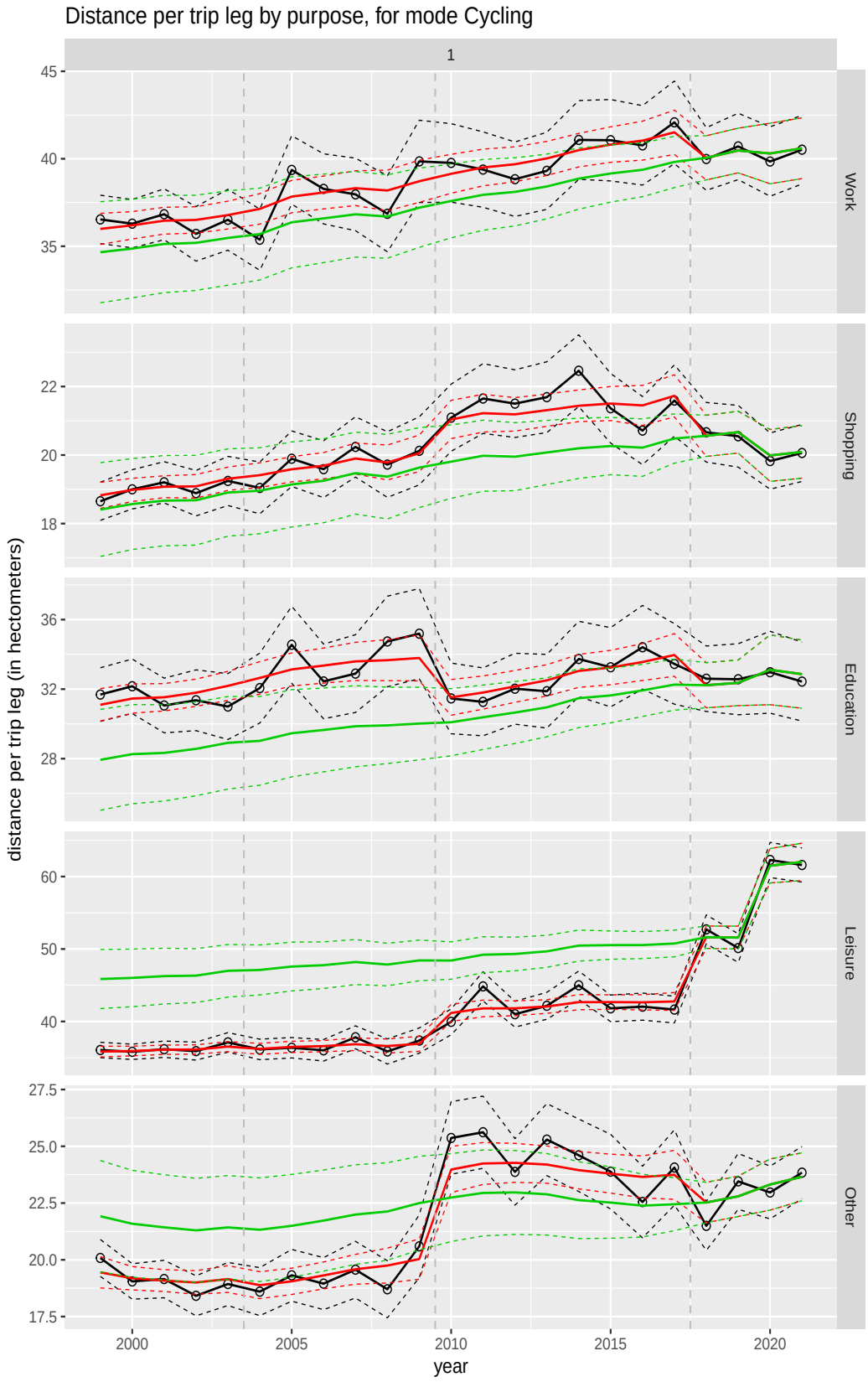

Figure A.96 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

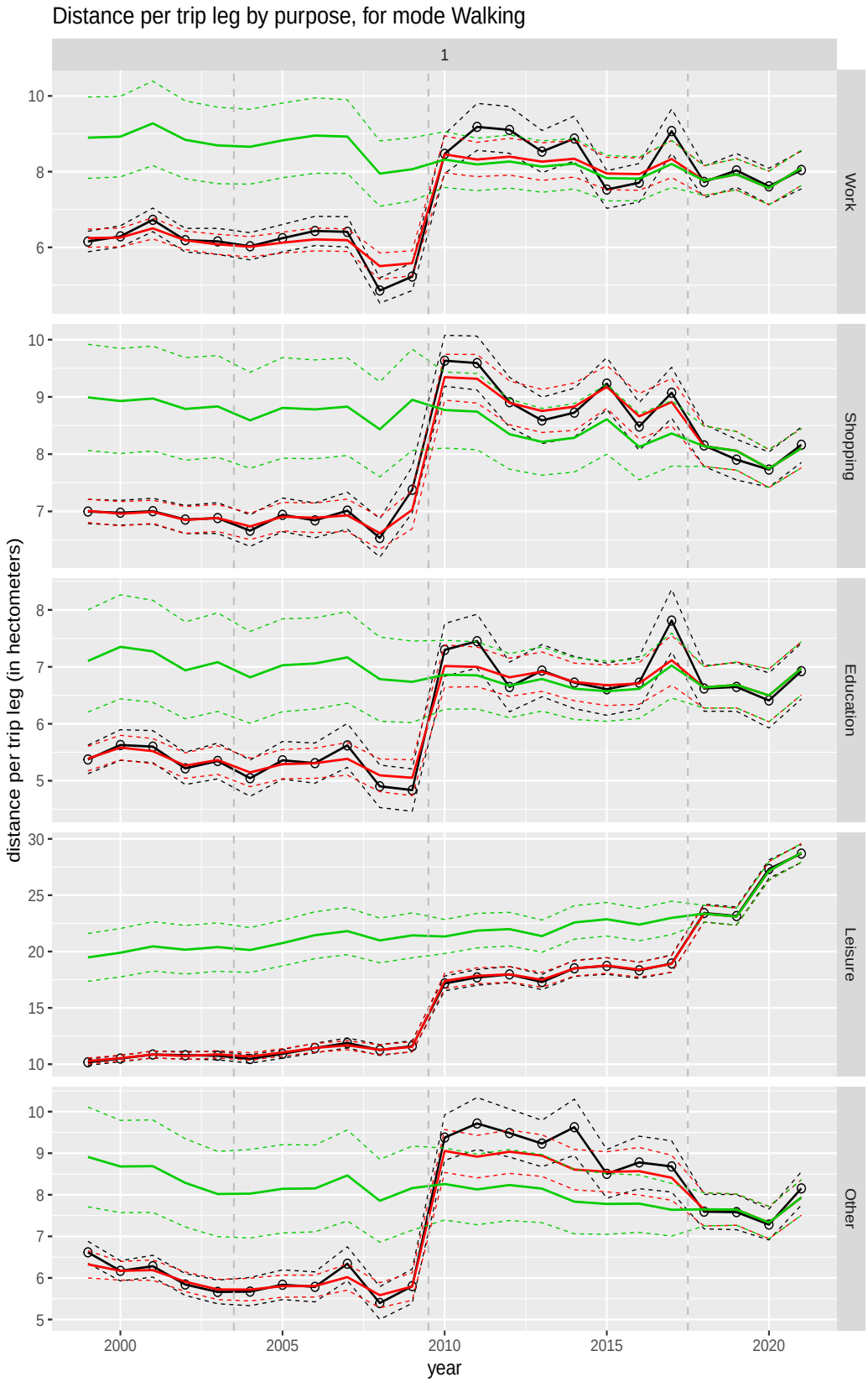

Figure A.97 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

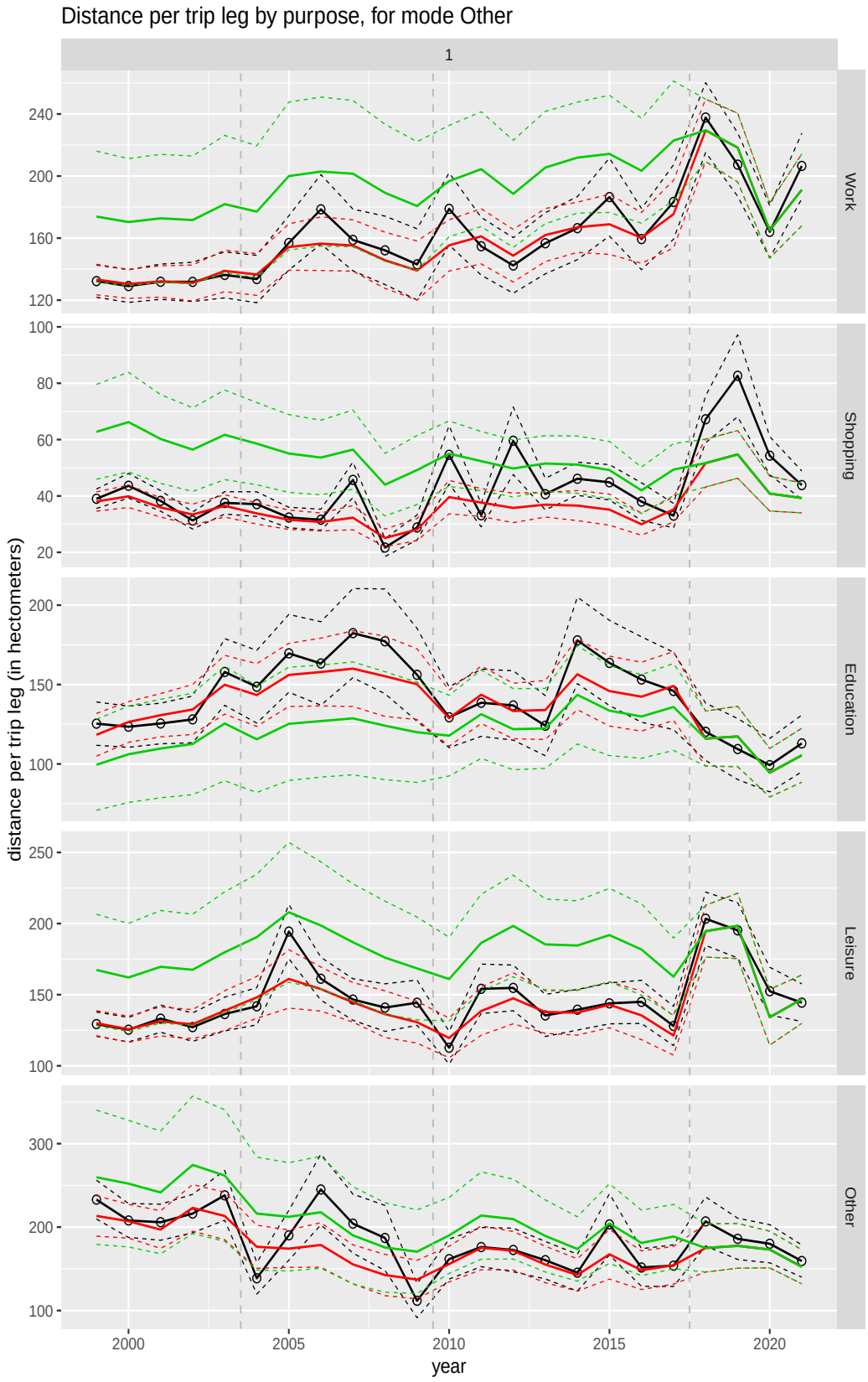

Figure A.98 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

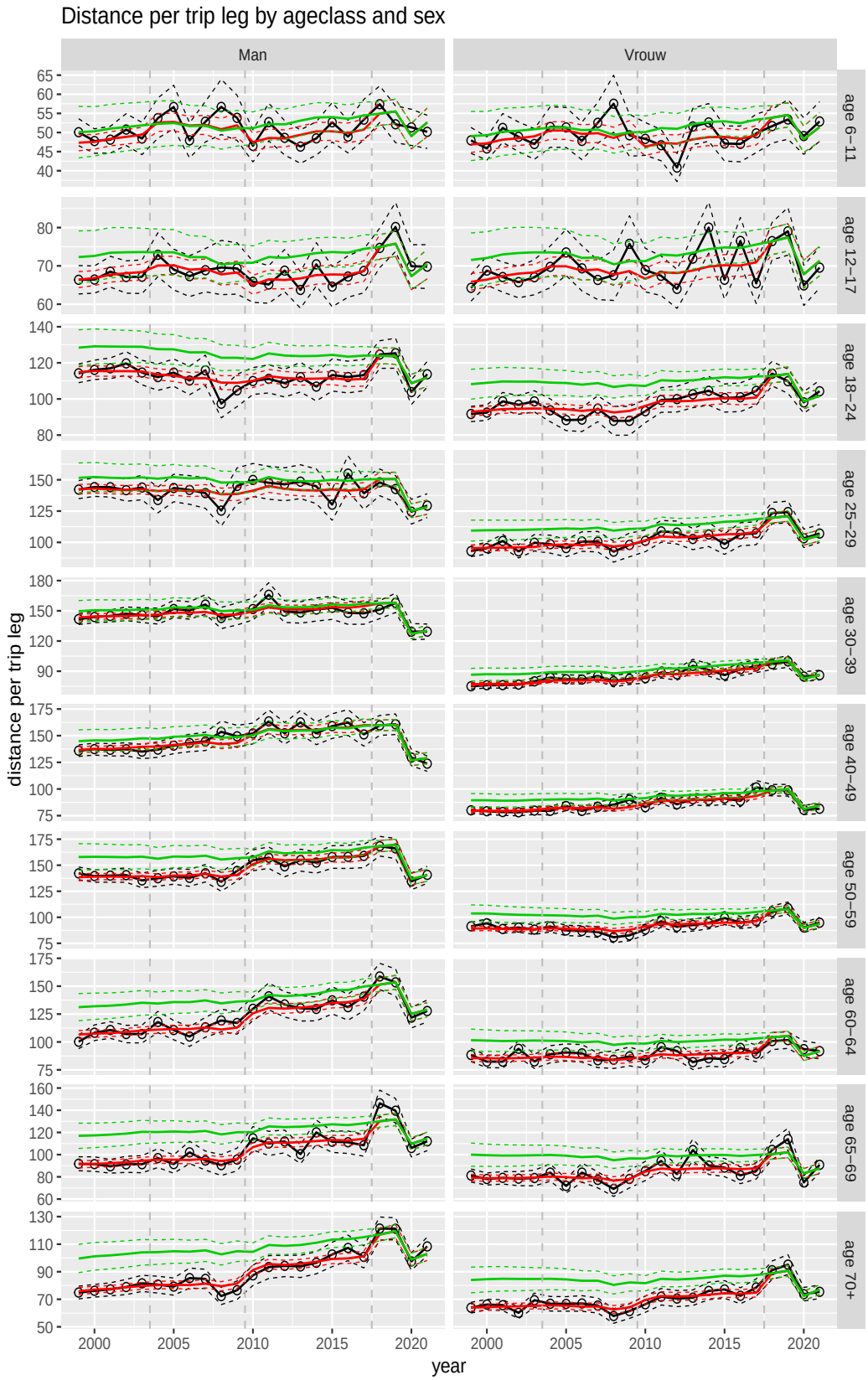

Figure A.99 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.100 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.101 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.102 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

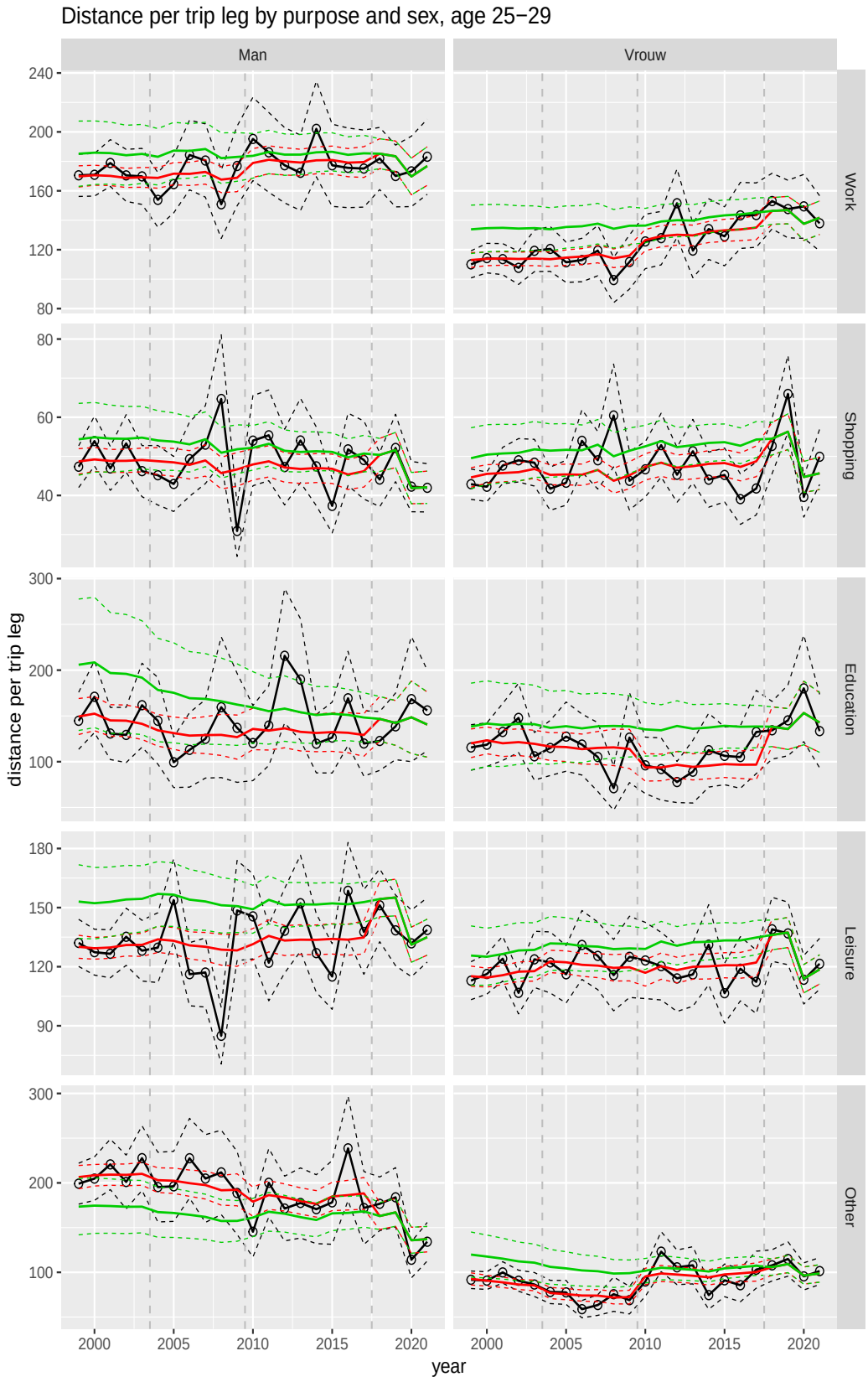

Figure A.103 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

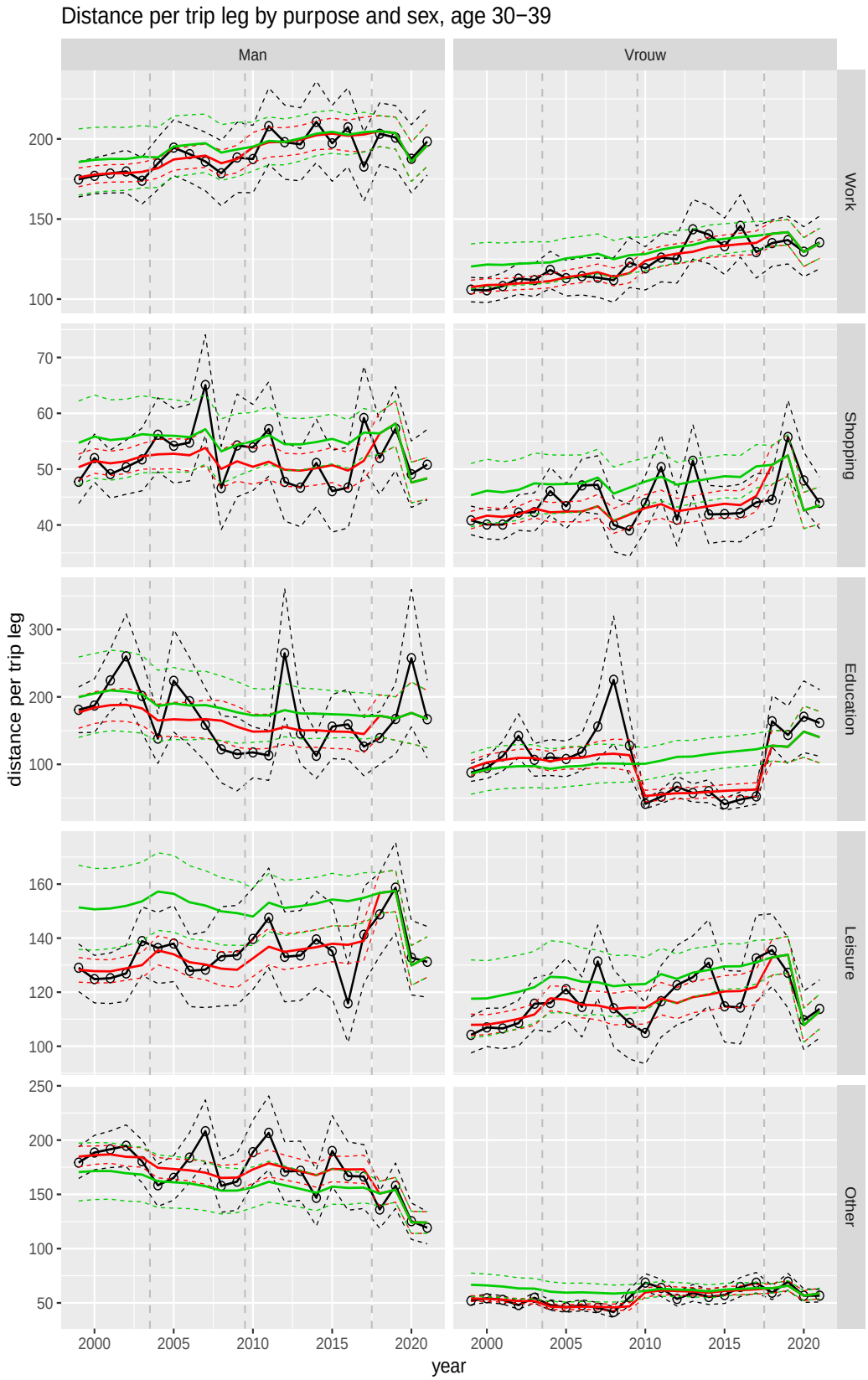

Figure A.104 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.105 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.106 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.107 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

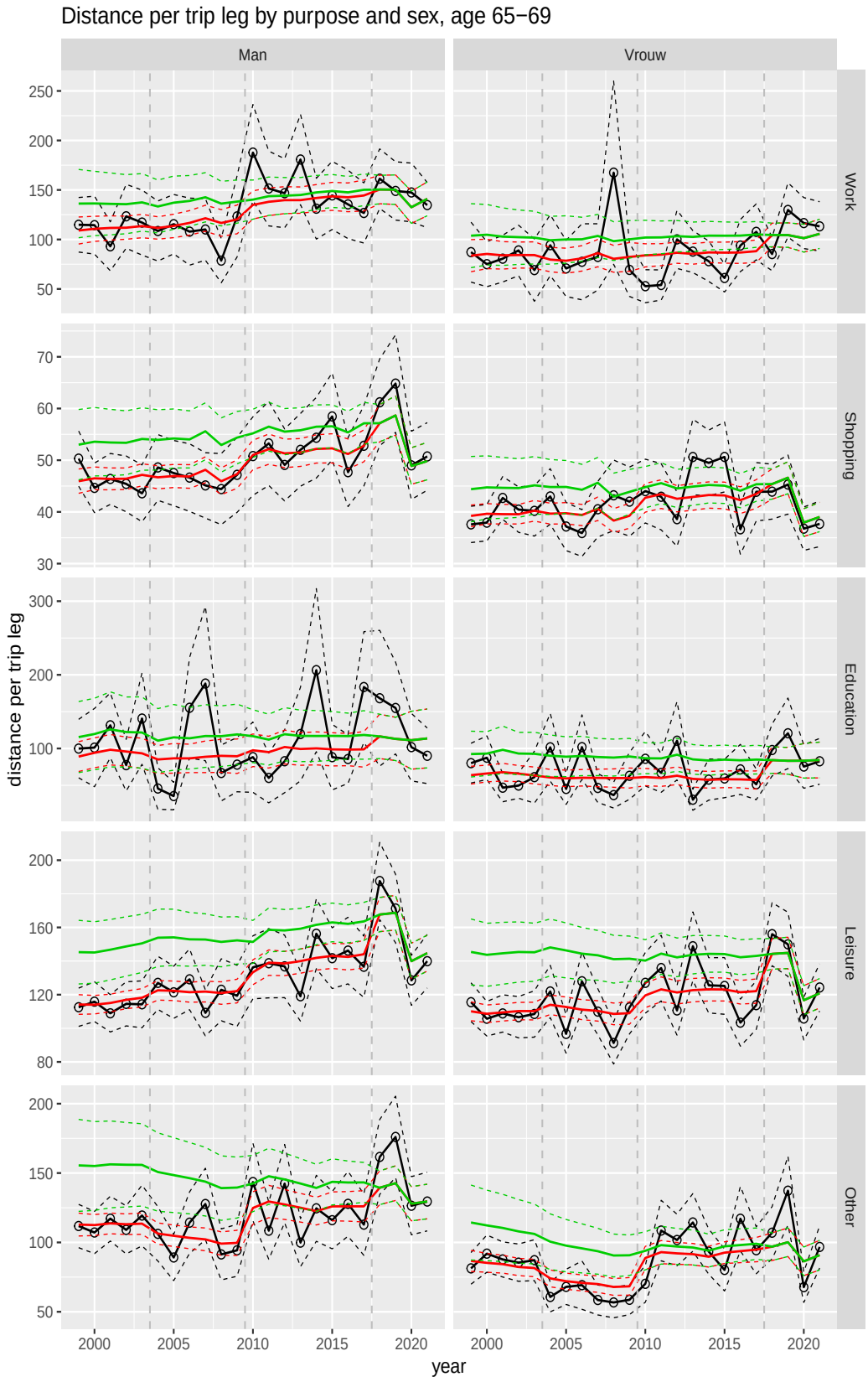

Figure A.108 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

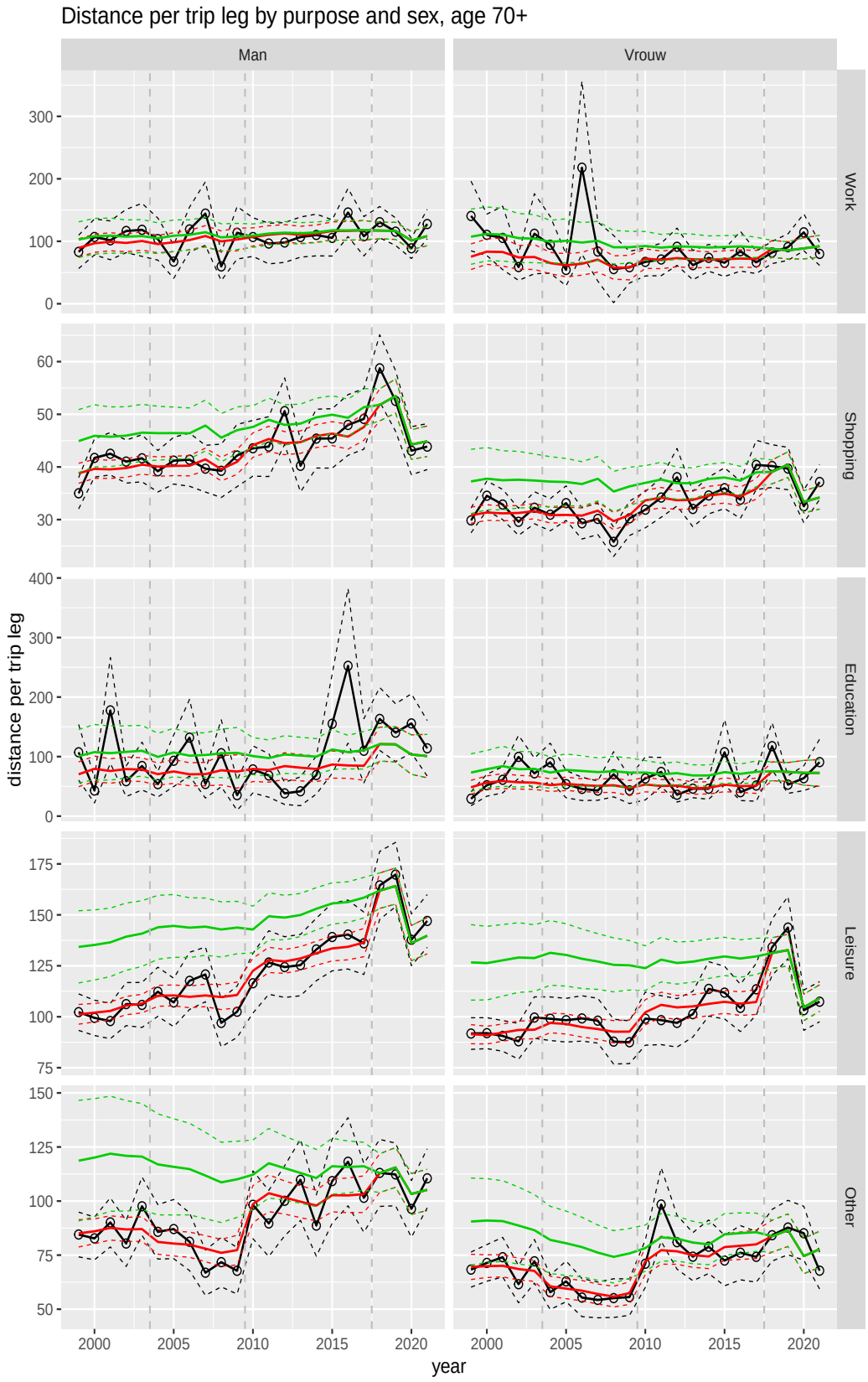

Figure A.109 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

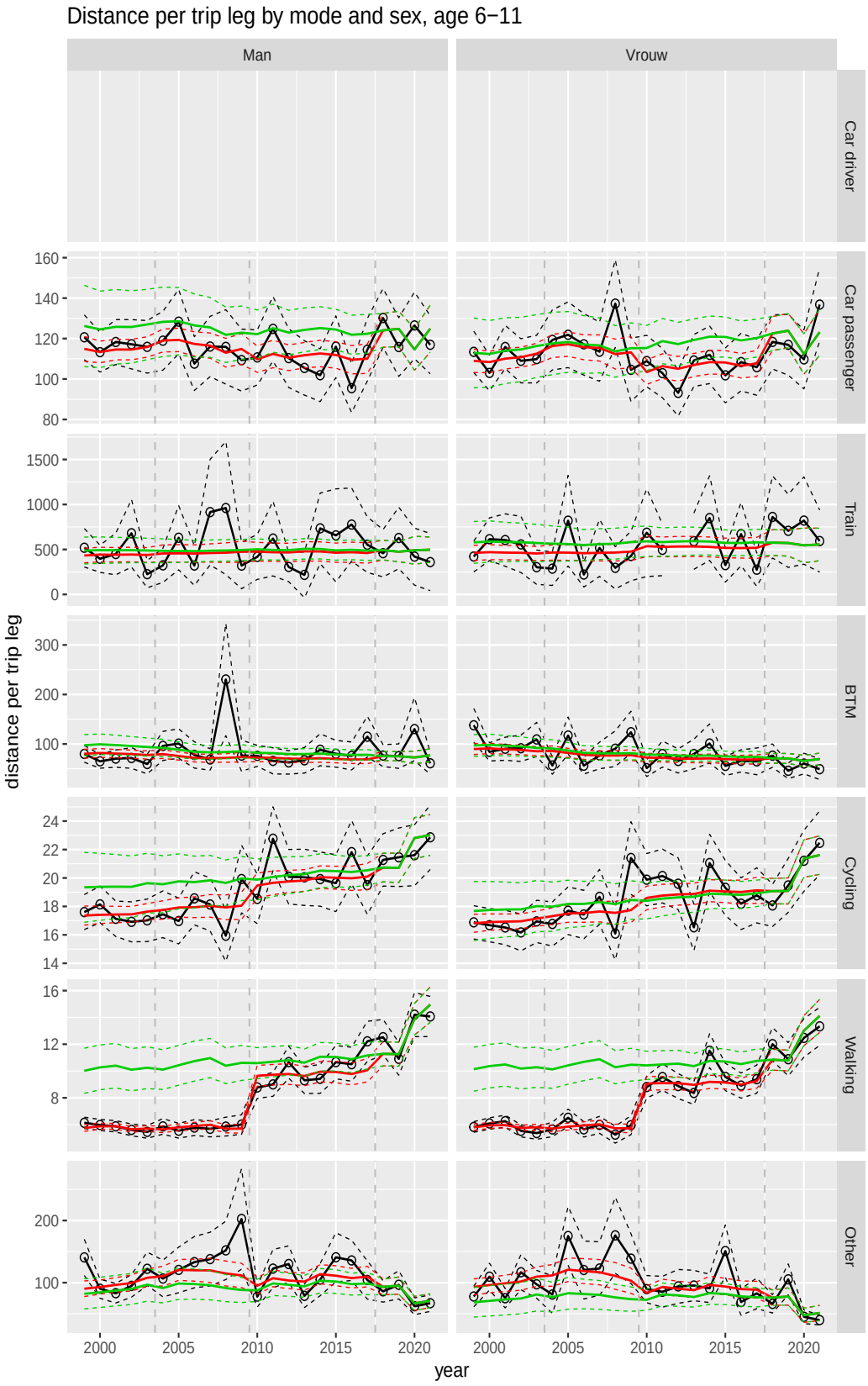

Figure A.110 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.111 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

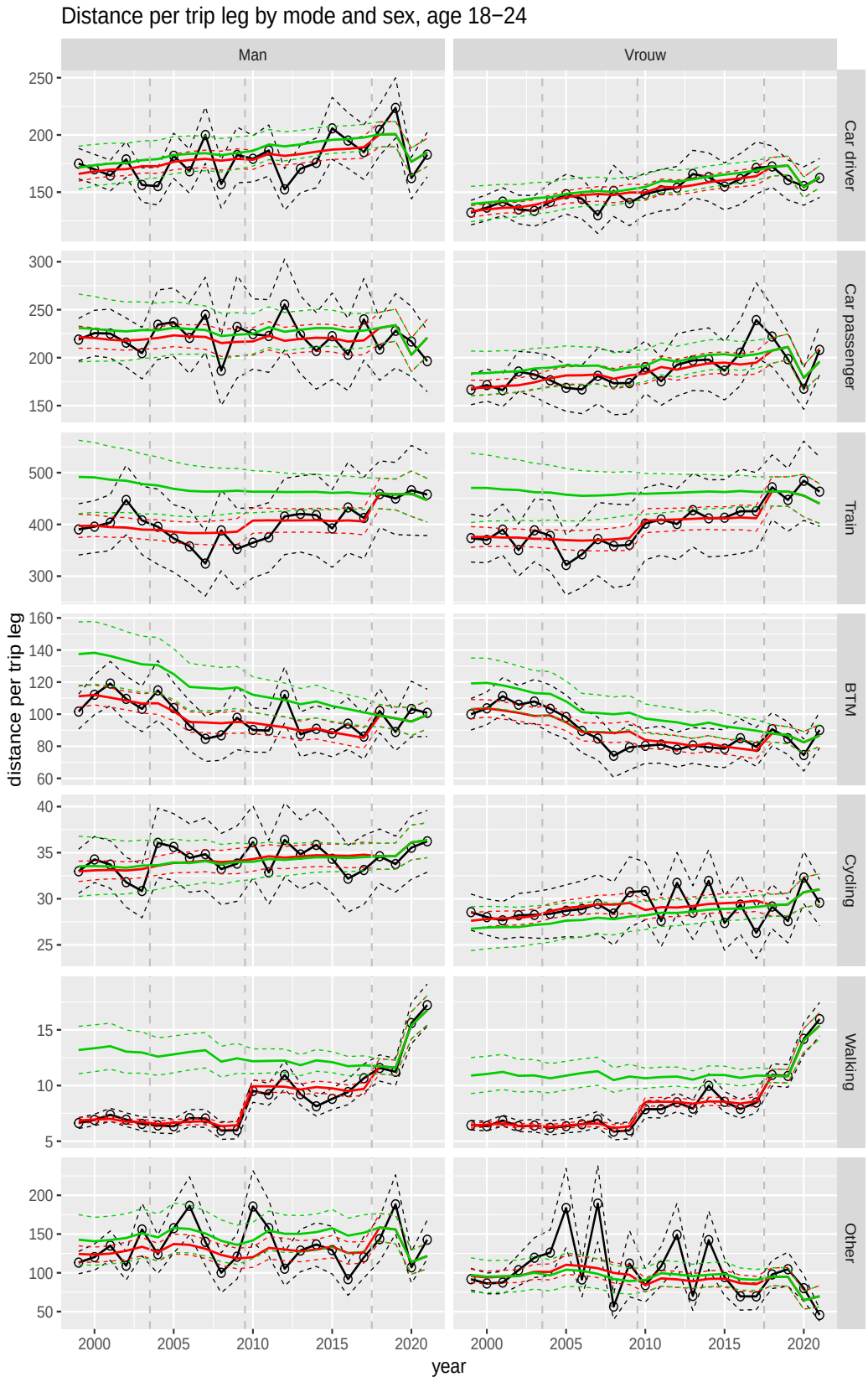

Figure A.112 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

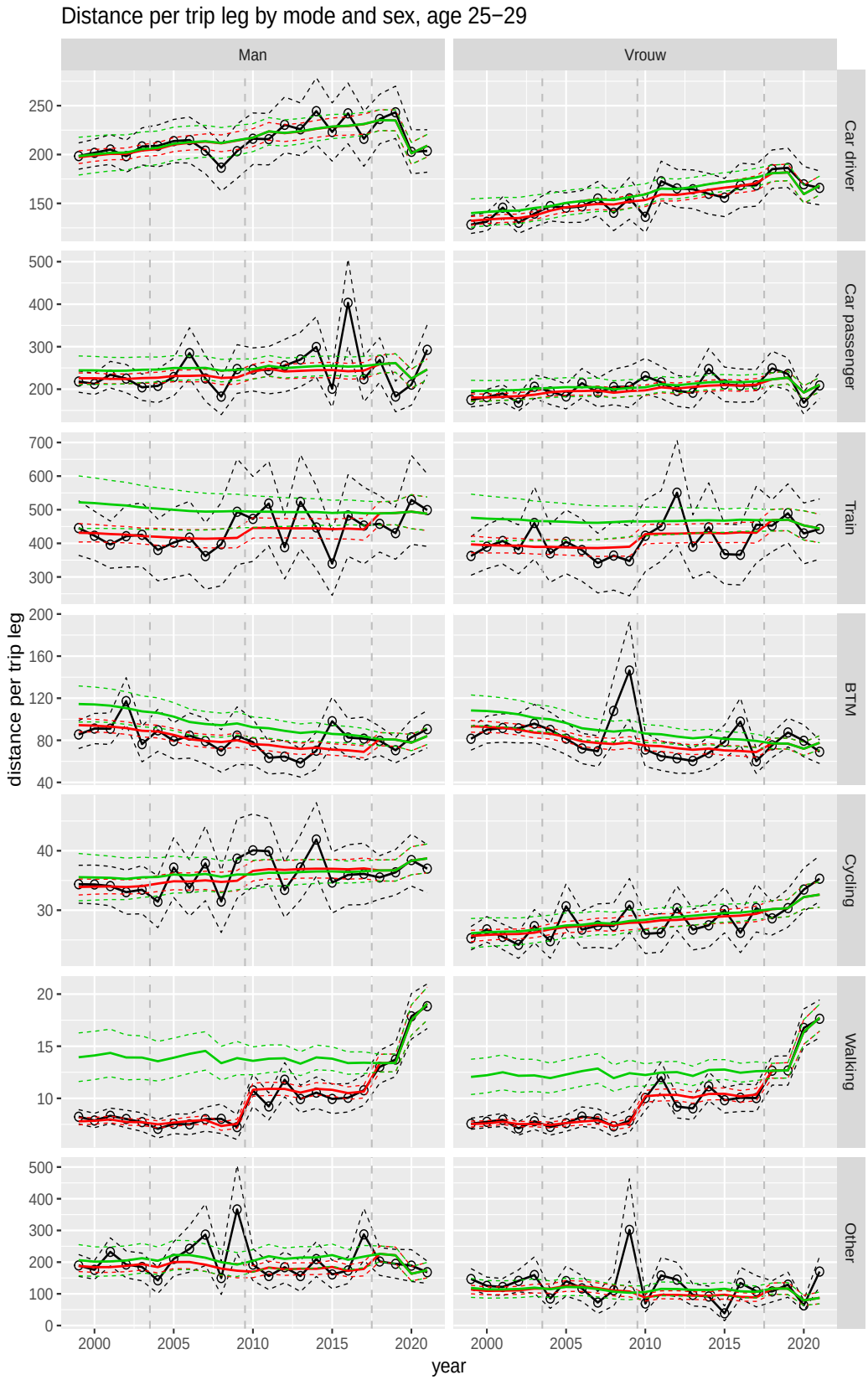

Figure A.113 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

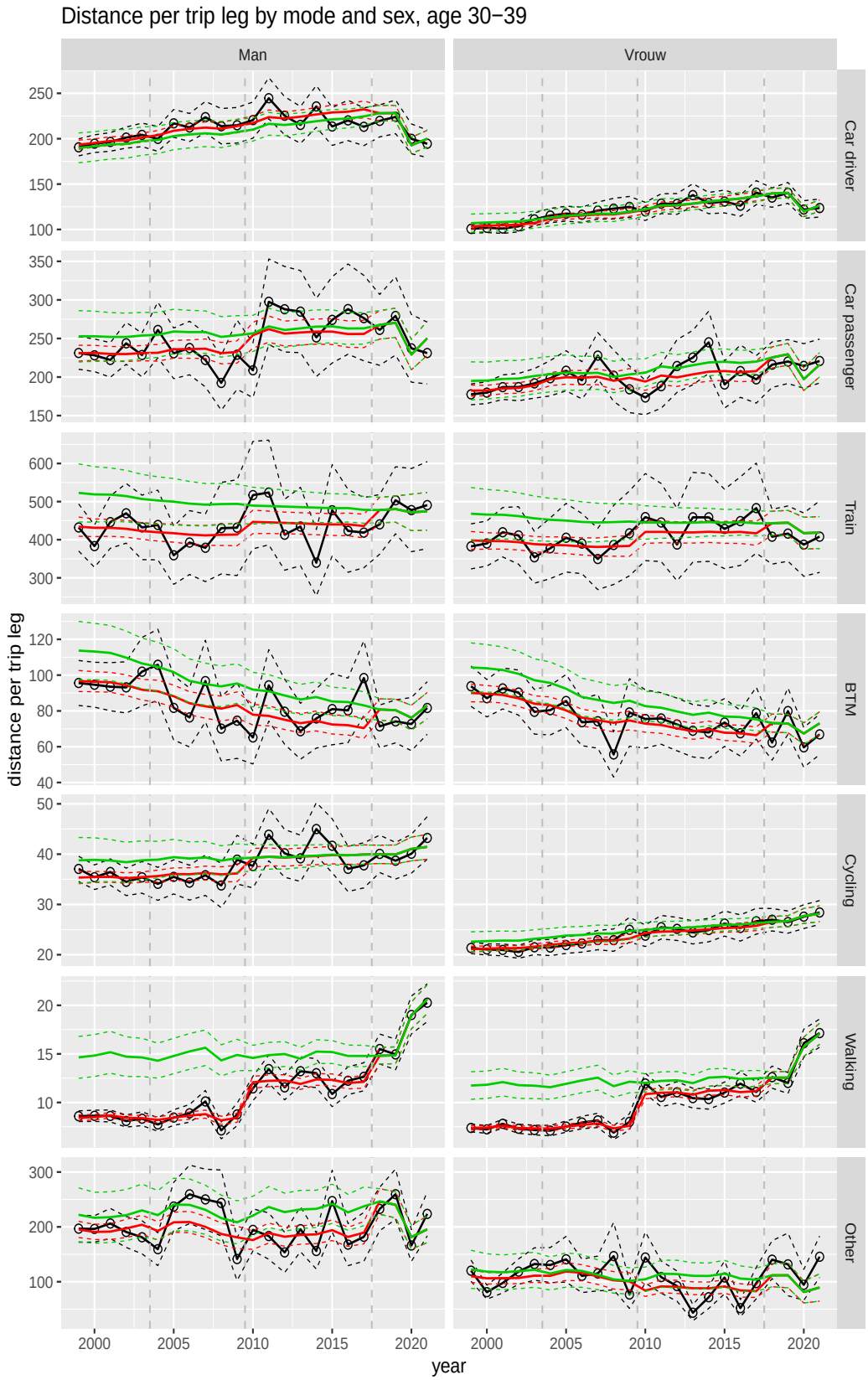

Figure A.114 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

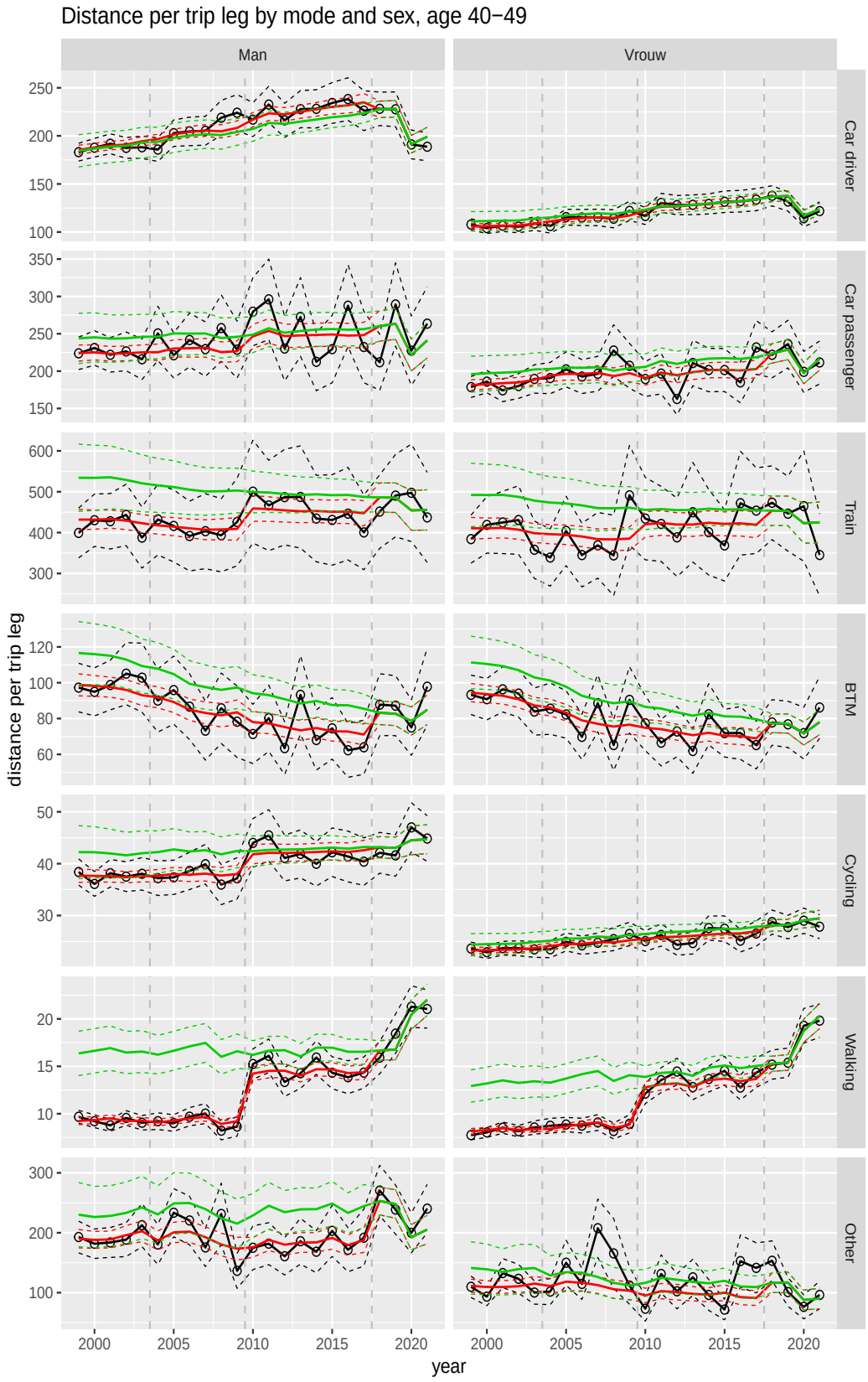

Figure A.115 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

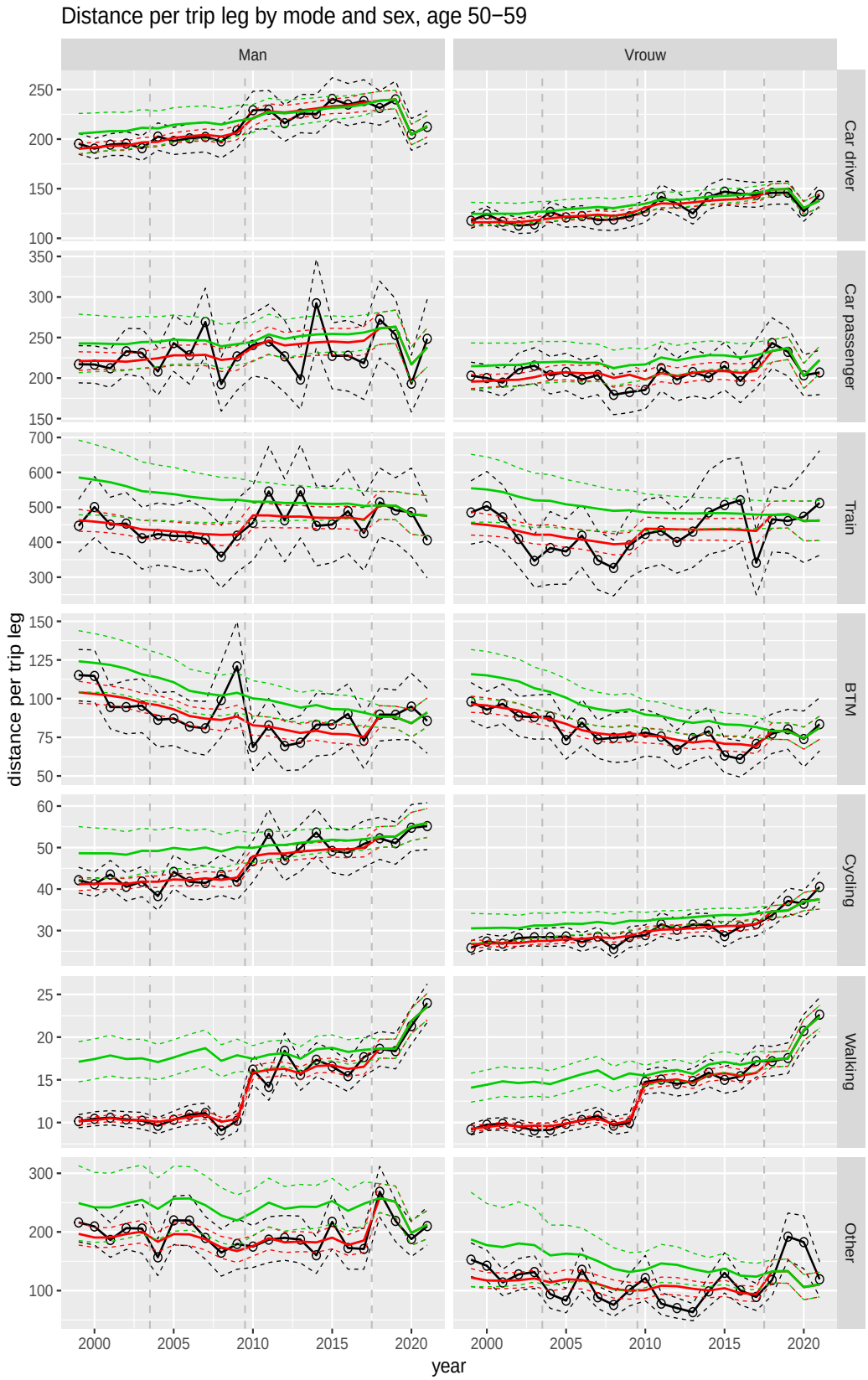

Figure A.116 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

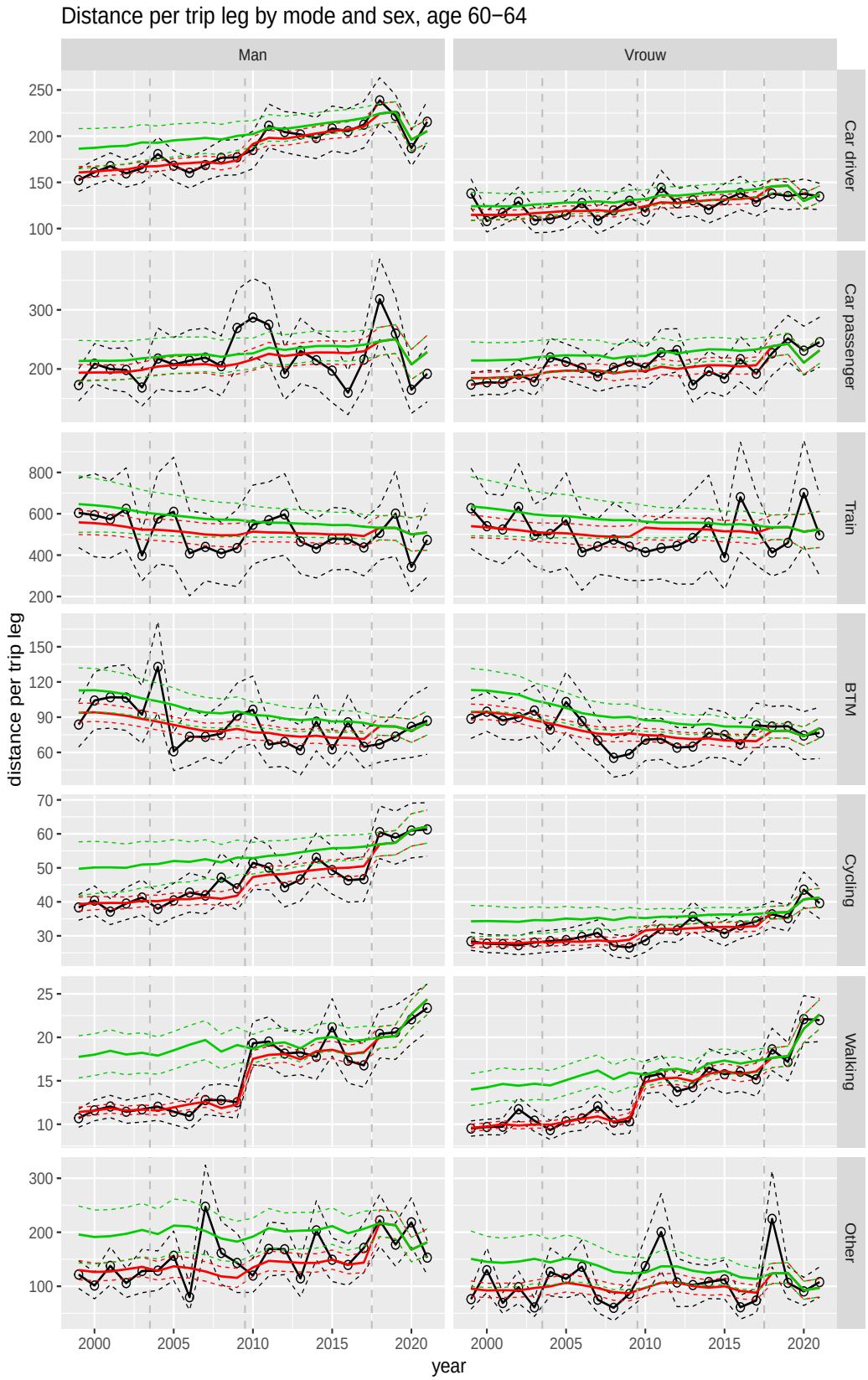

Figure A.117 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

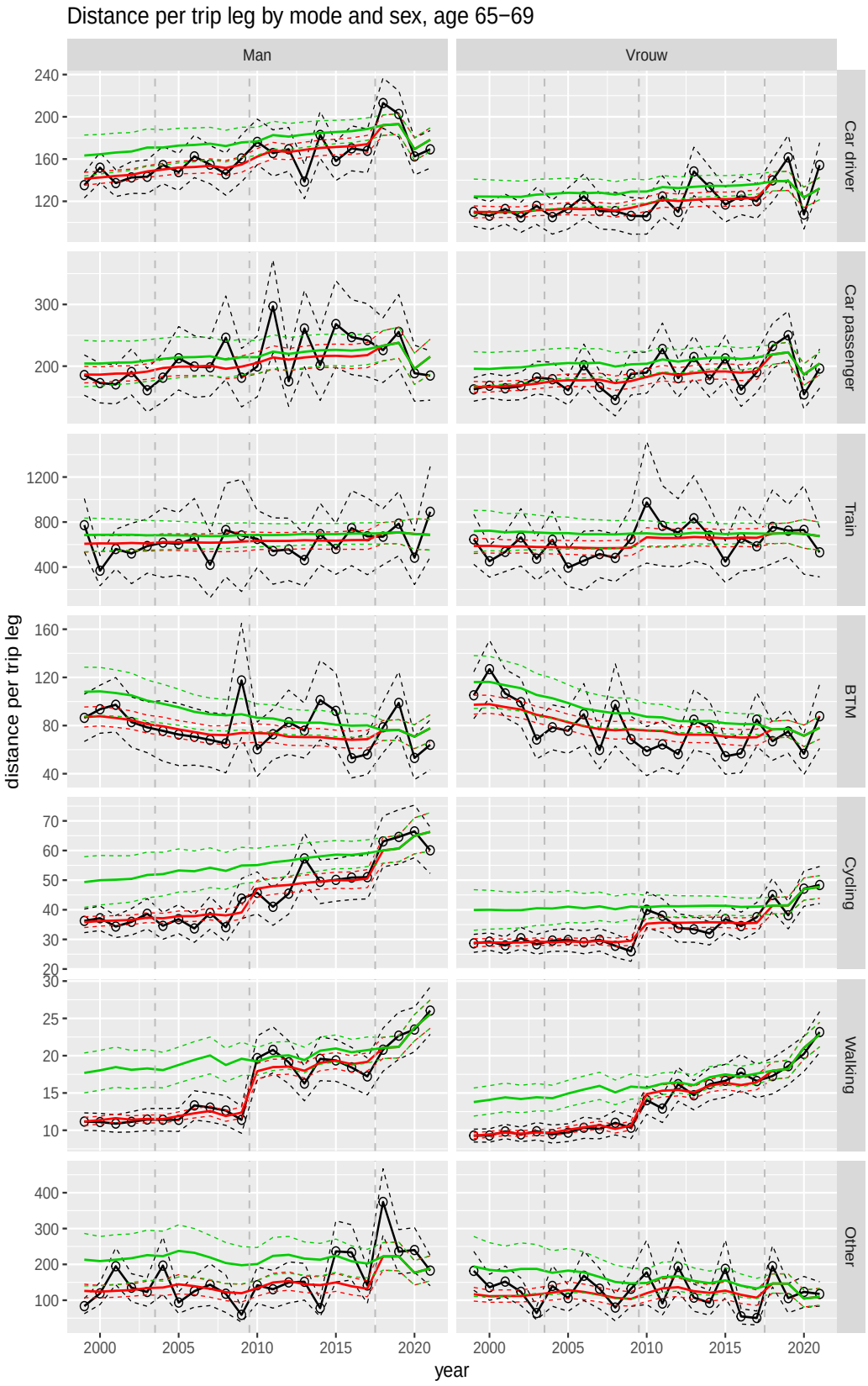

Figure A.118 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

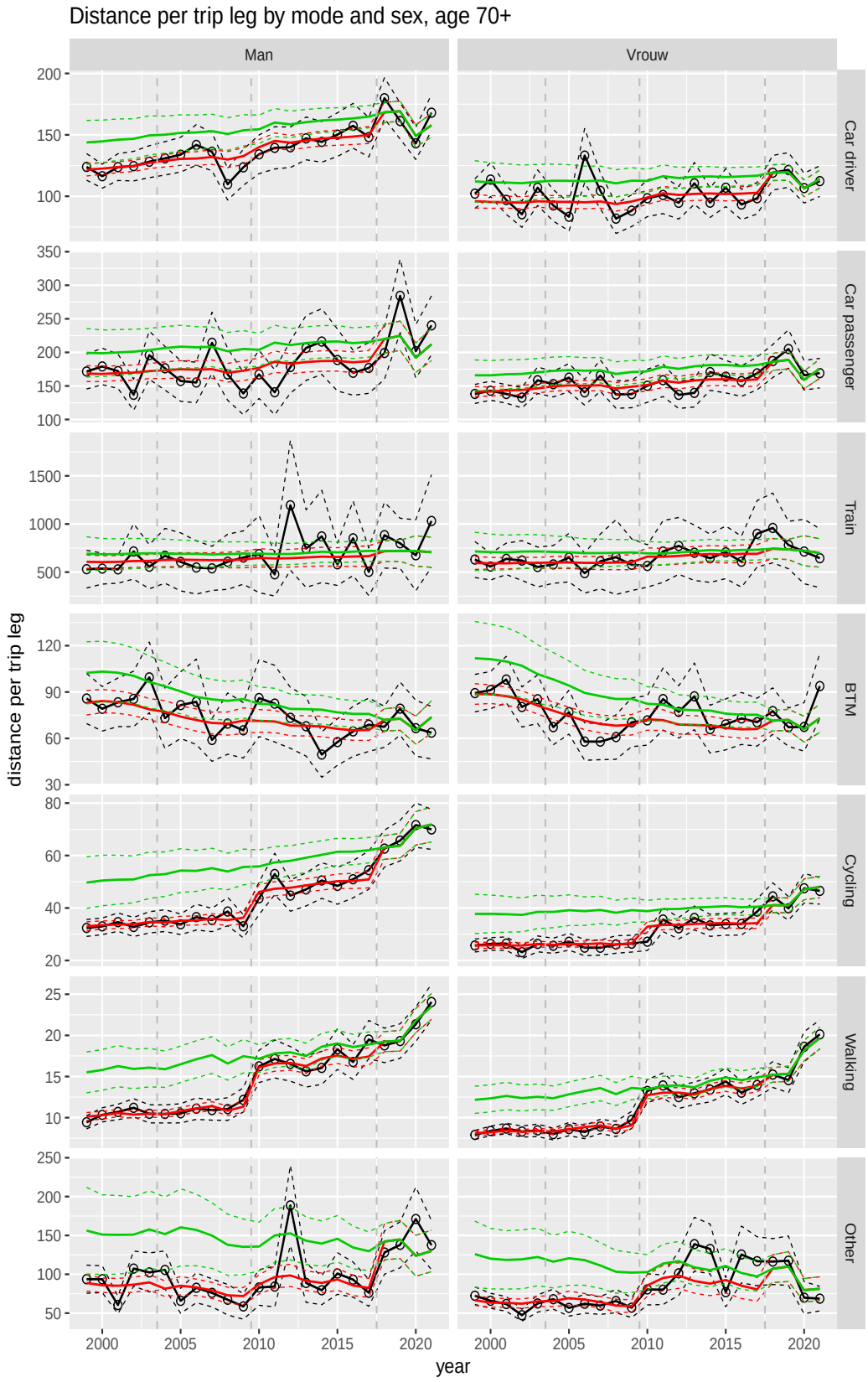

Figure A.119 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

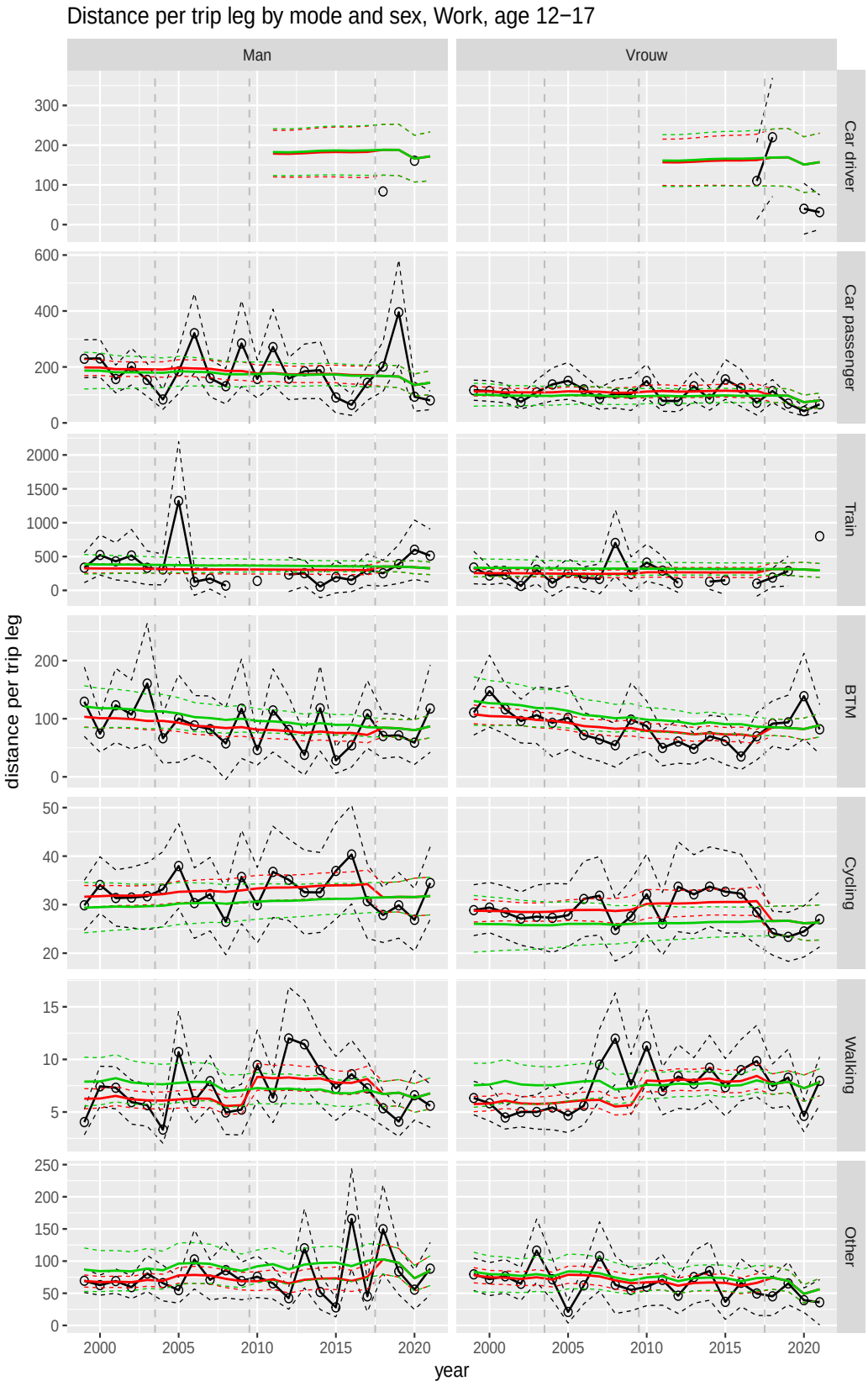

Figure A.120 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.121 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

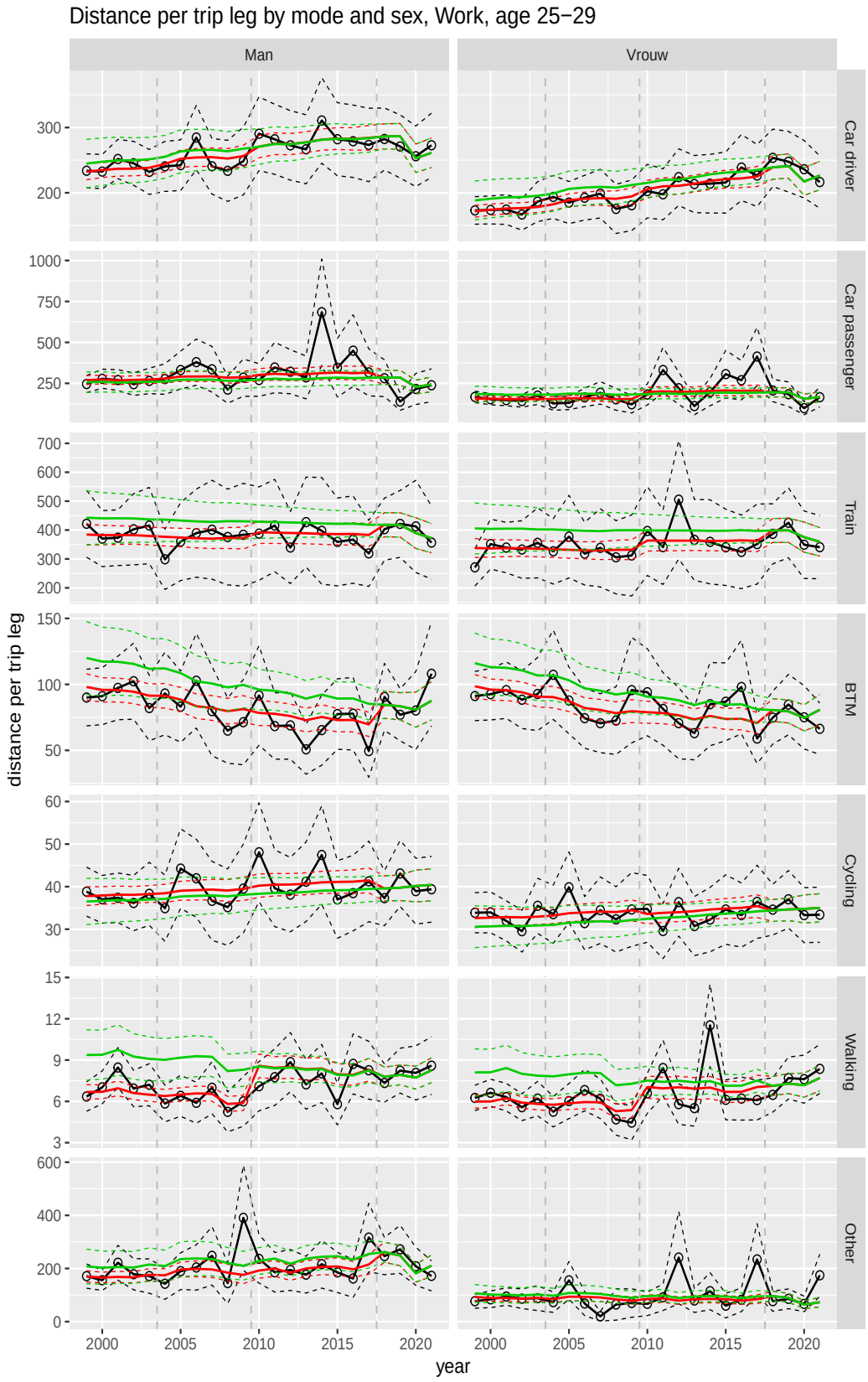

Figure A.122 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

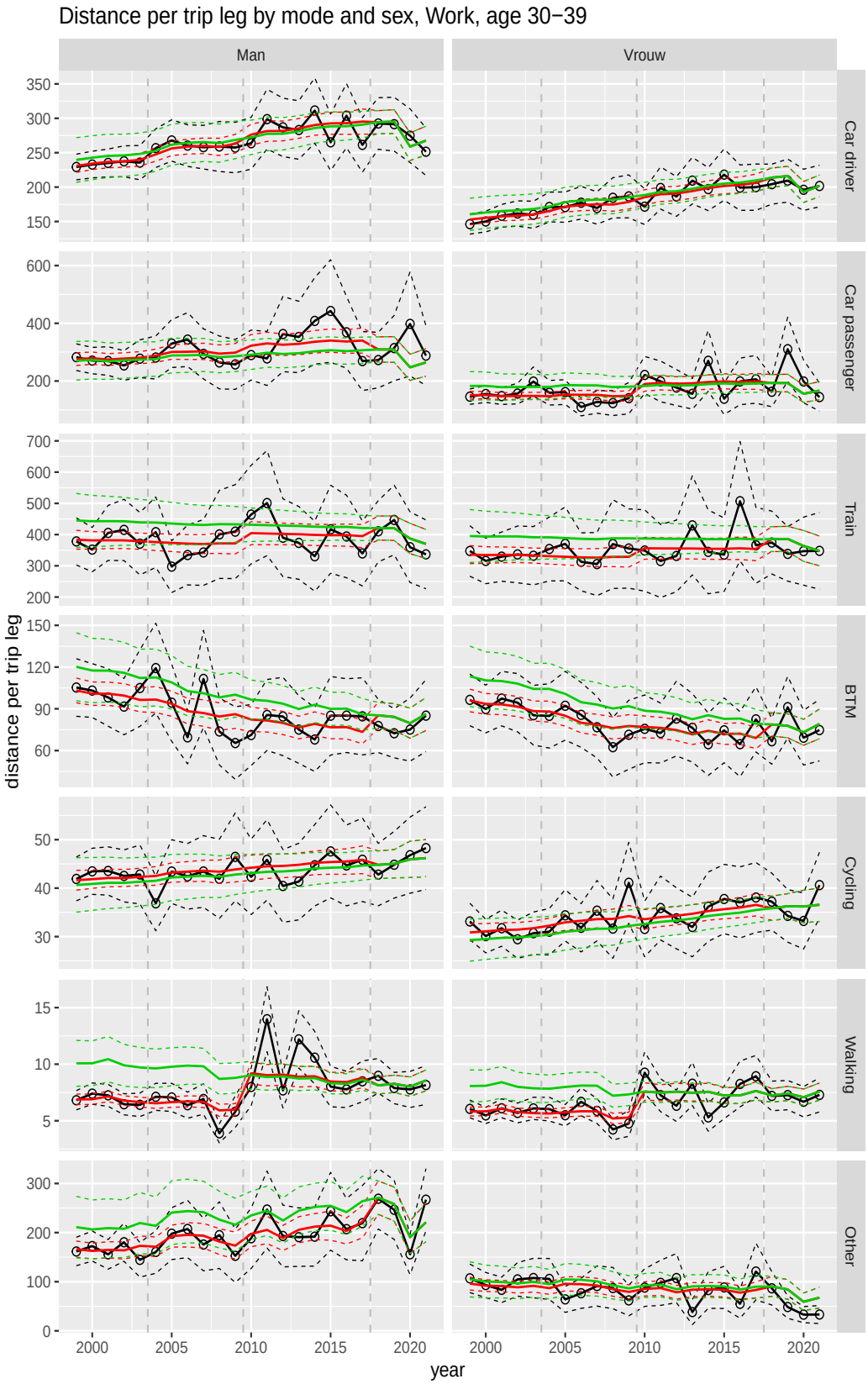

Figure A.123 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

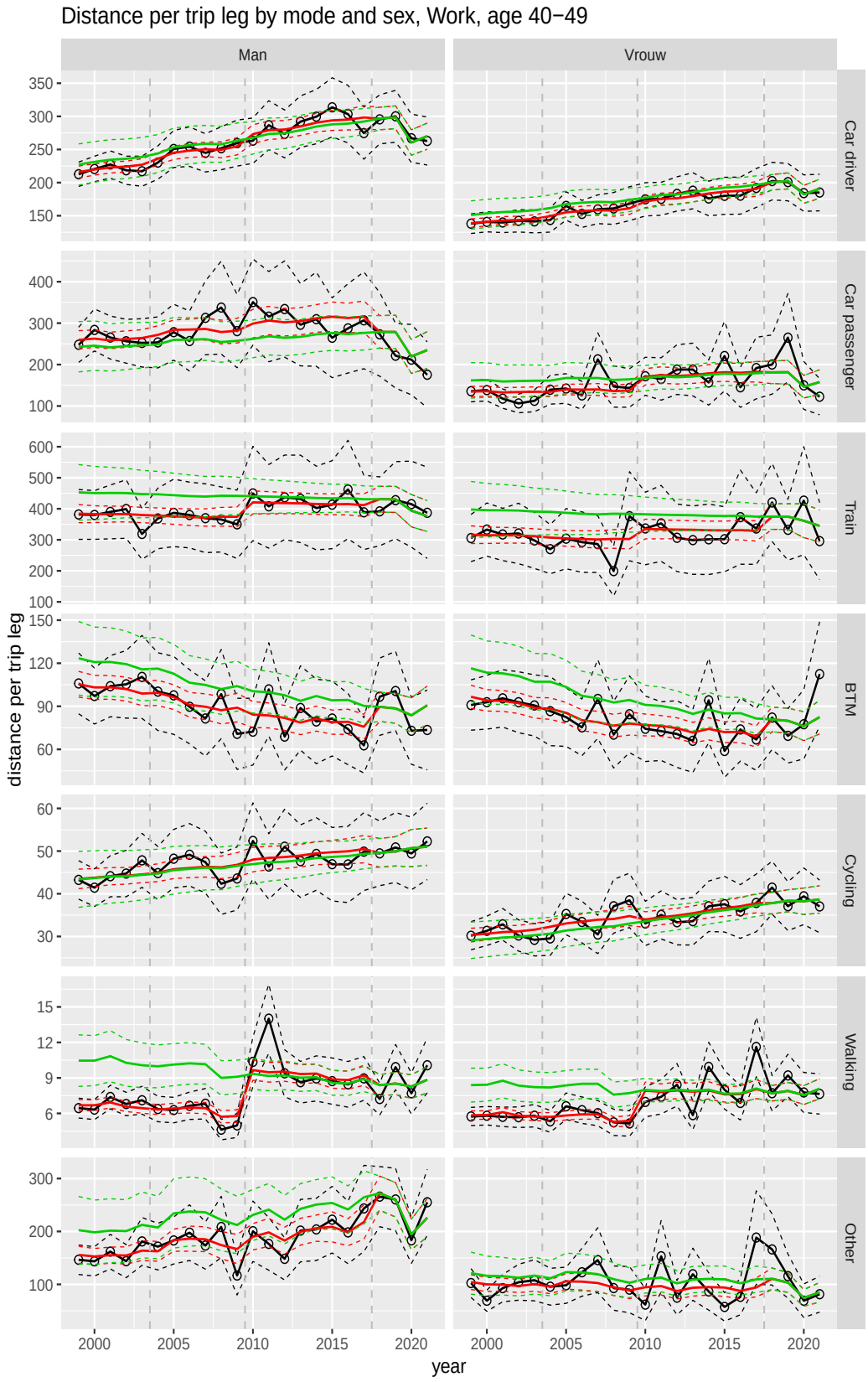

Figure A.124 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

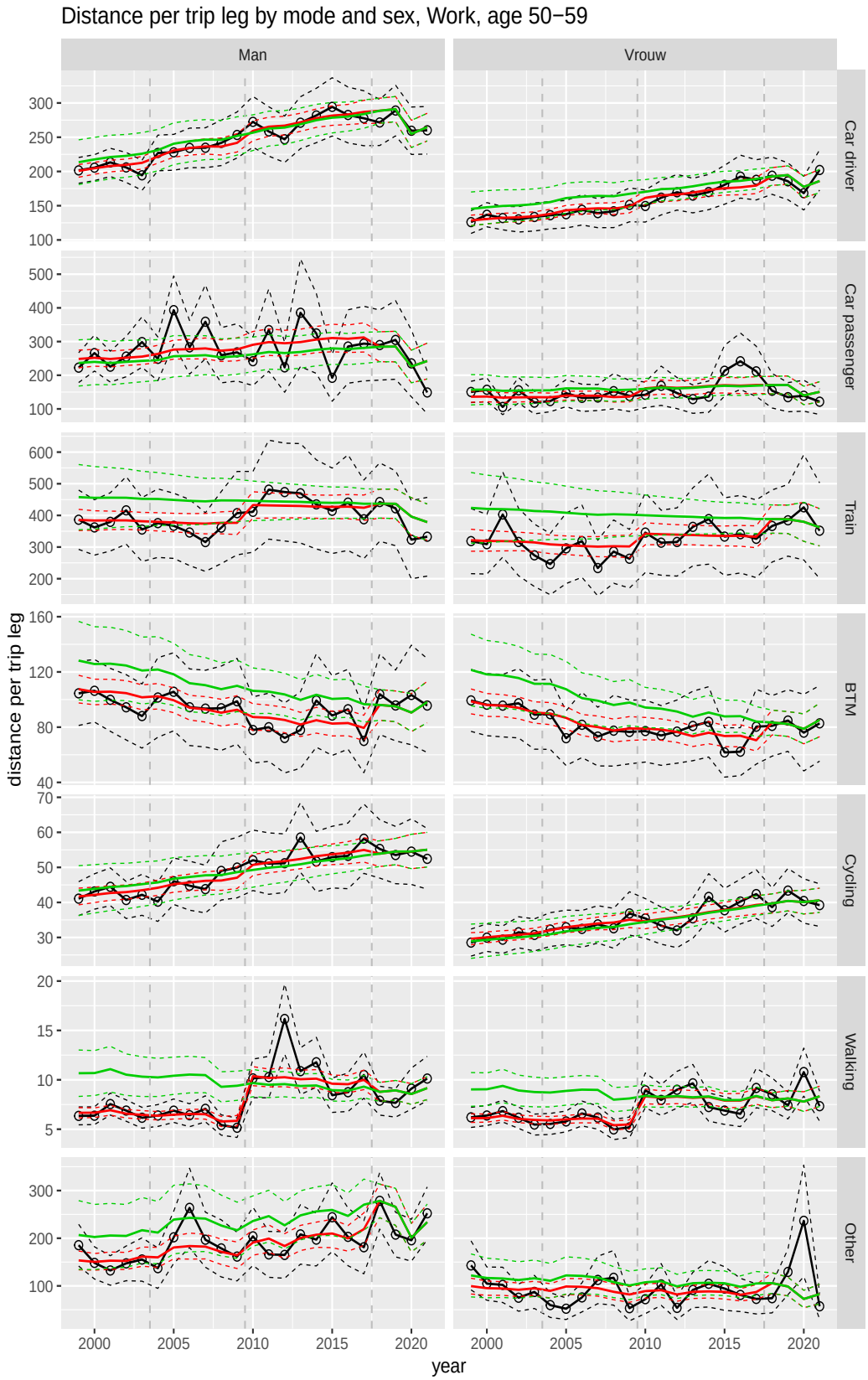

Figure A.125 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

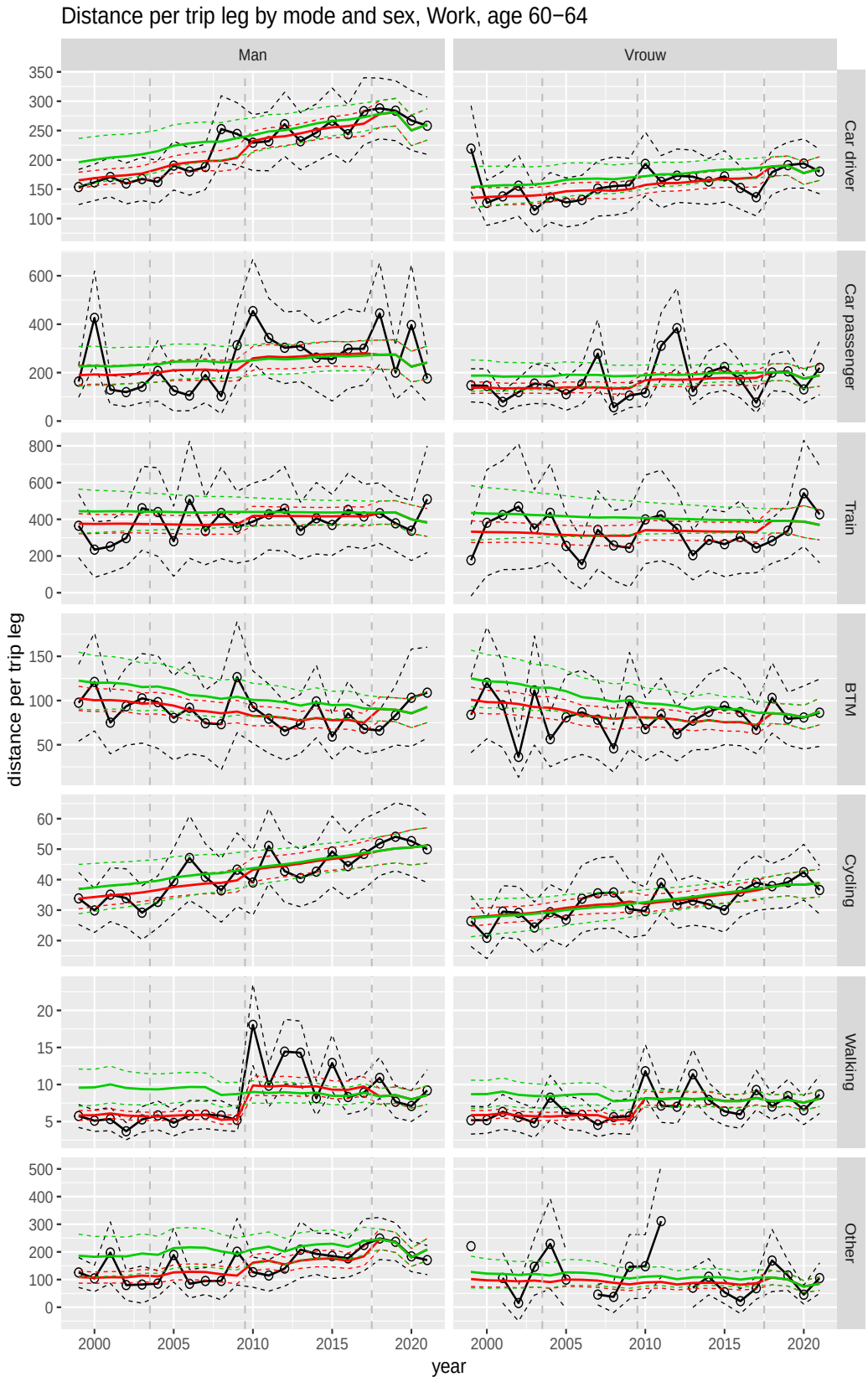

Figure A.126 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

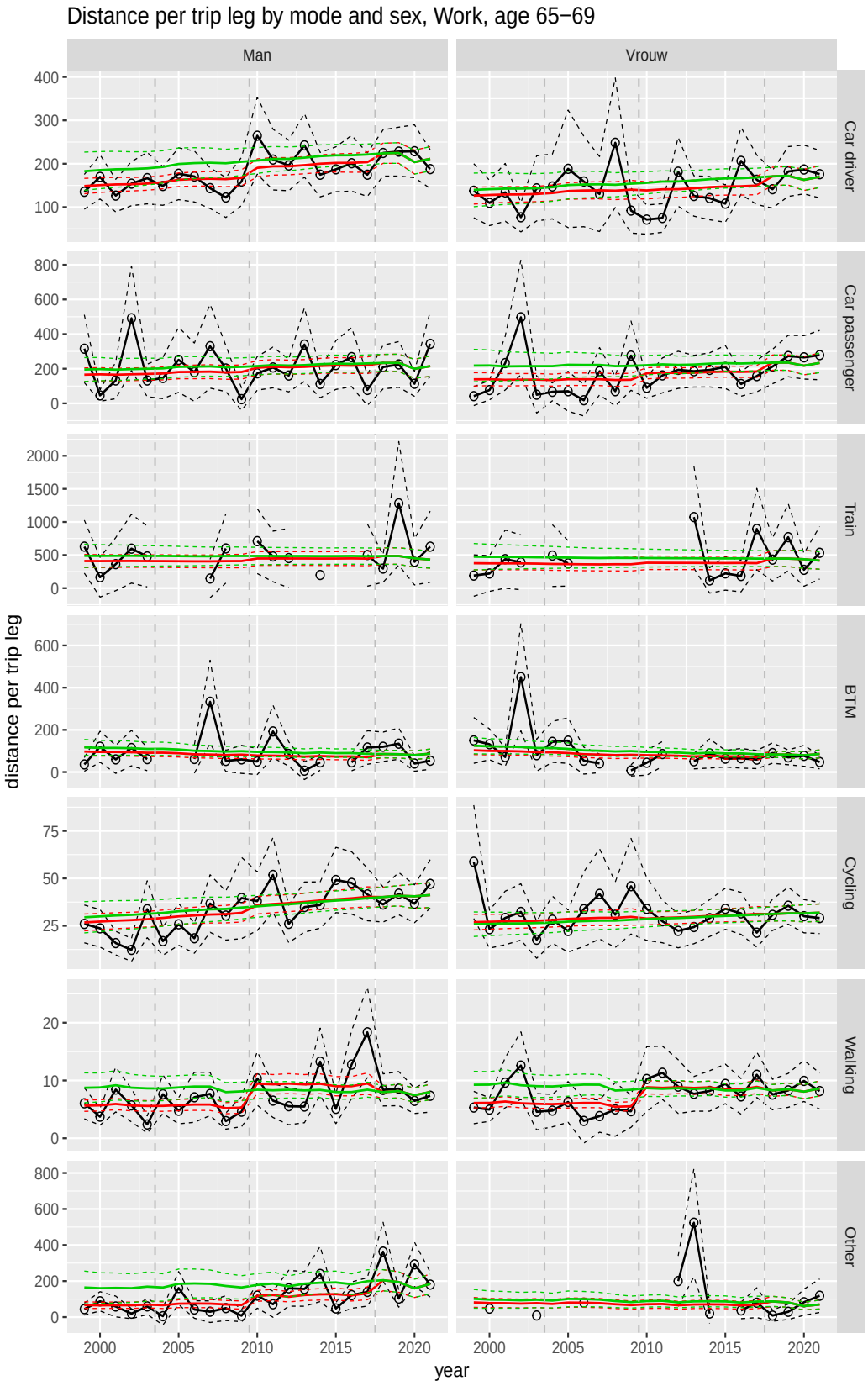

Figure A.127 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

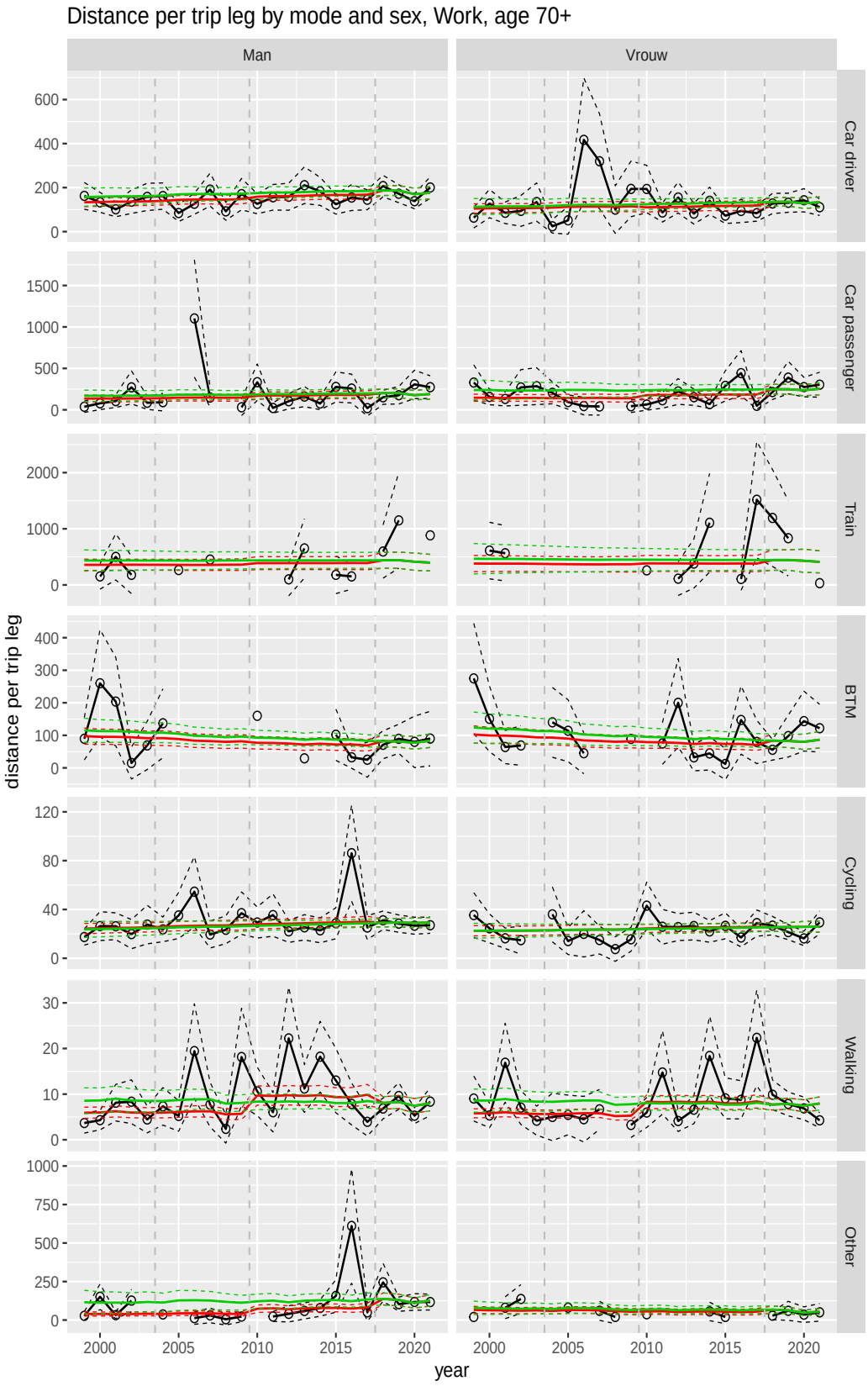

Figure A.128 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

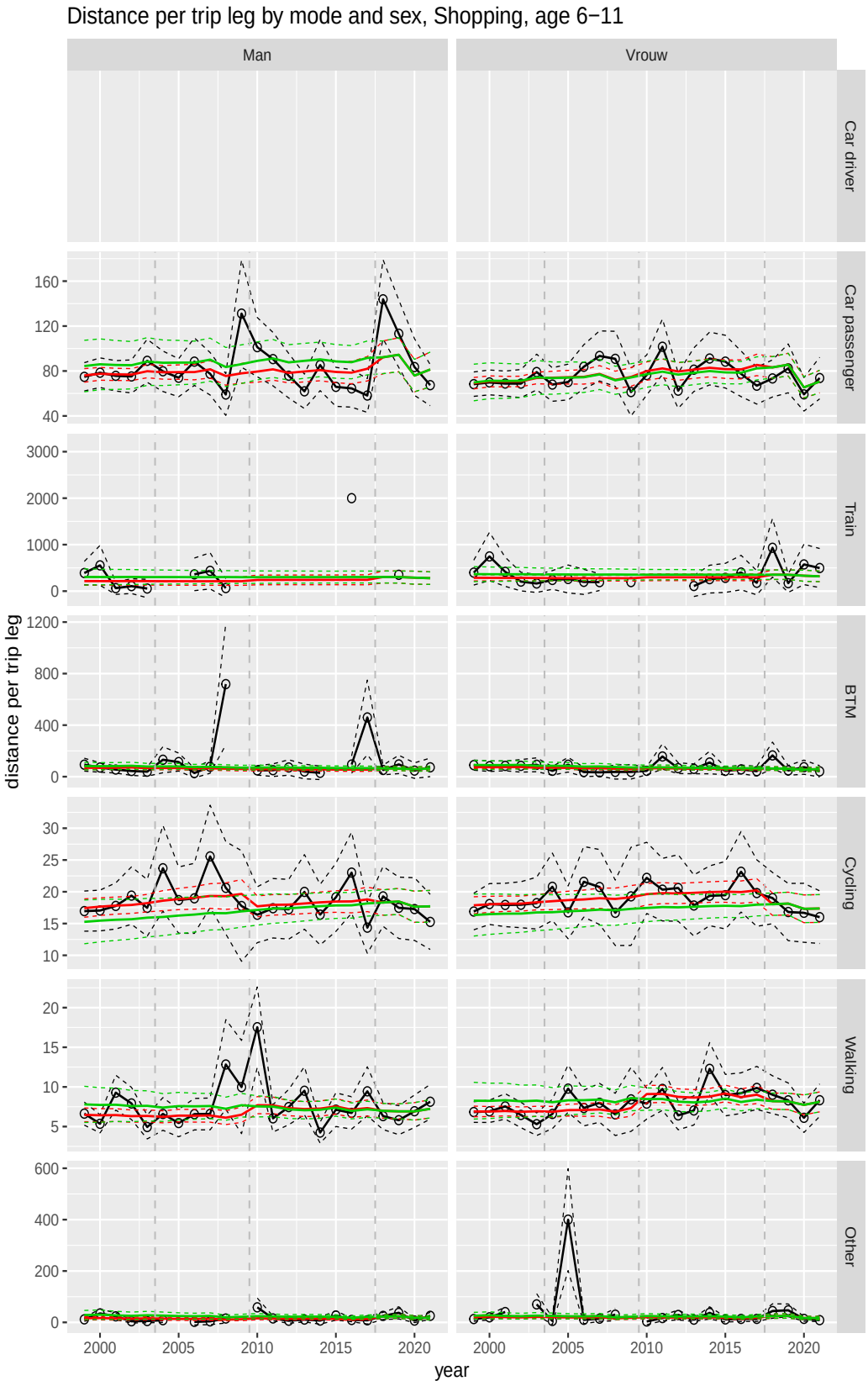

Figure A.129 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

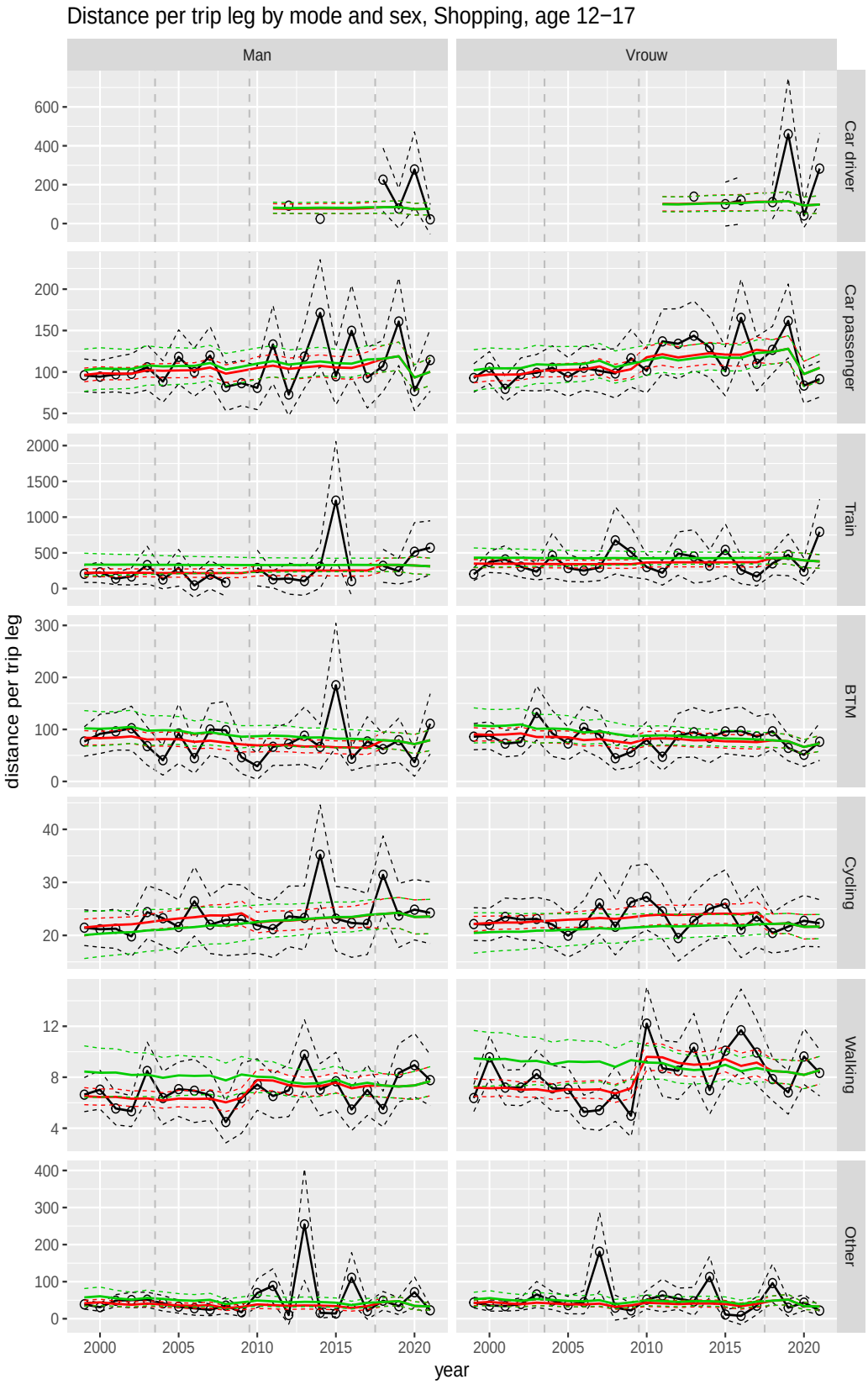

Figure A.130 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

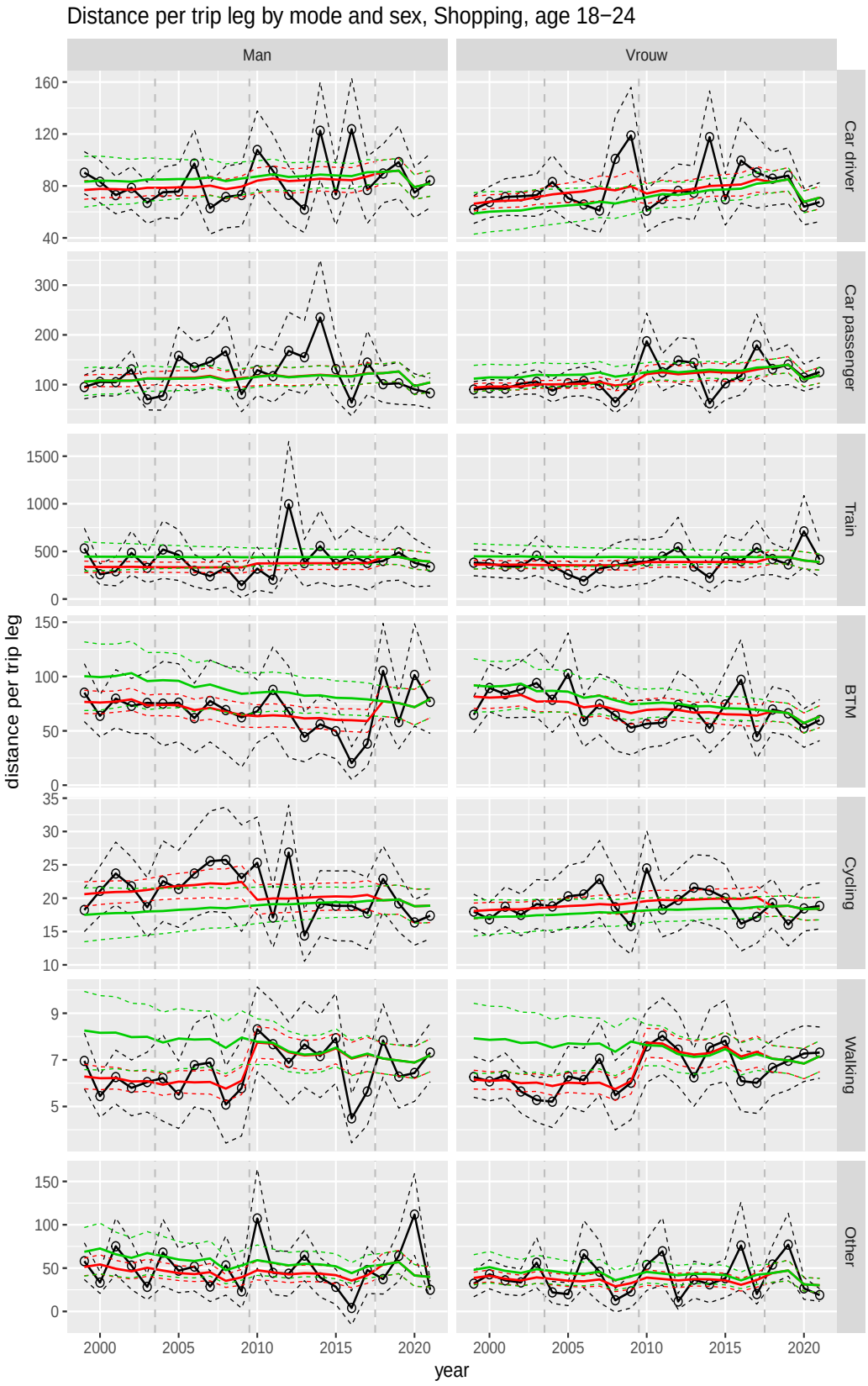

Figure A.131 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

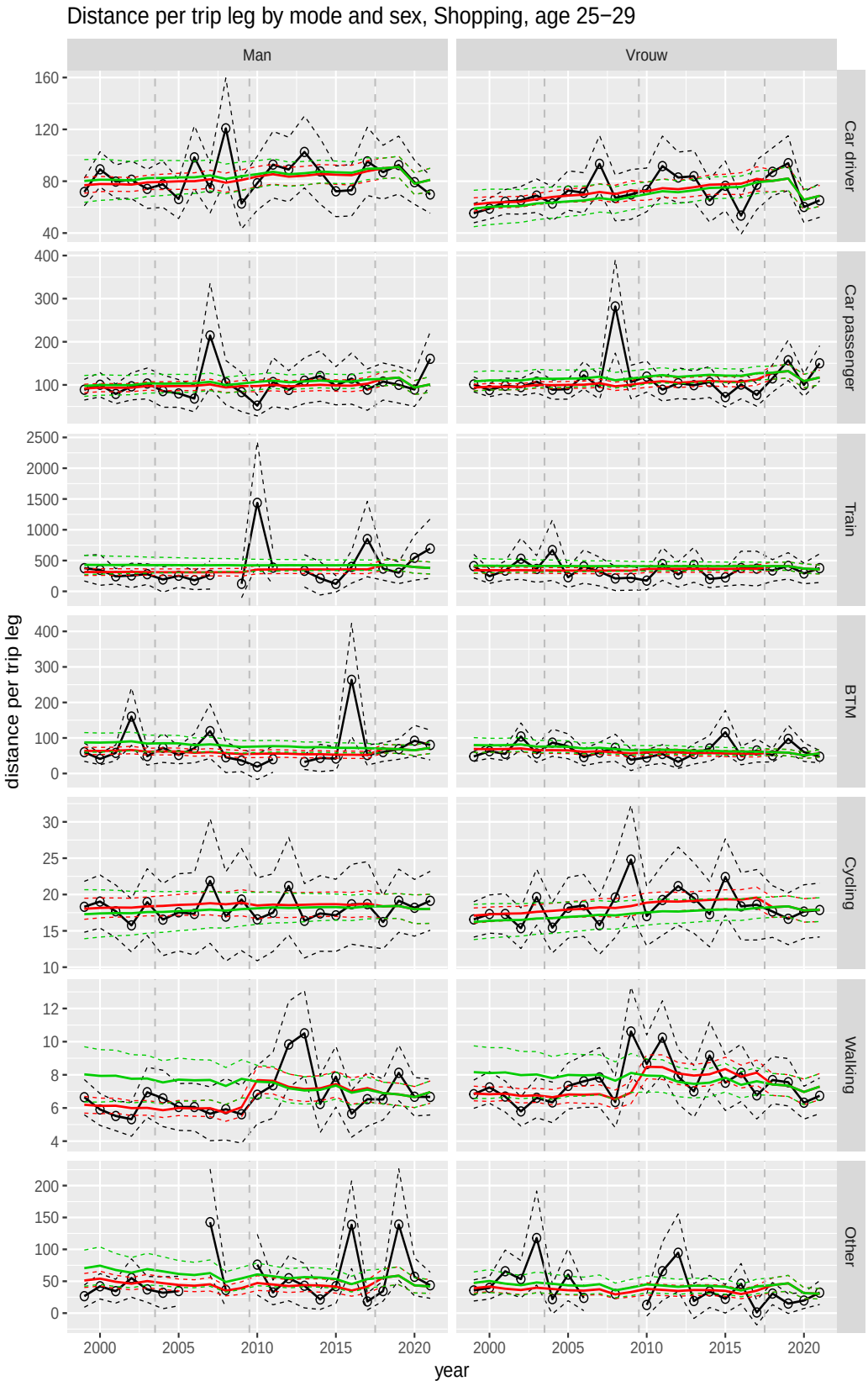

Figure A.132 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

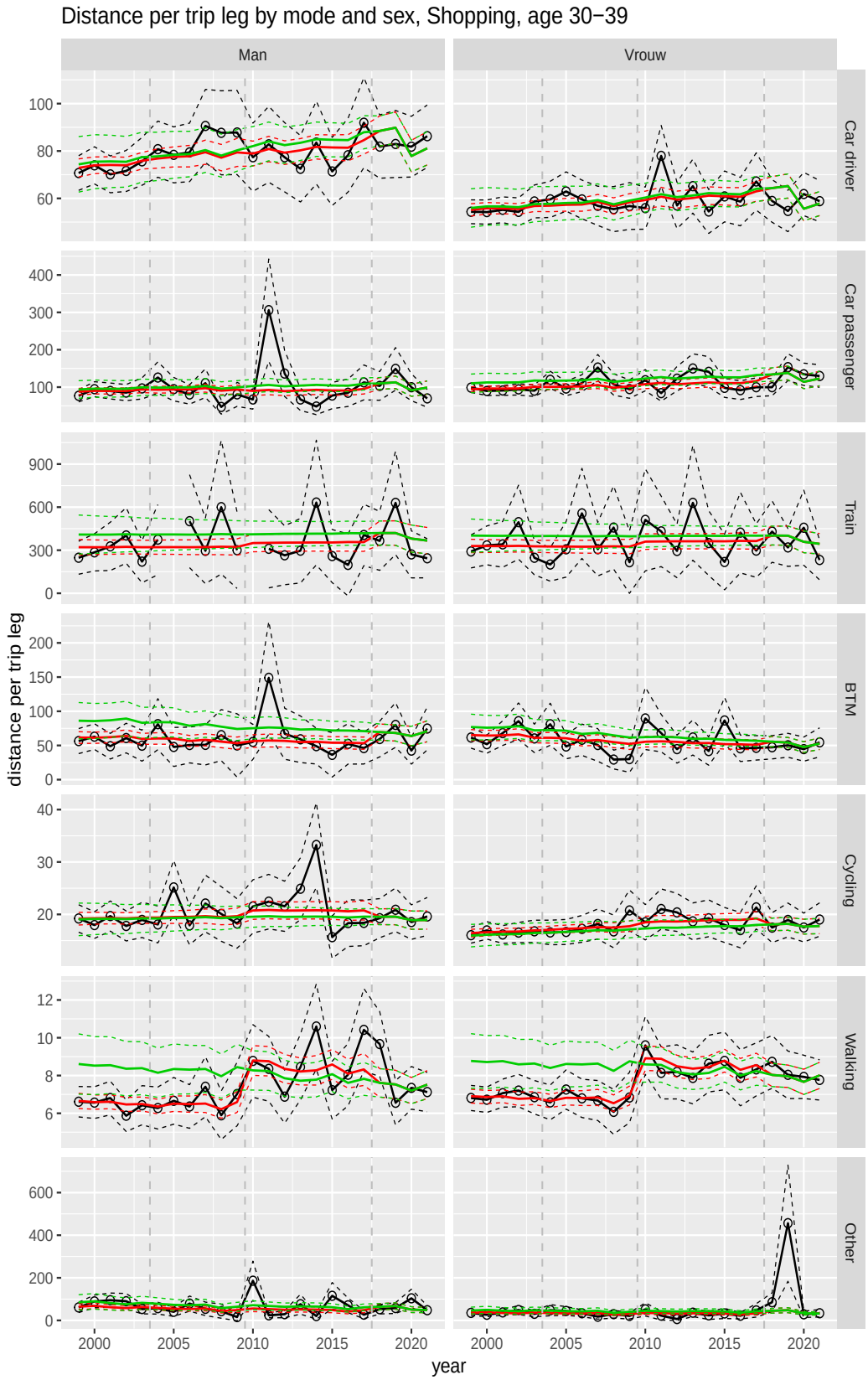

Figure A.133 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

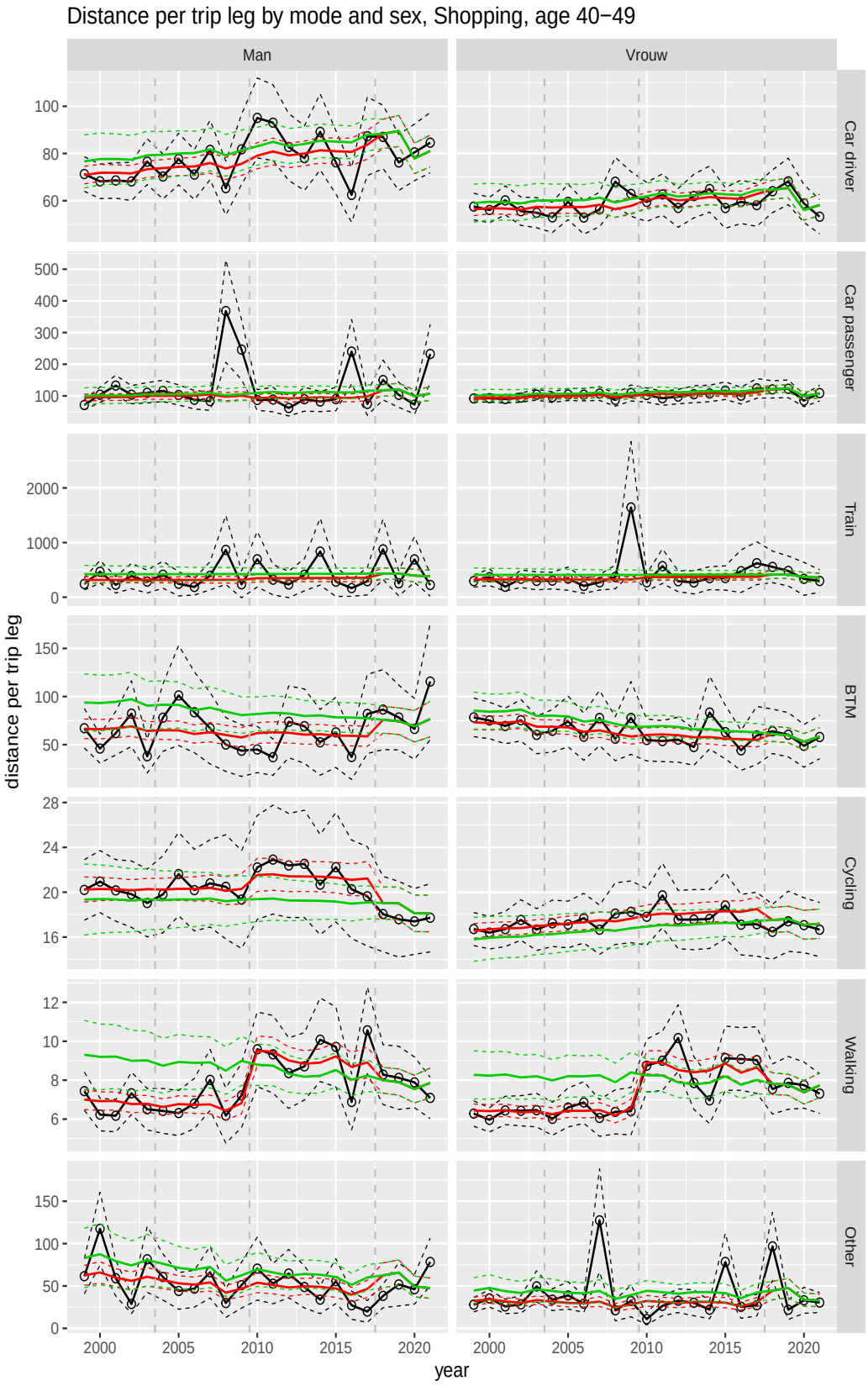

Figure A.134 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

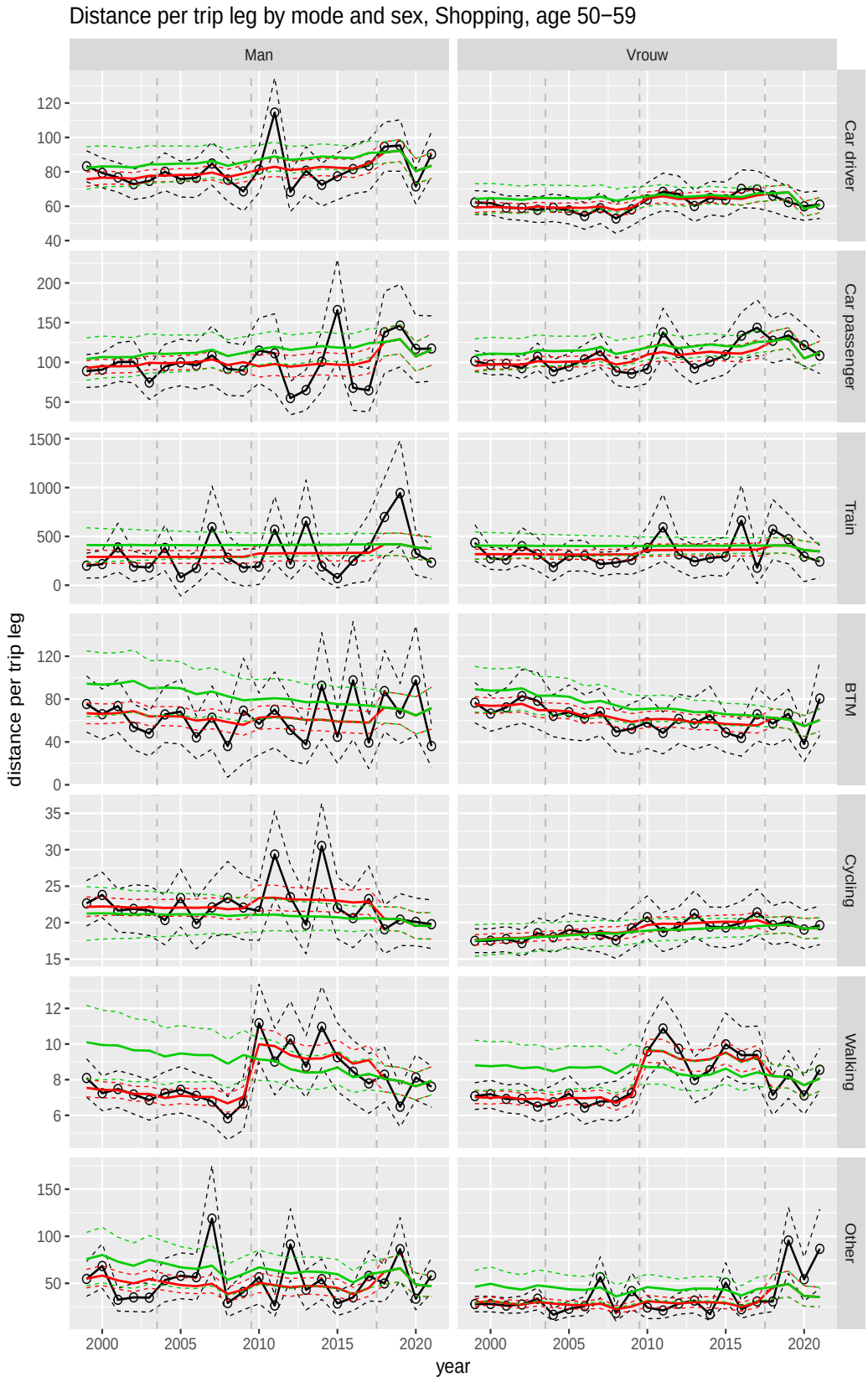

Figure A.135 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

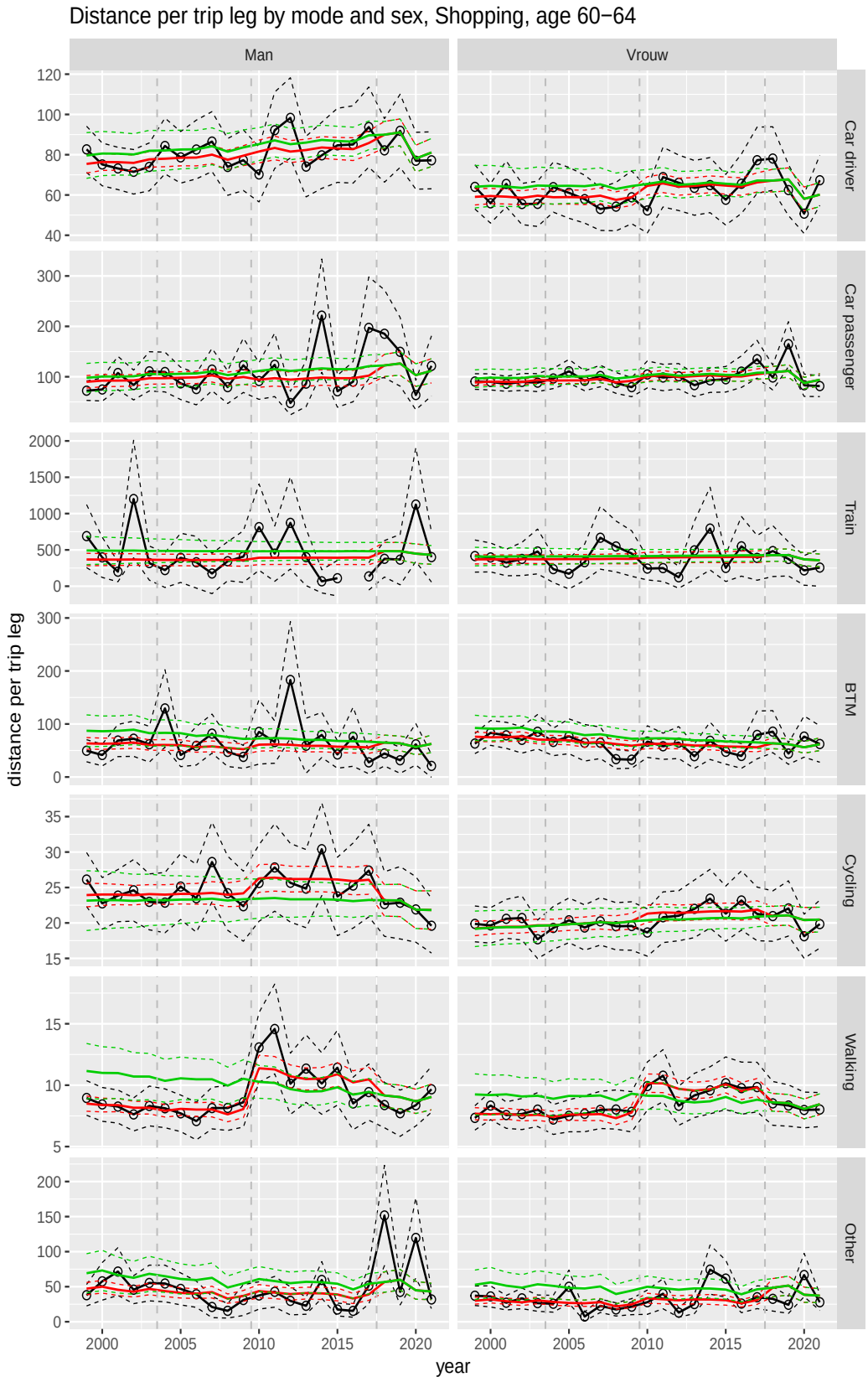

Figure A.136 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

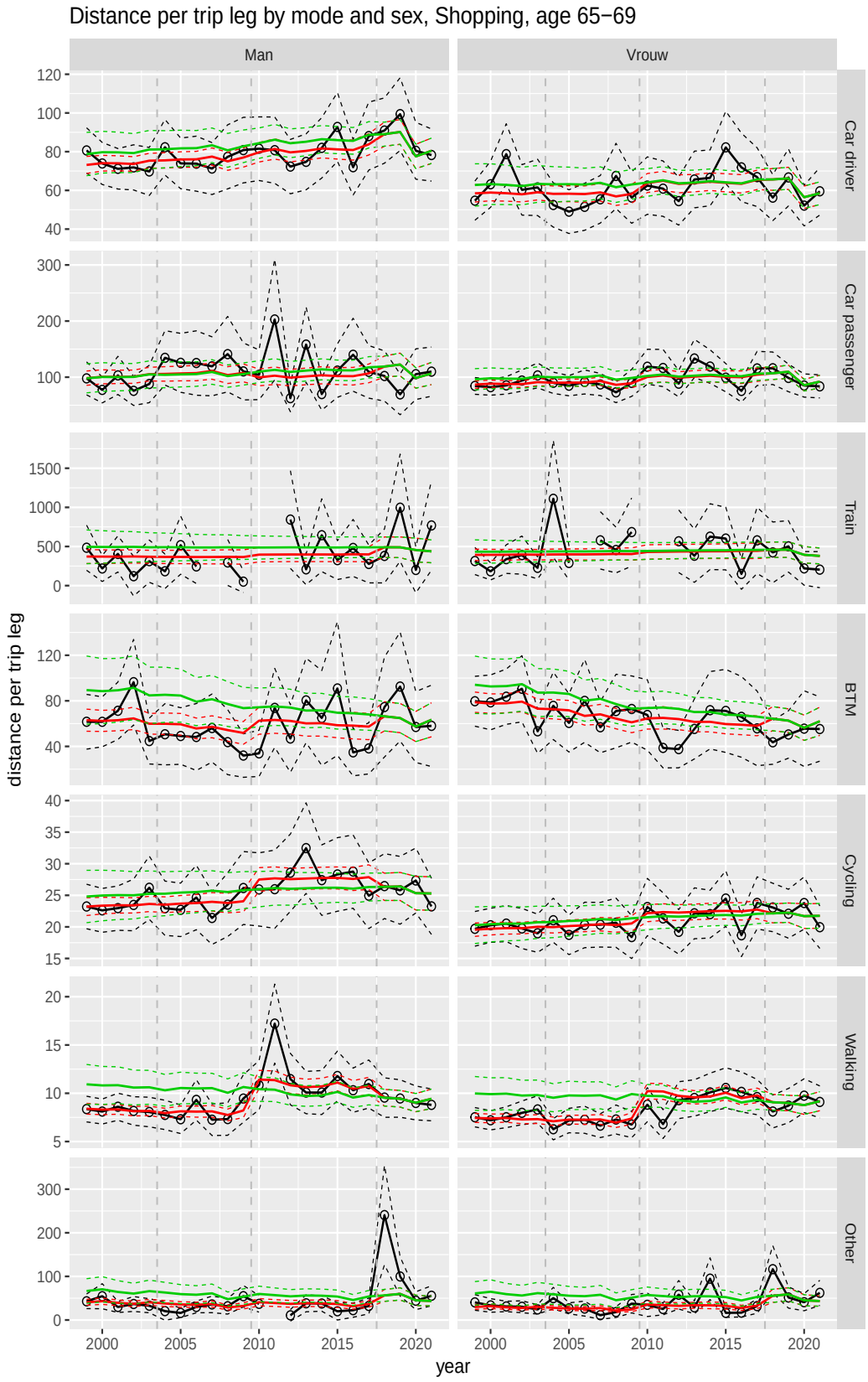

Figure A.137 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

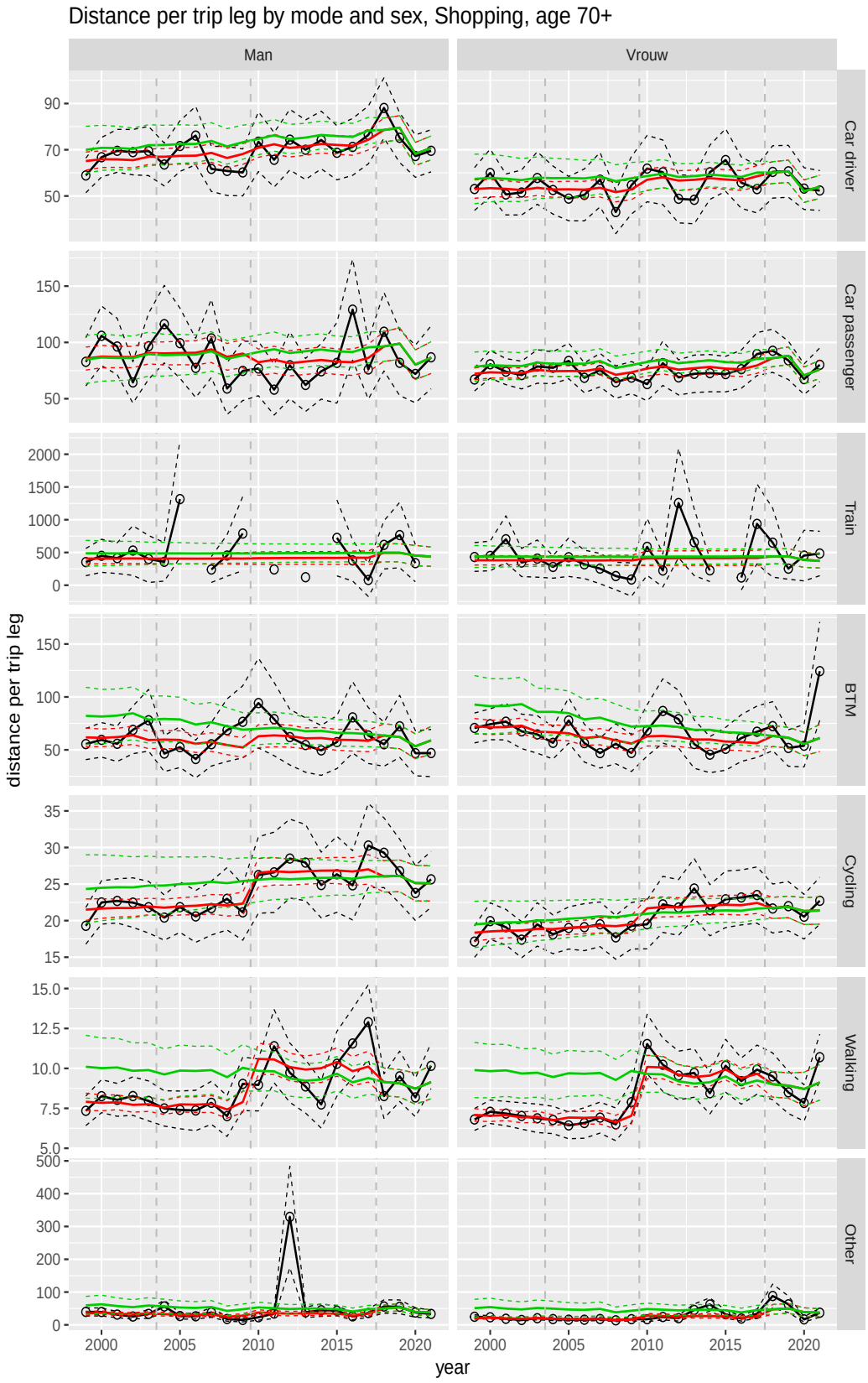

Figure A.138 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

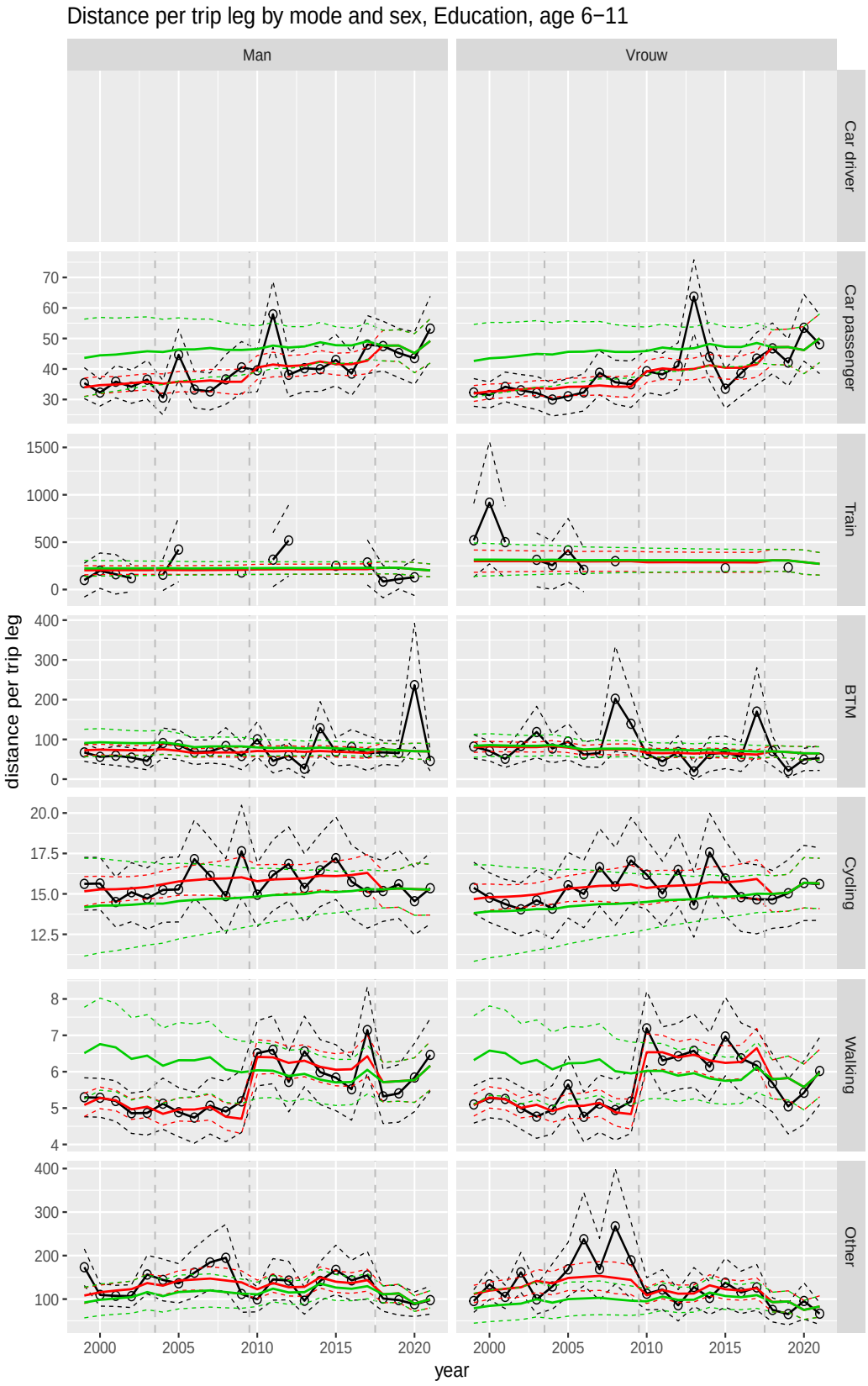

Figure A.139 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

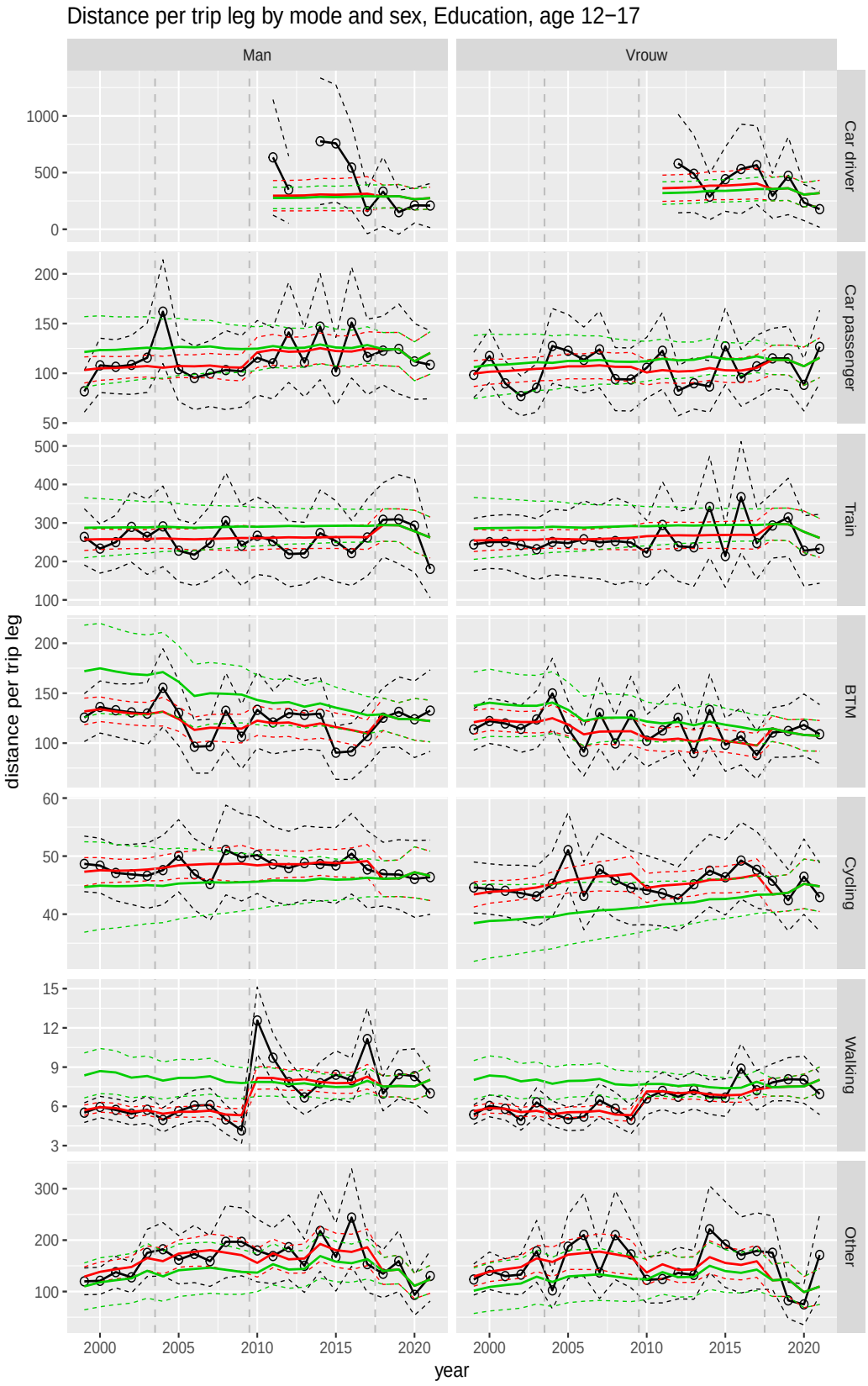

Figure A.140 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.141 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

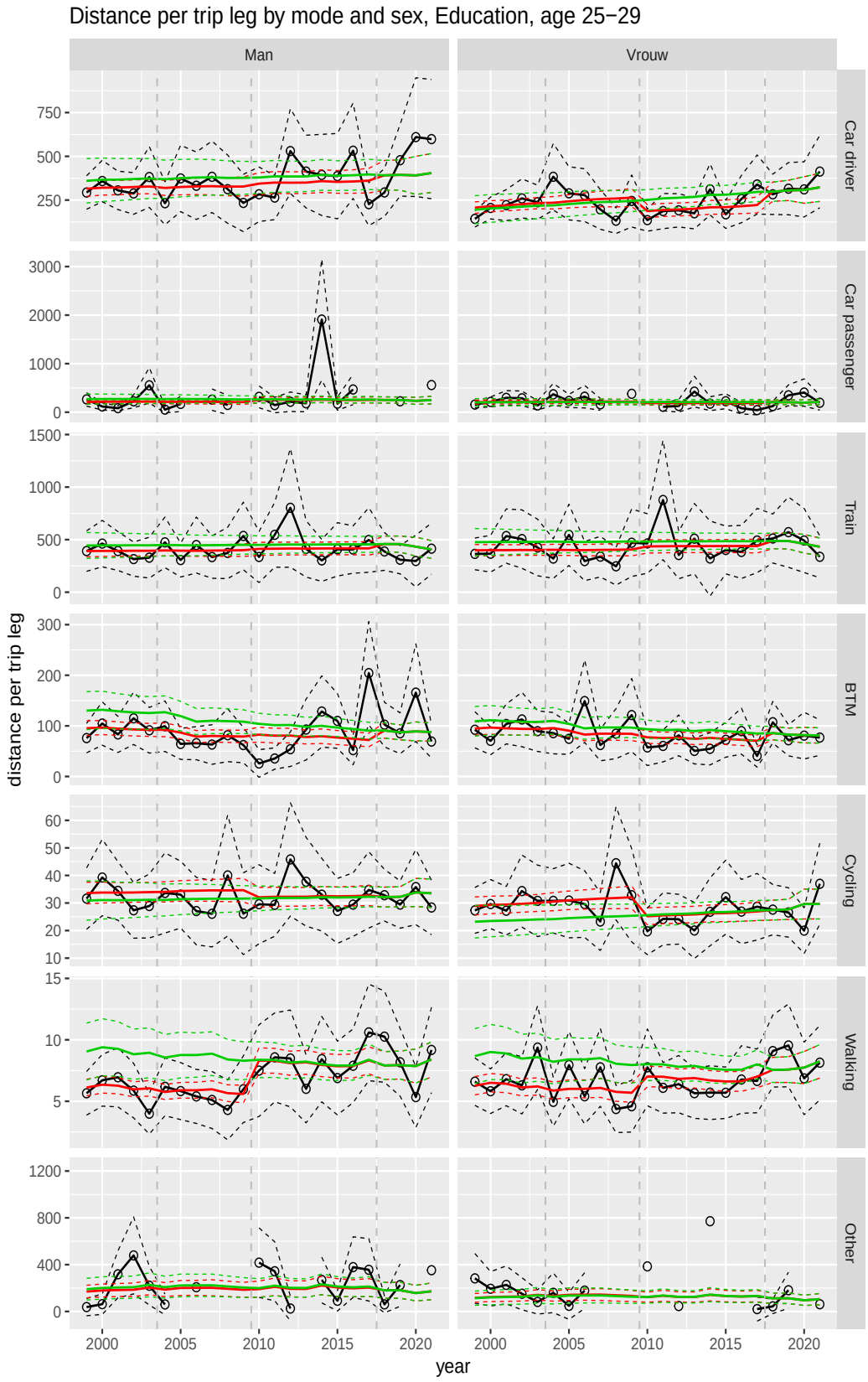

Figure A.142 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

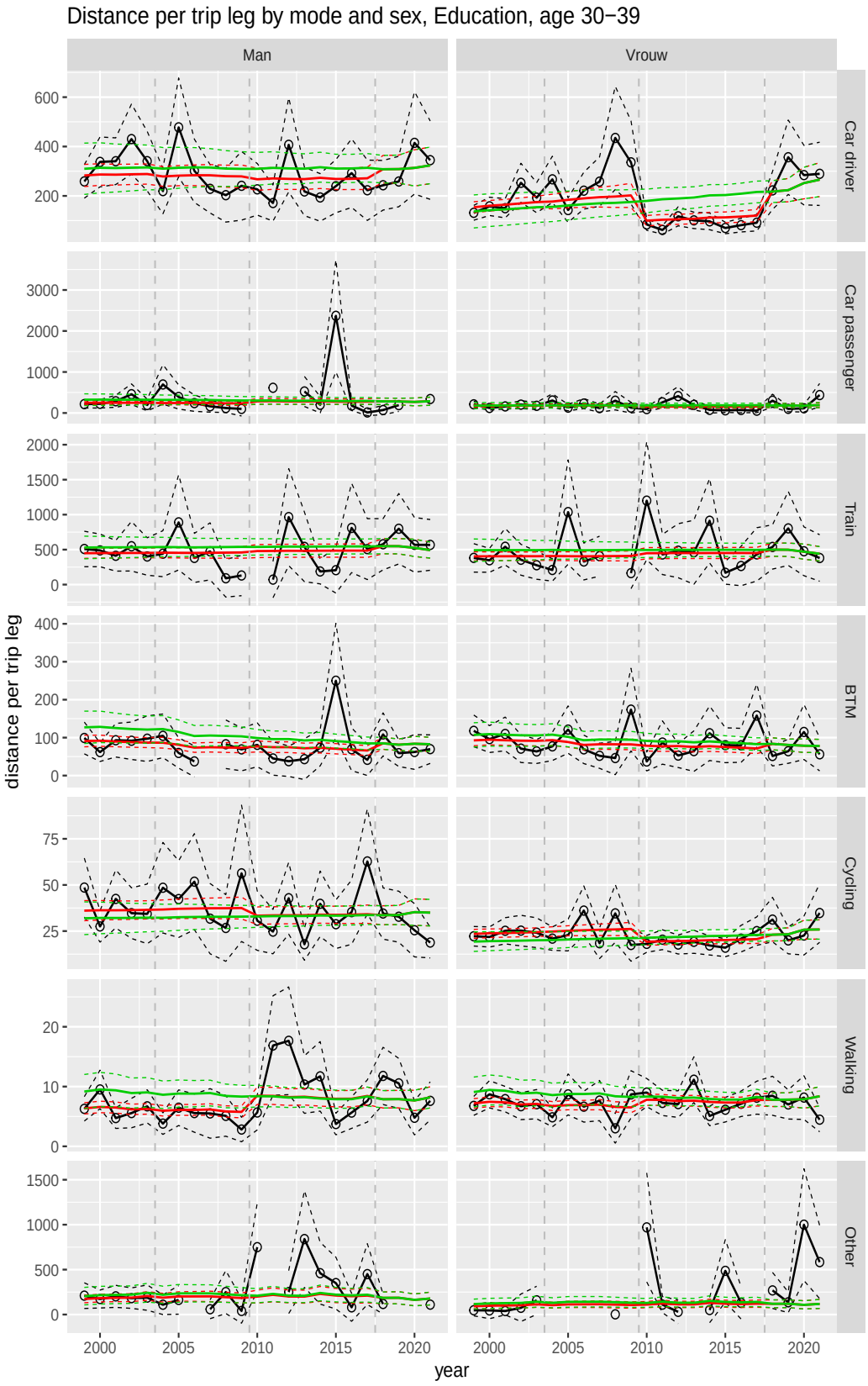

Figure A.143 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

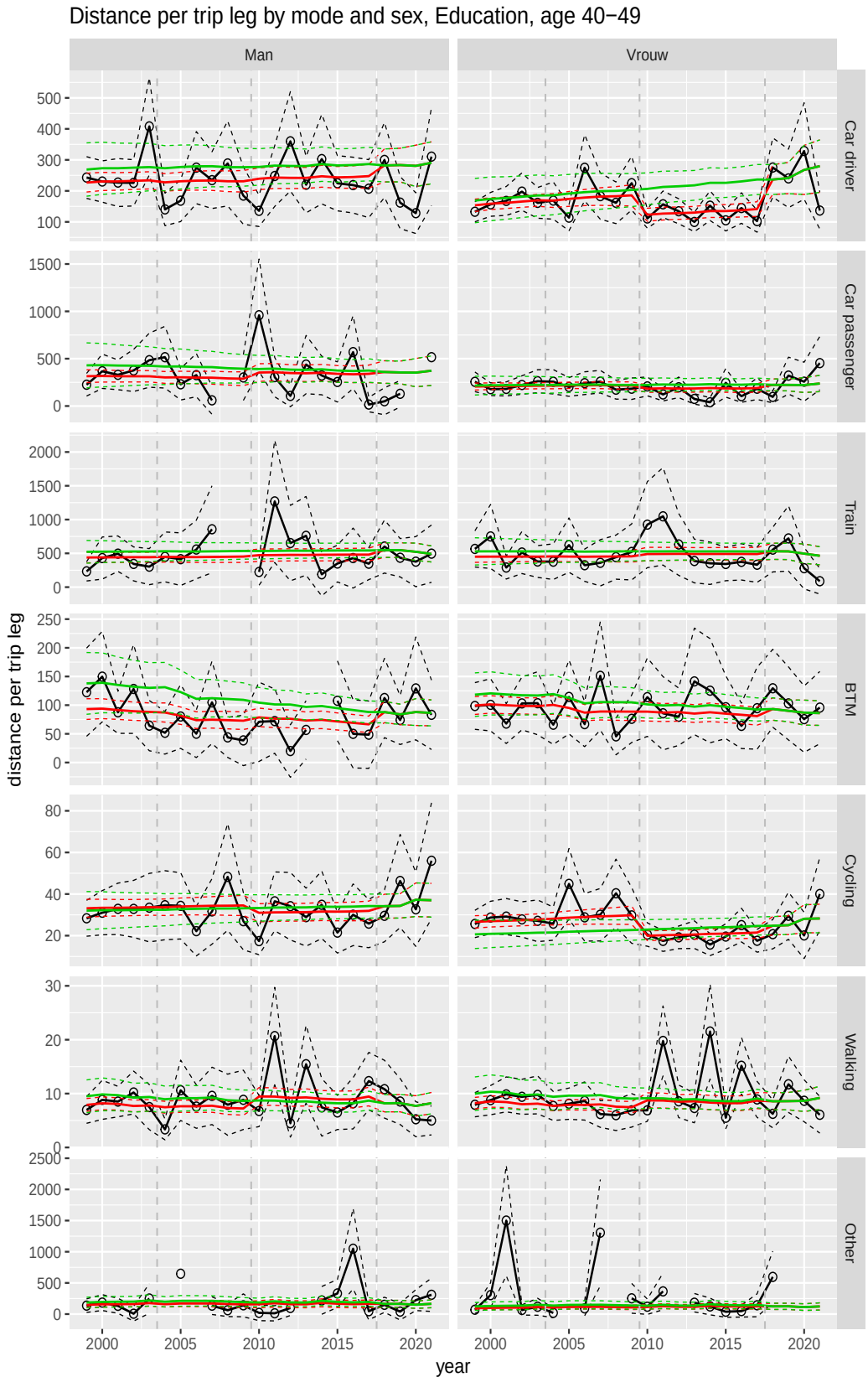

Figure A.144 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

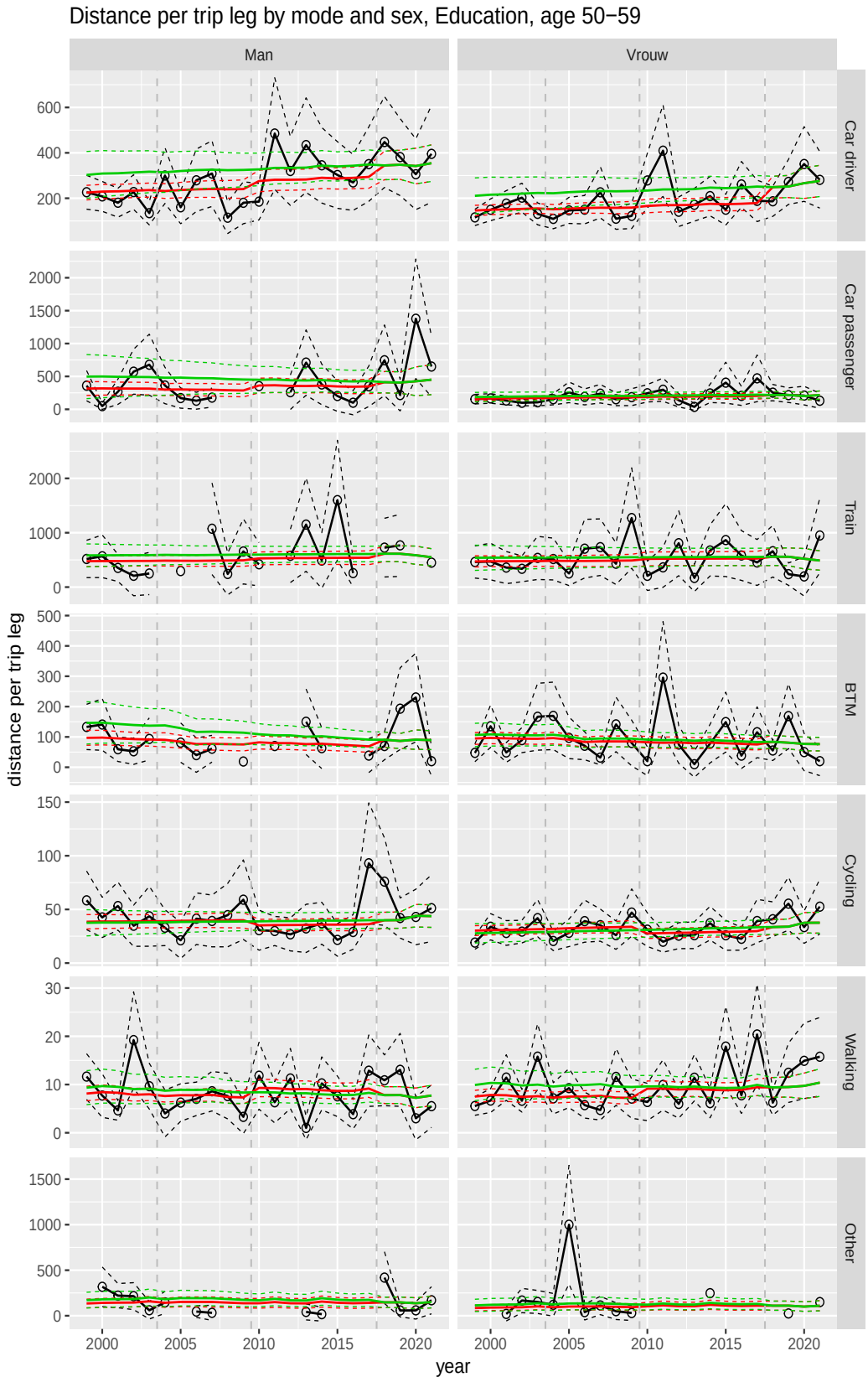

Figure A.145 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

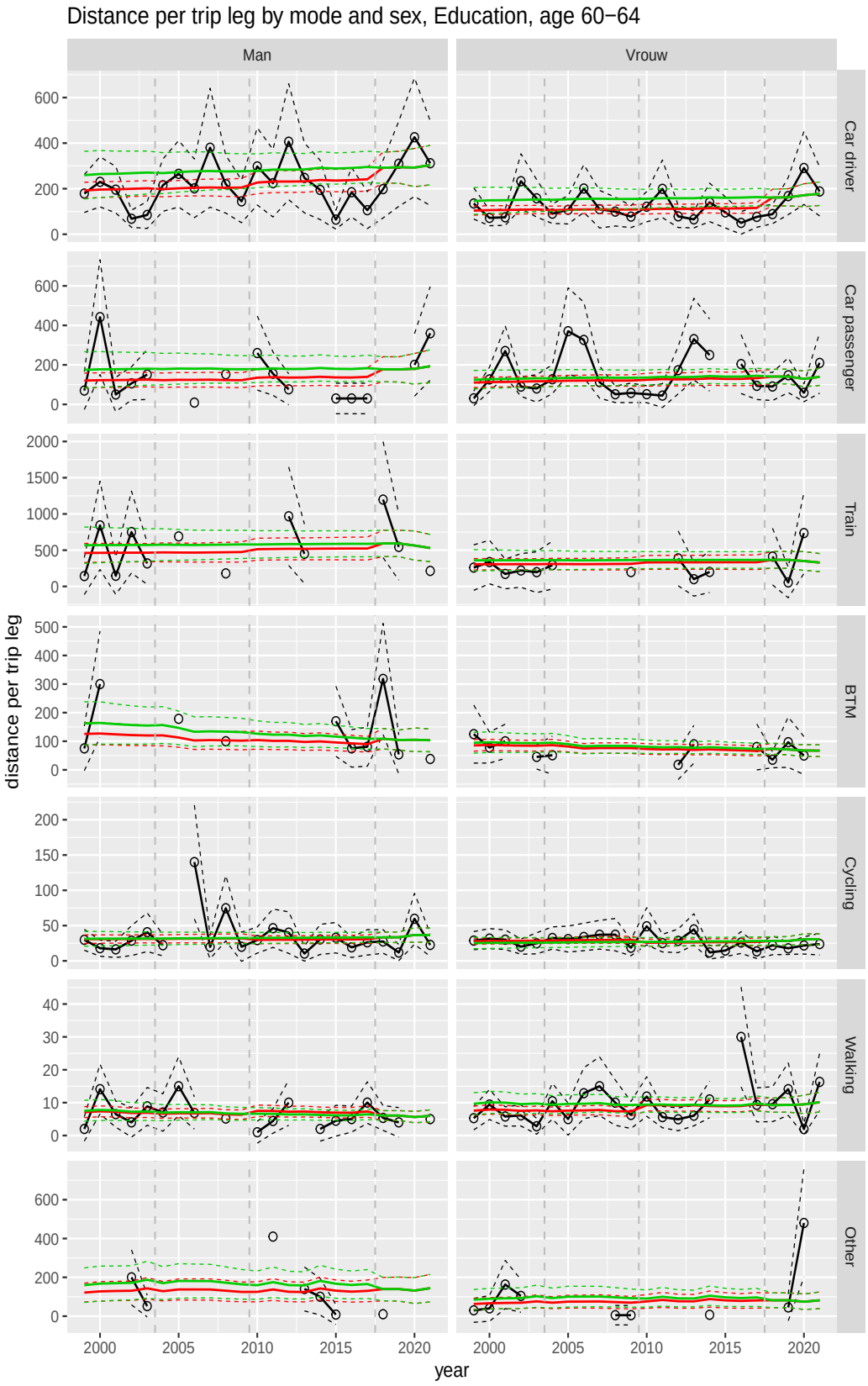

Figure A.146 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

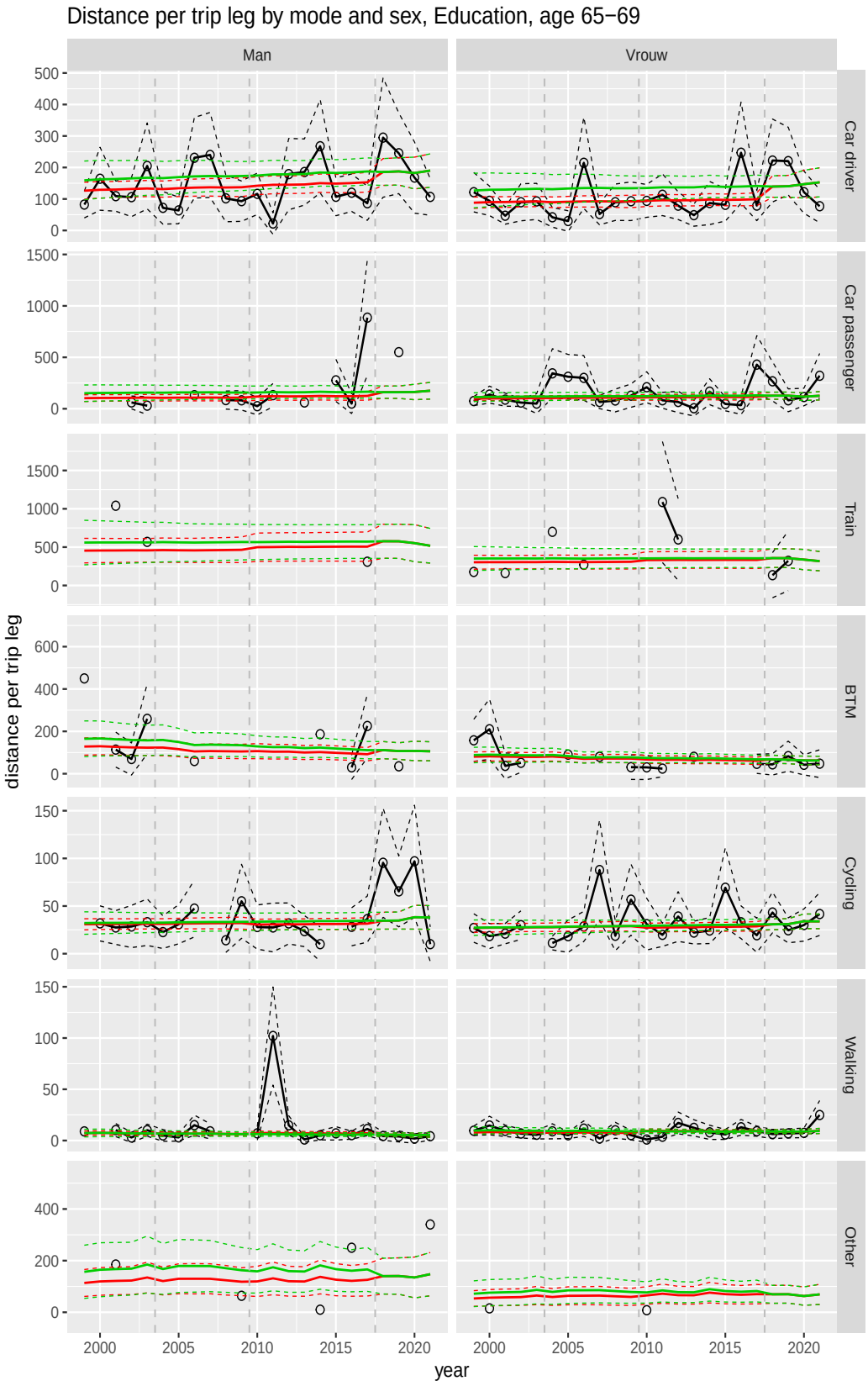

Figure A.147 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

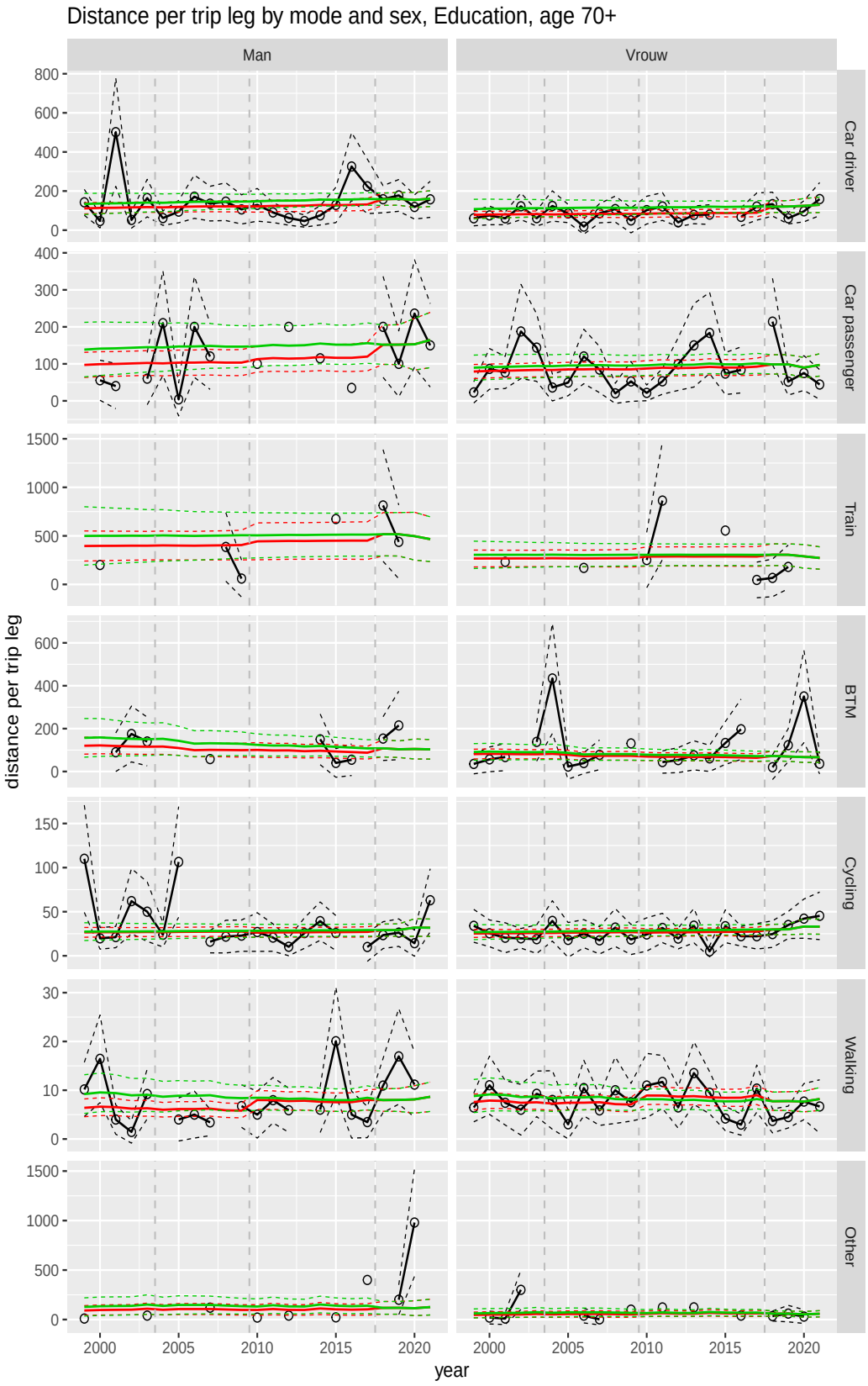

Figure A.148 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

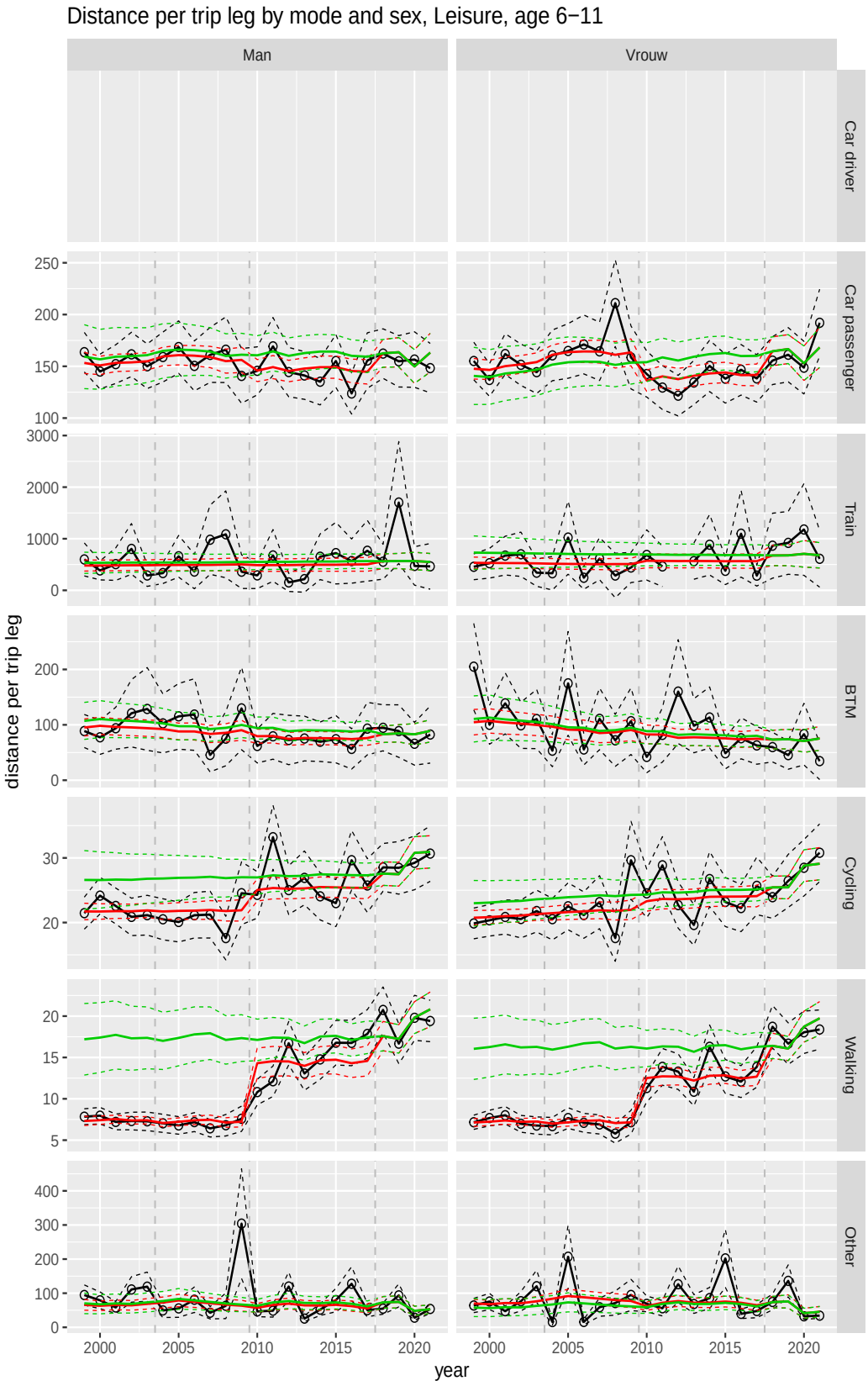

Figure A.149 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

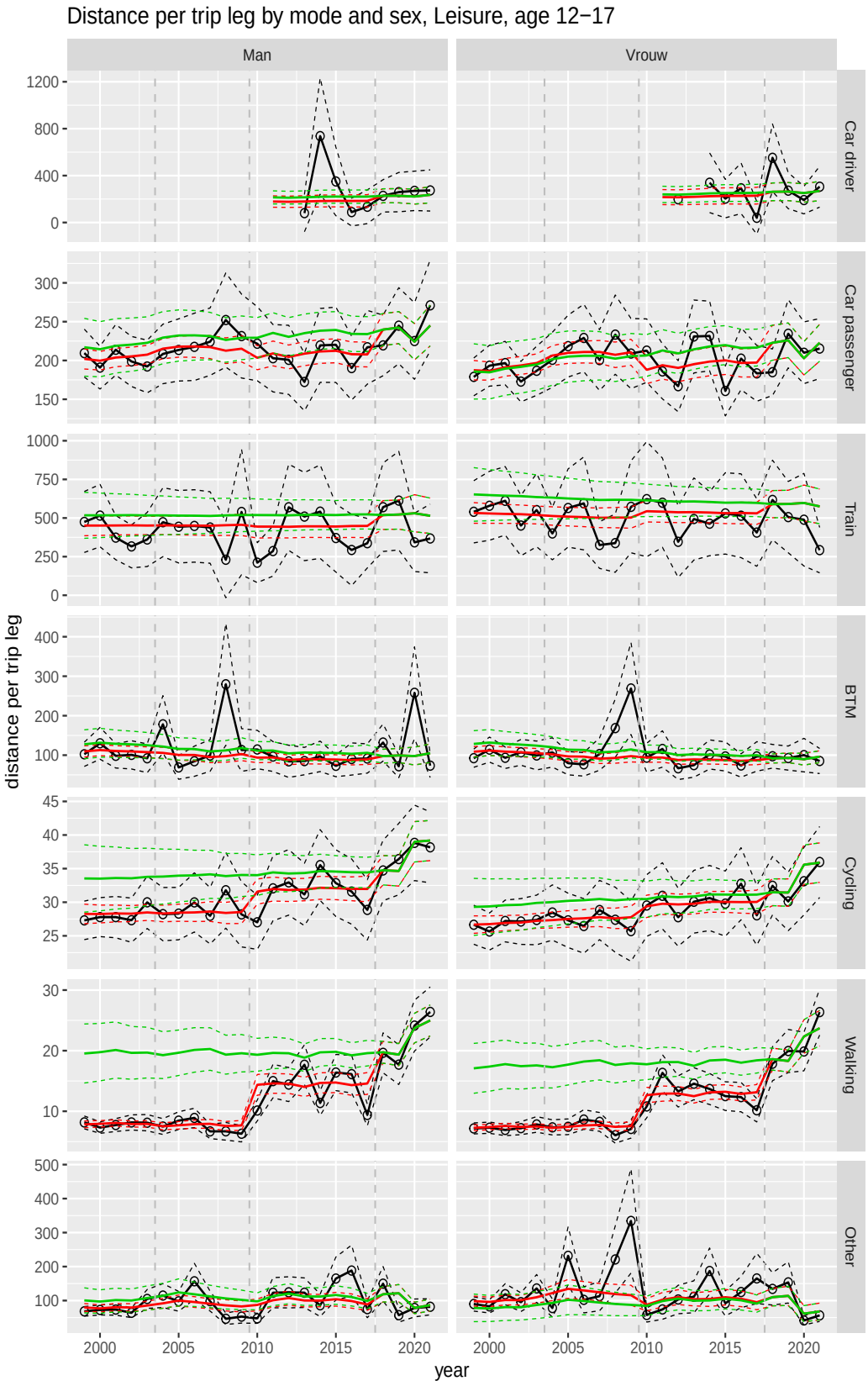

Figure A.150 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

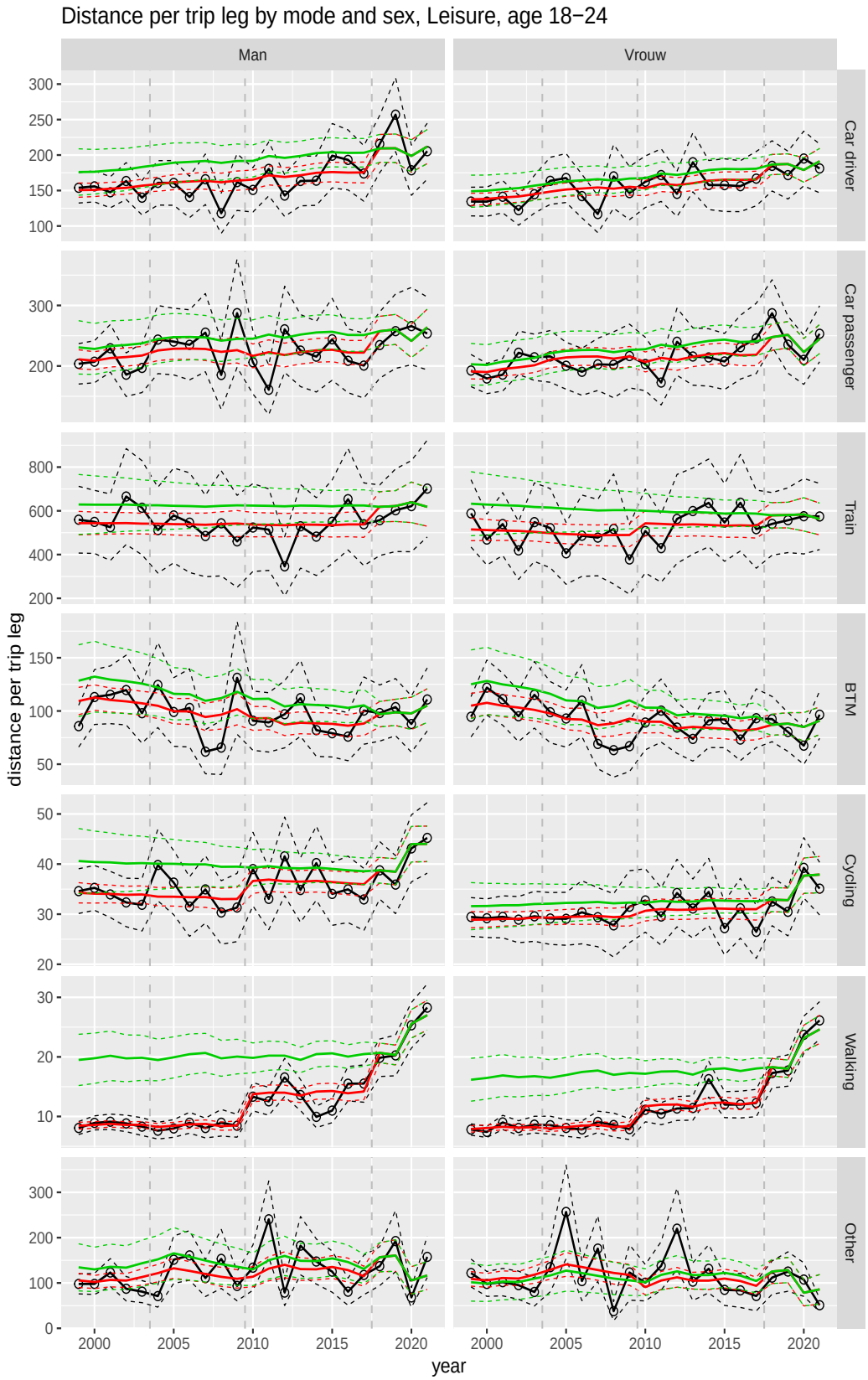

Figure A.151 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

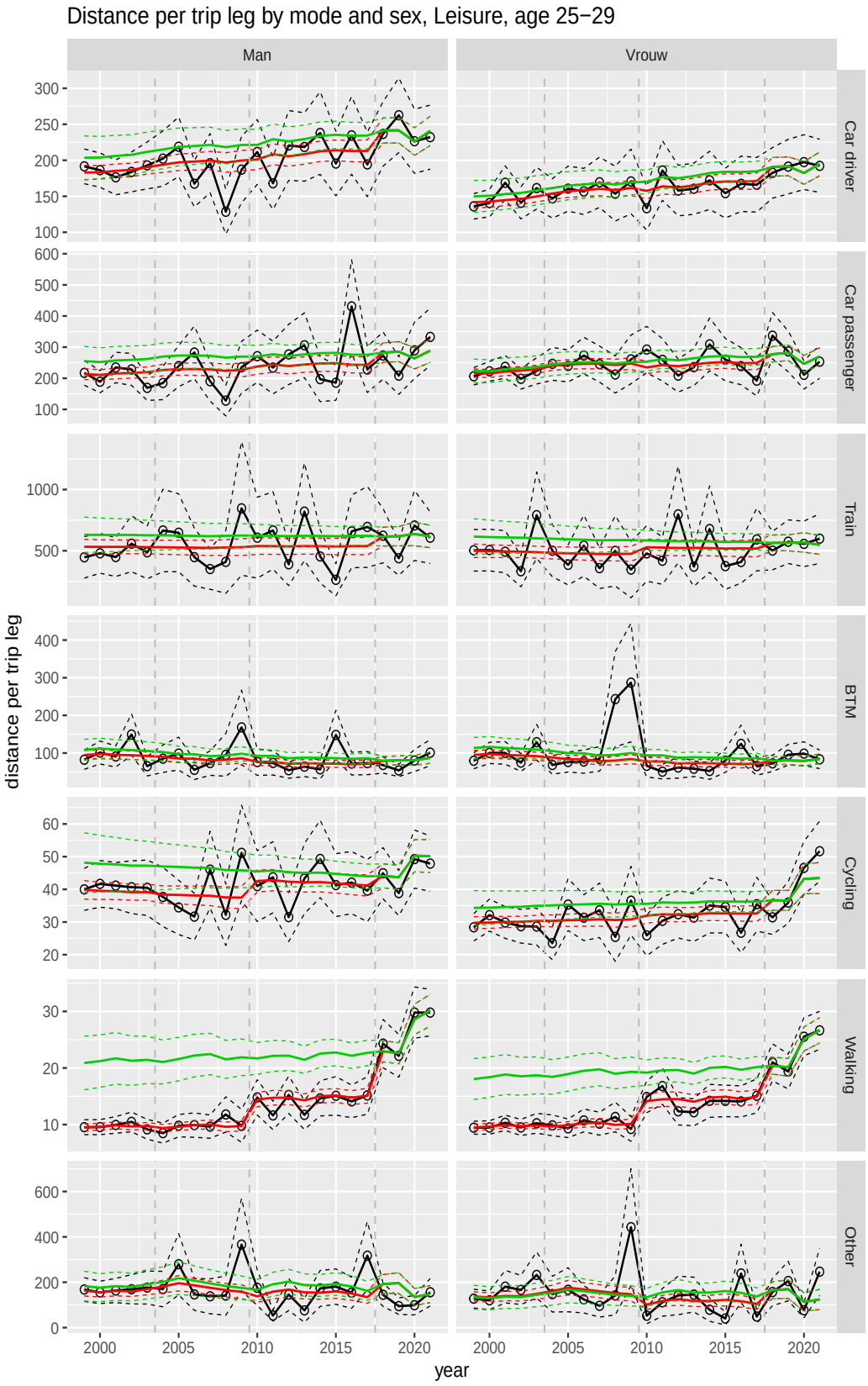

Figure A.152 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

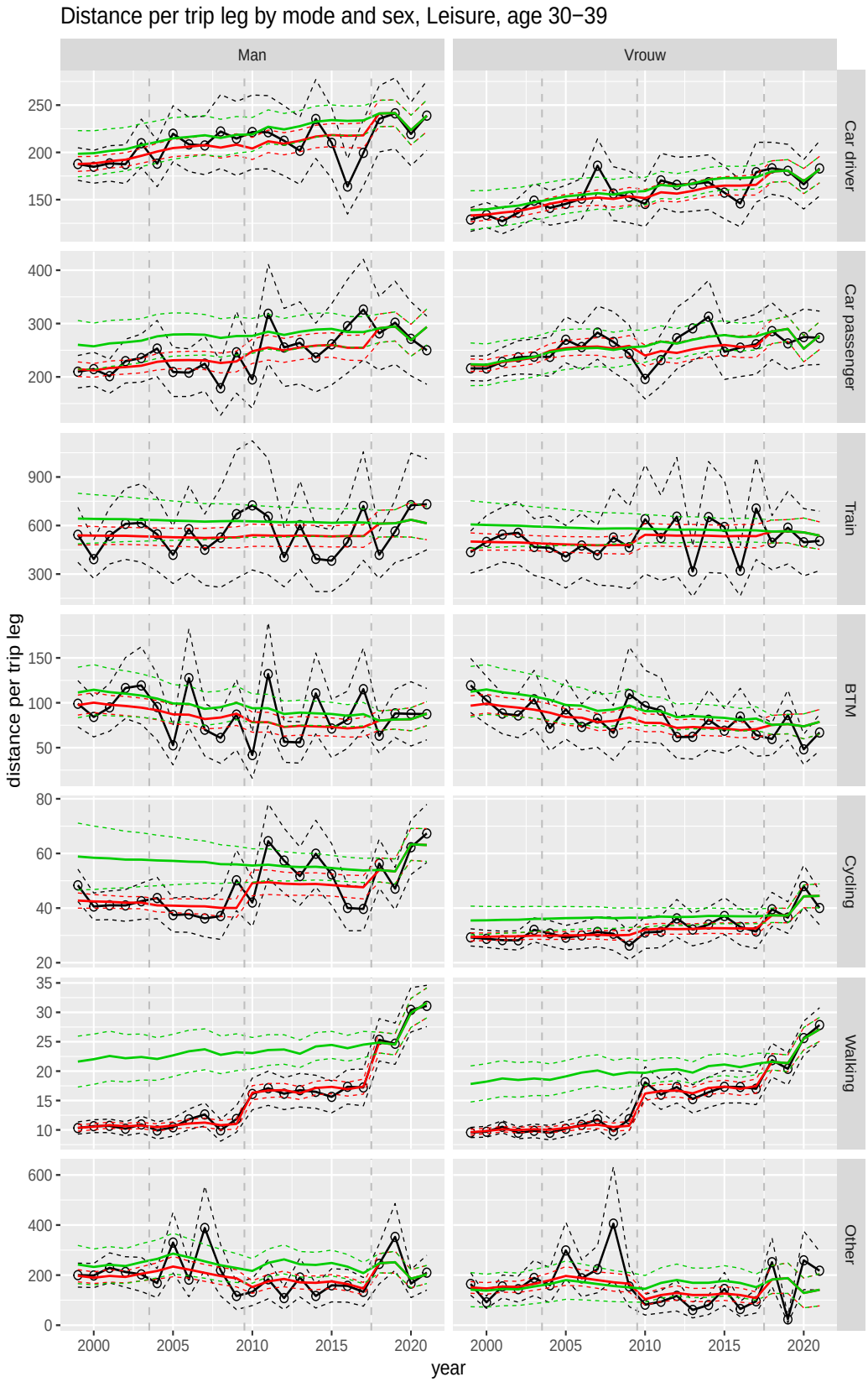

Figure A.153 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

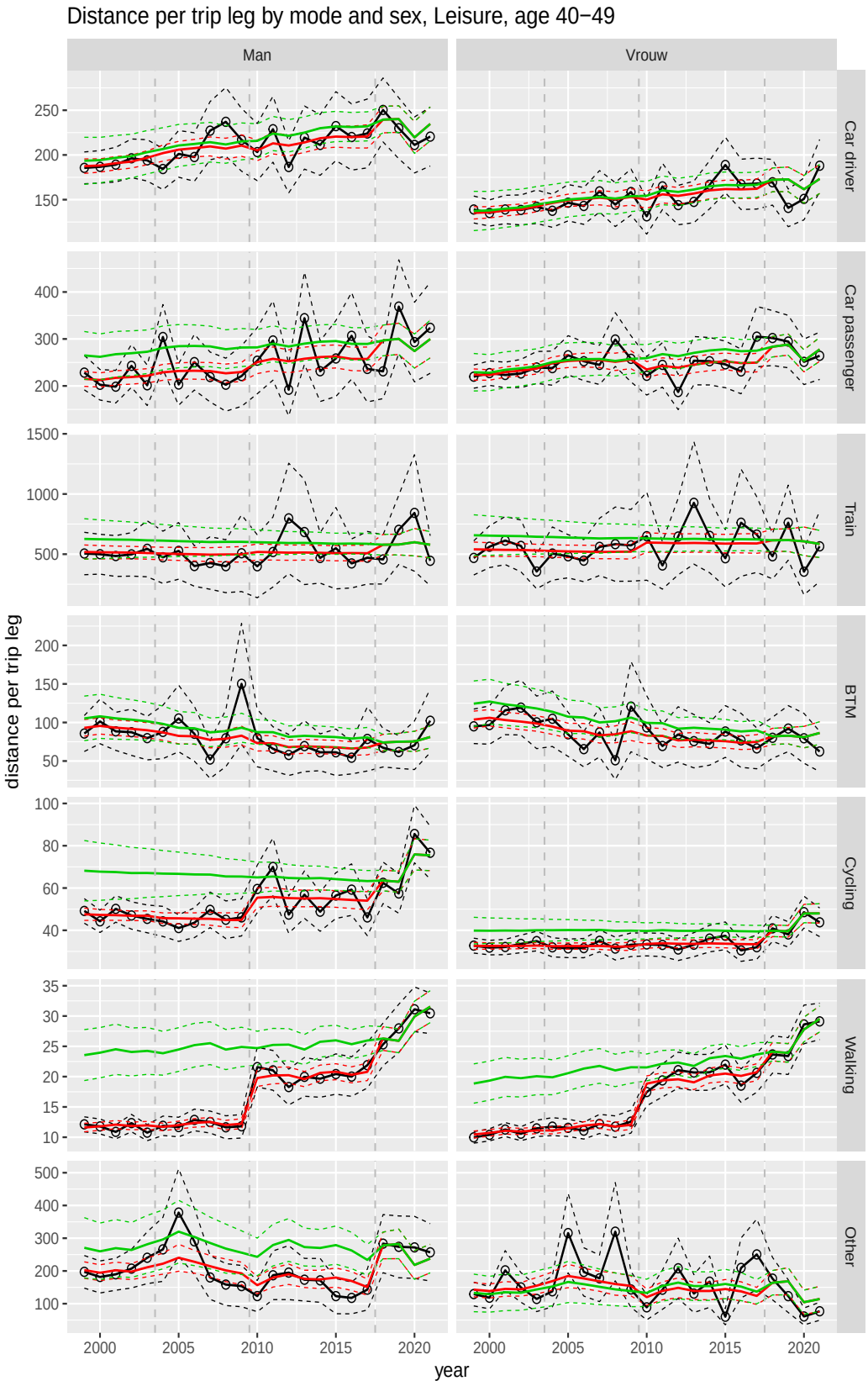

Figure A.154 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.155 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.156 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

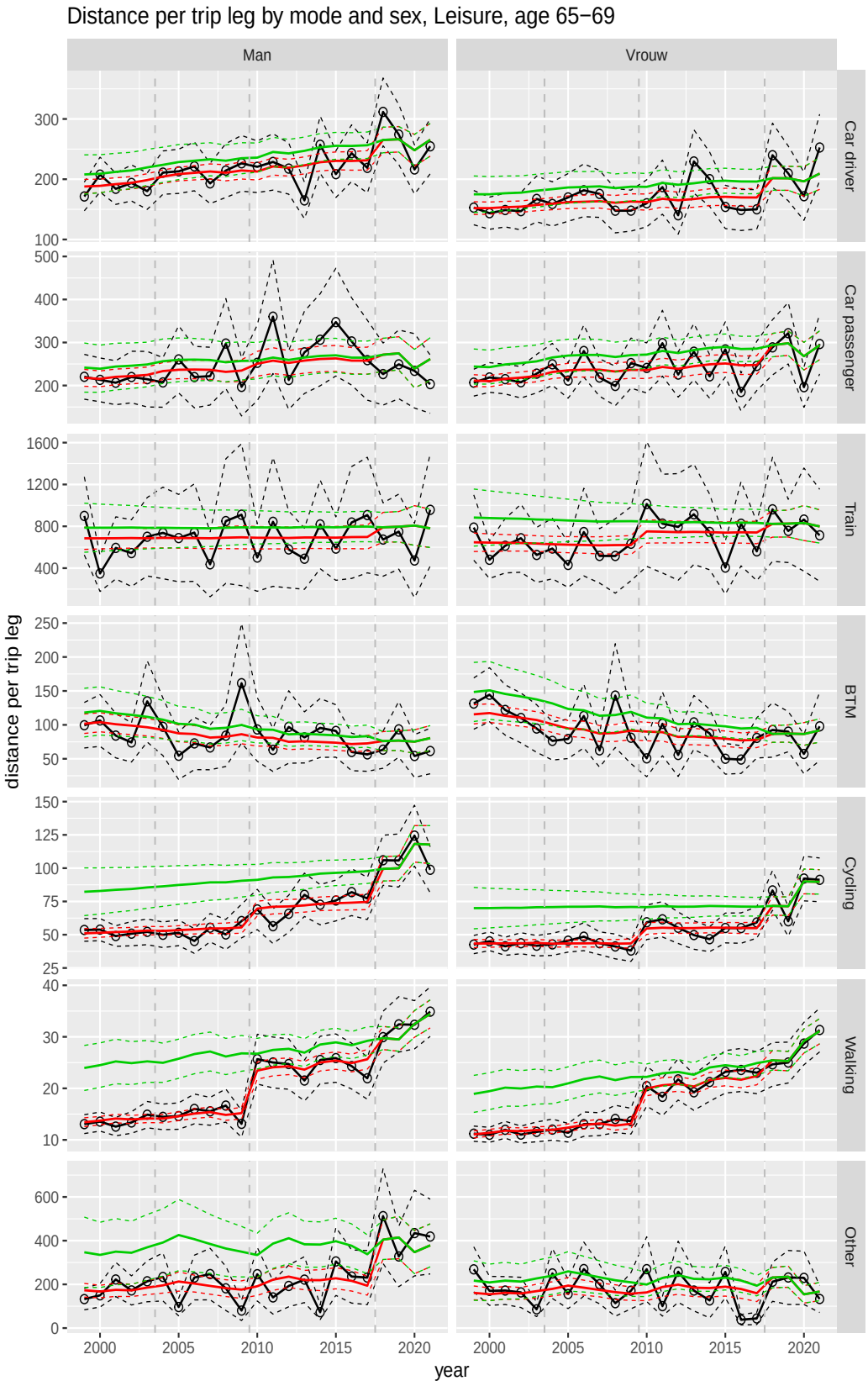

Figure A.157 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

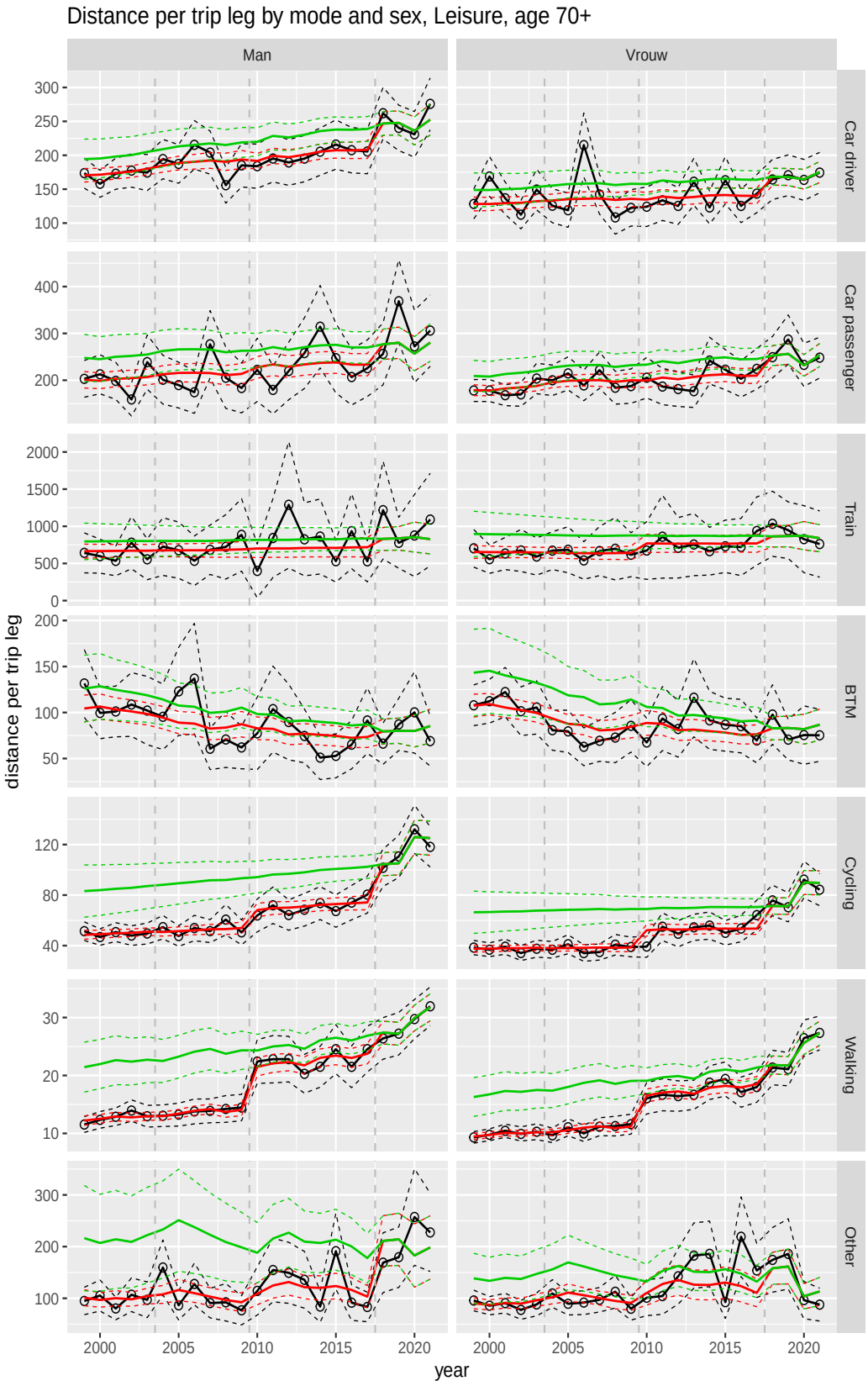

Figure A.158 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

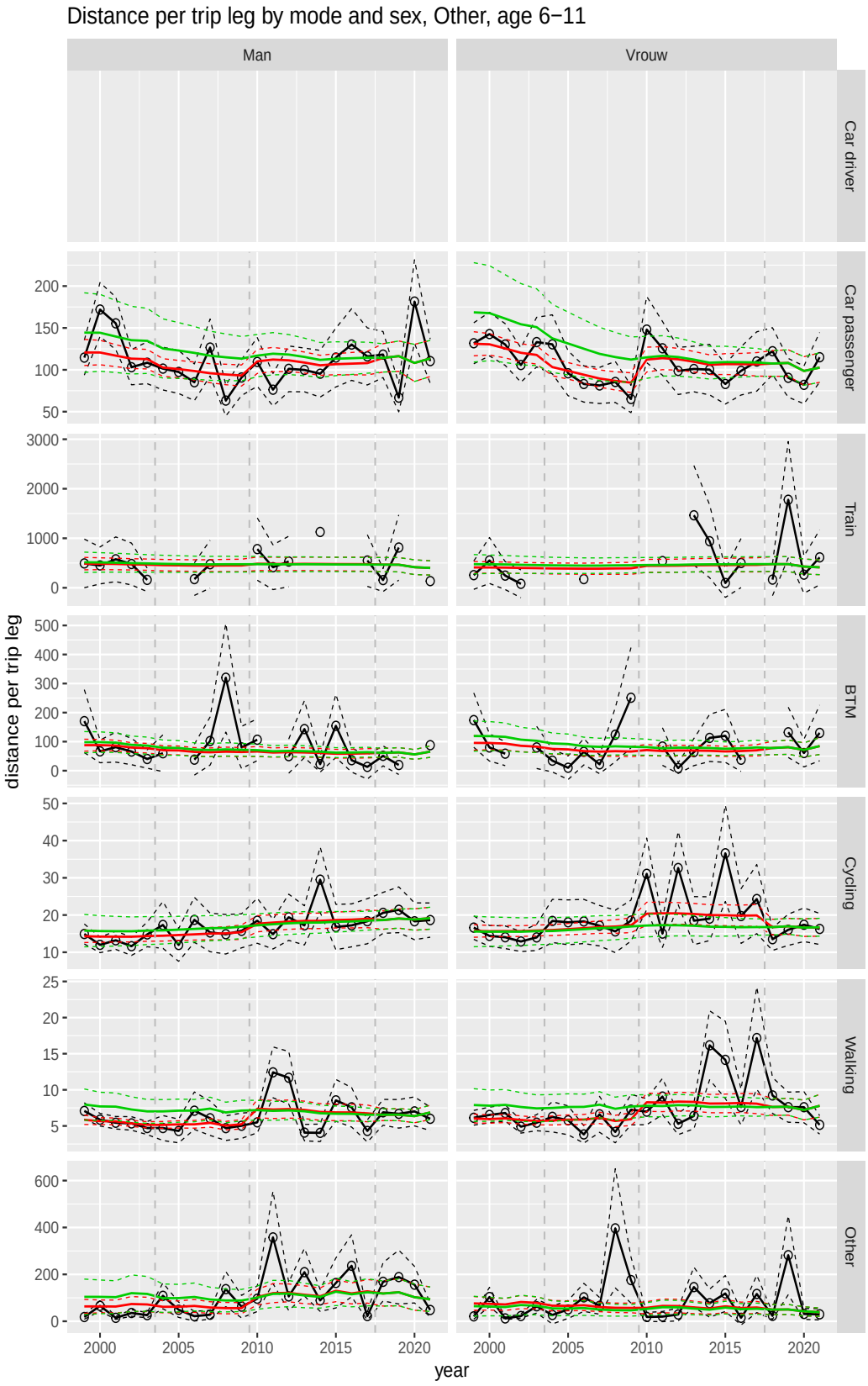

Figure A.159 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

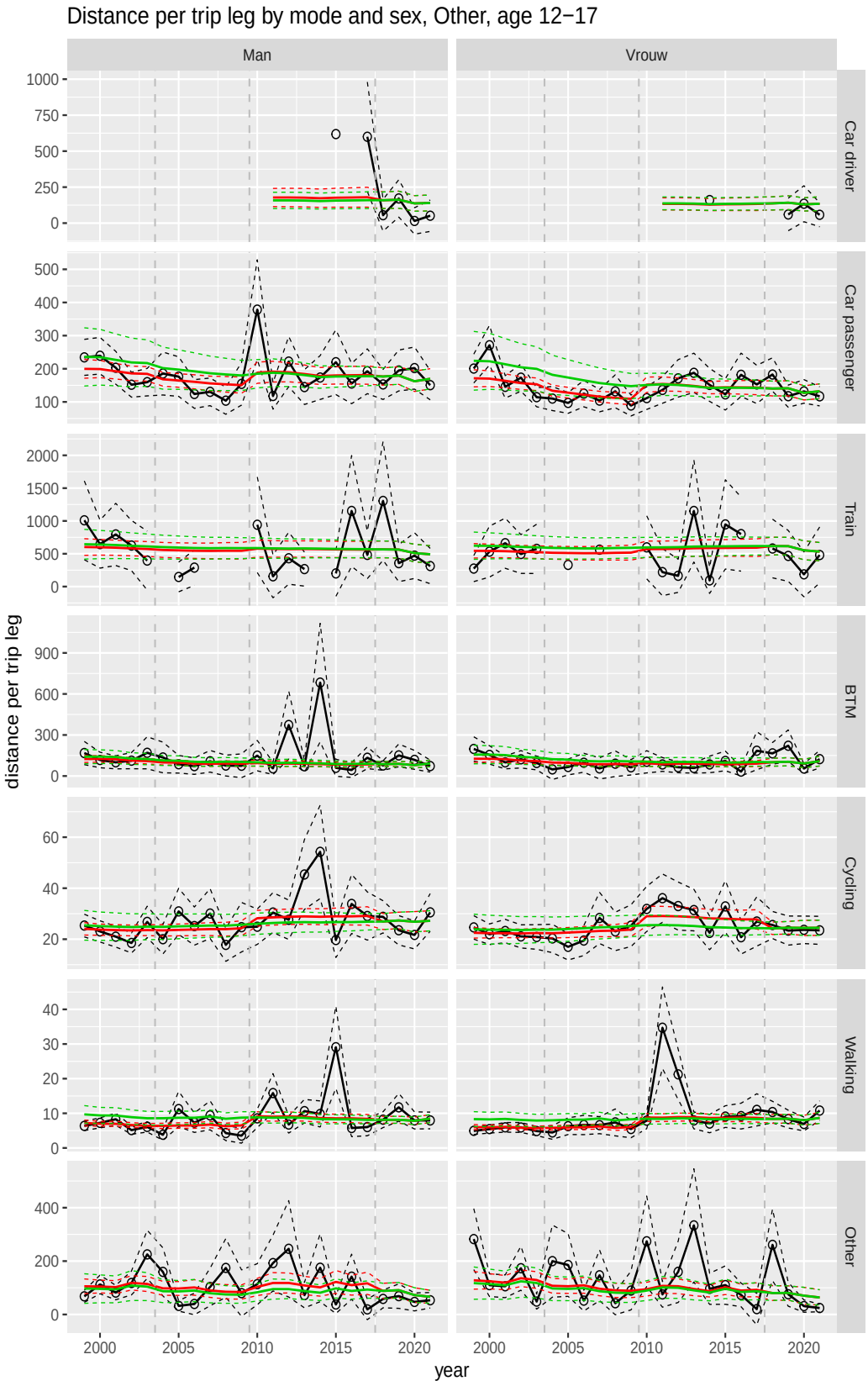

Figure A.160 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

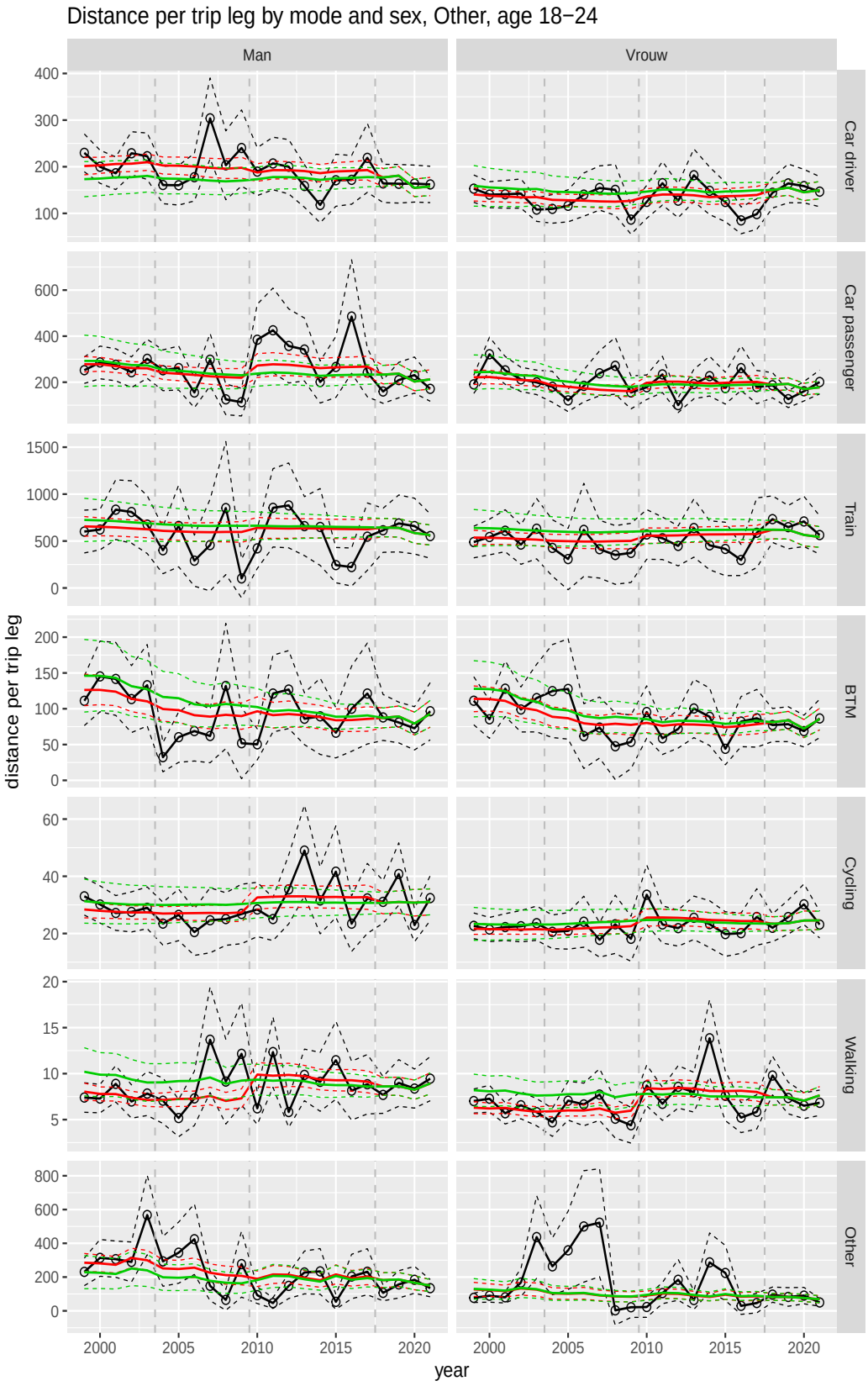

Figure A.161 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

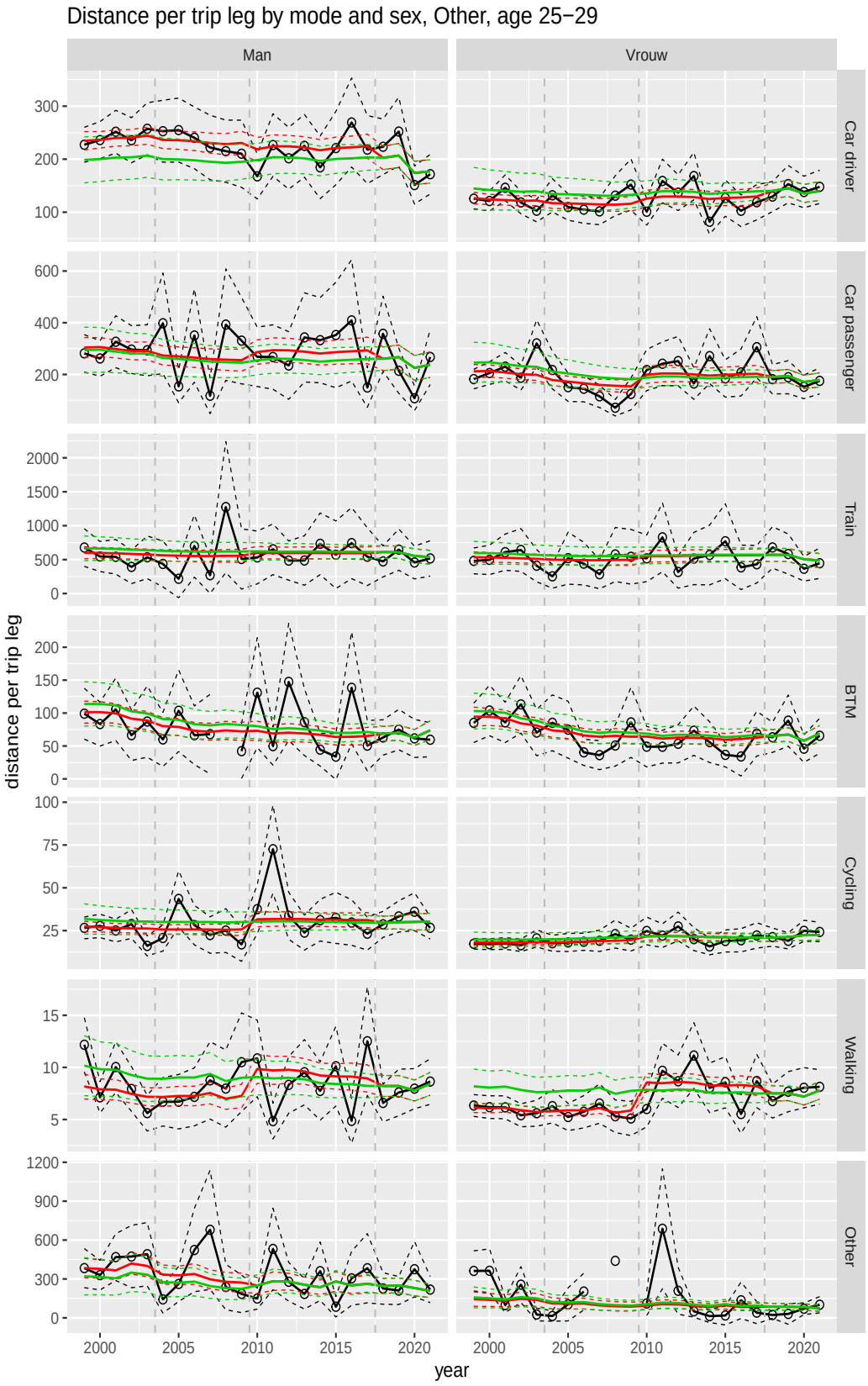

Figure A.162 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

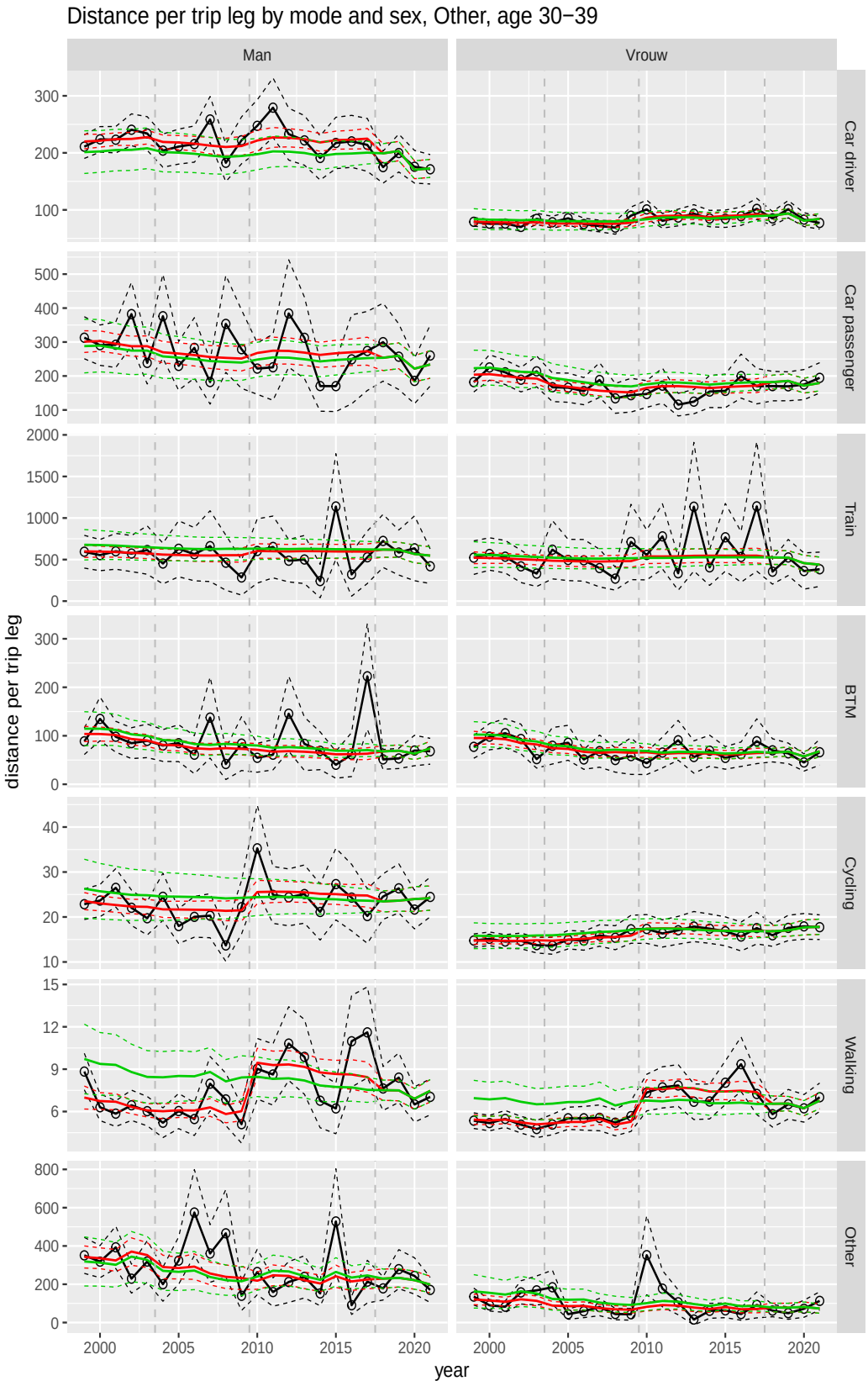

Figure A.163 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

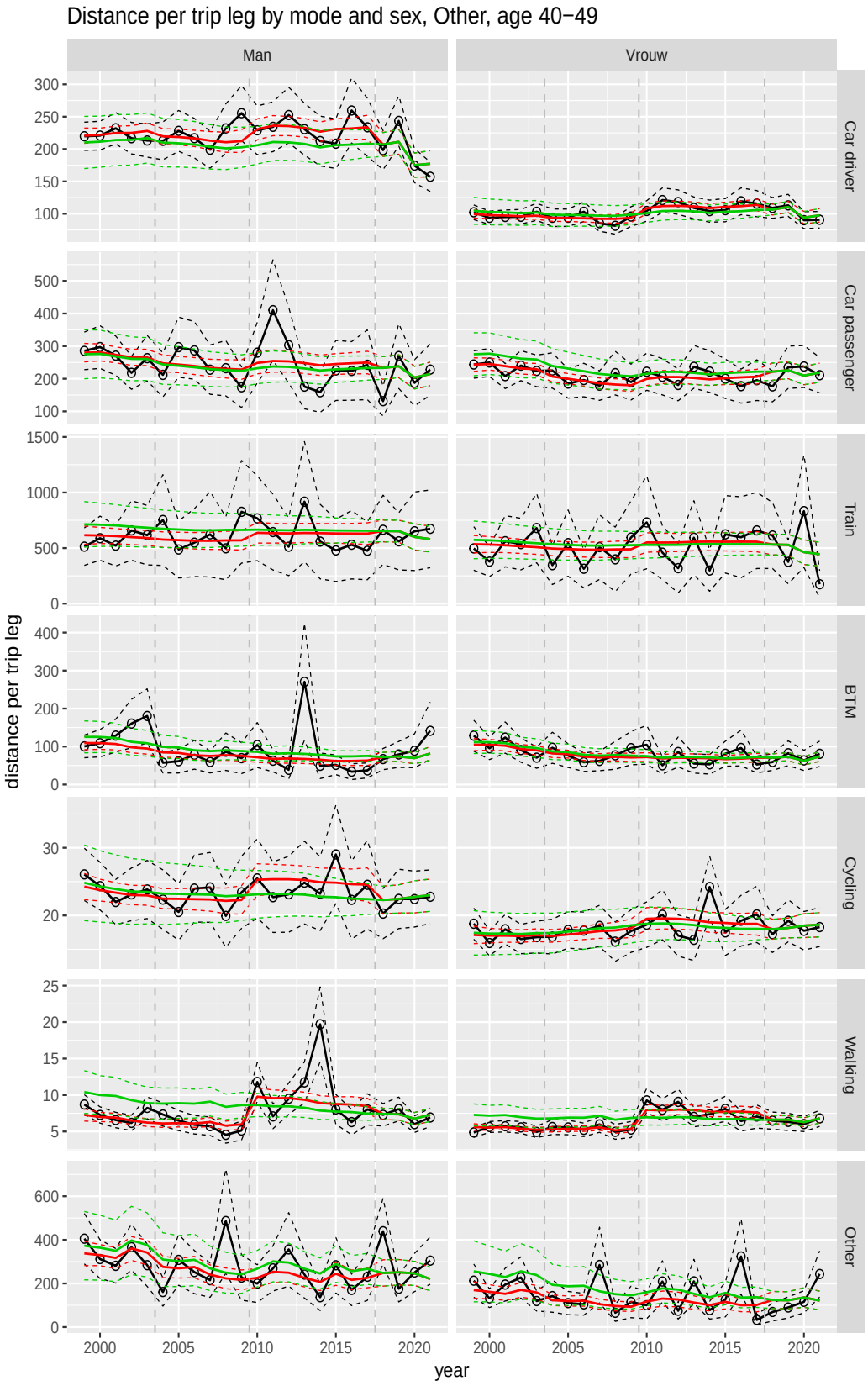

Figure A.164 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

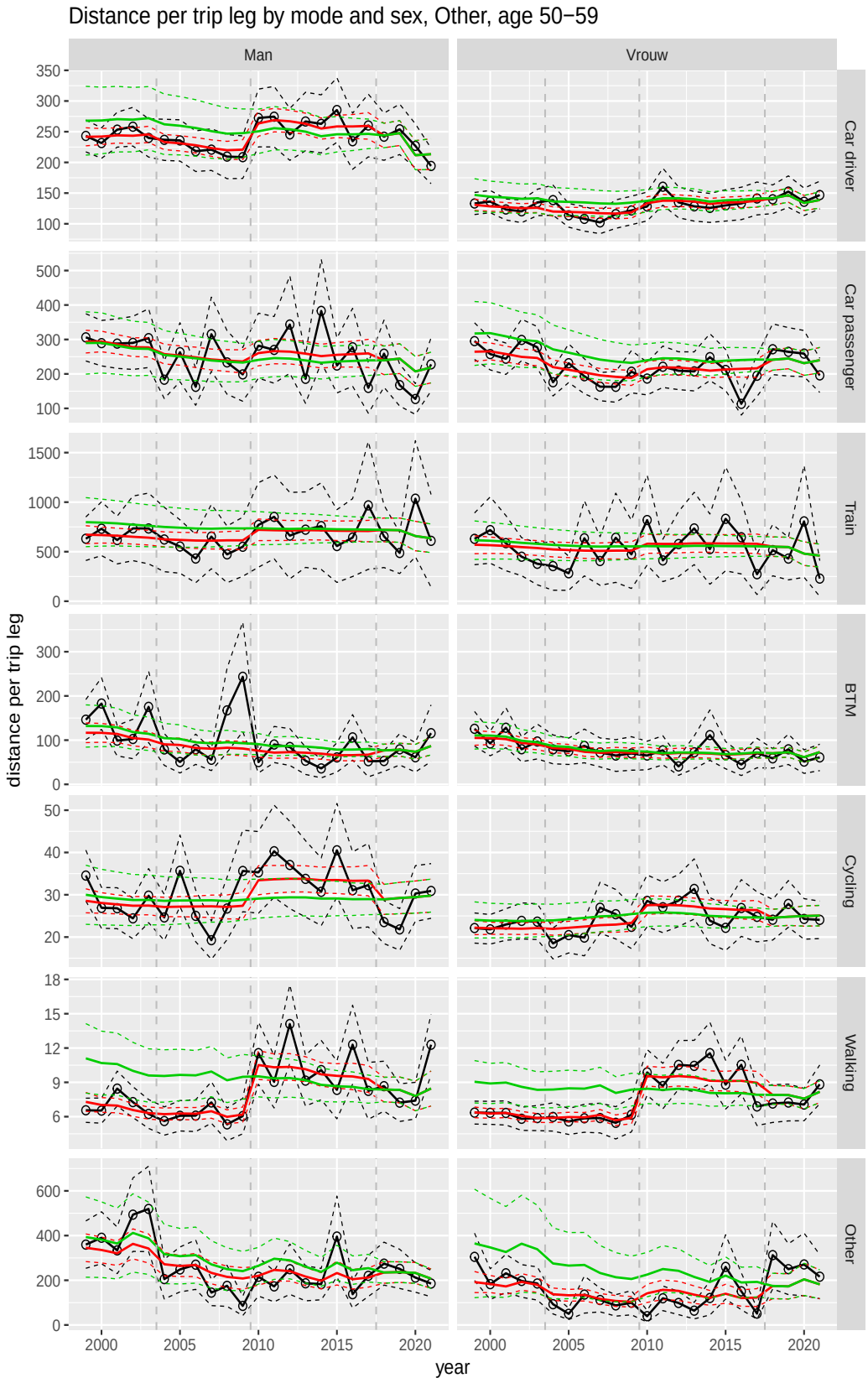

Figure A.165 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

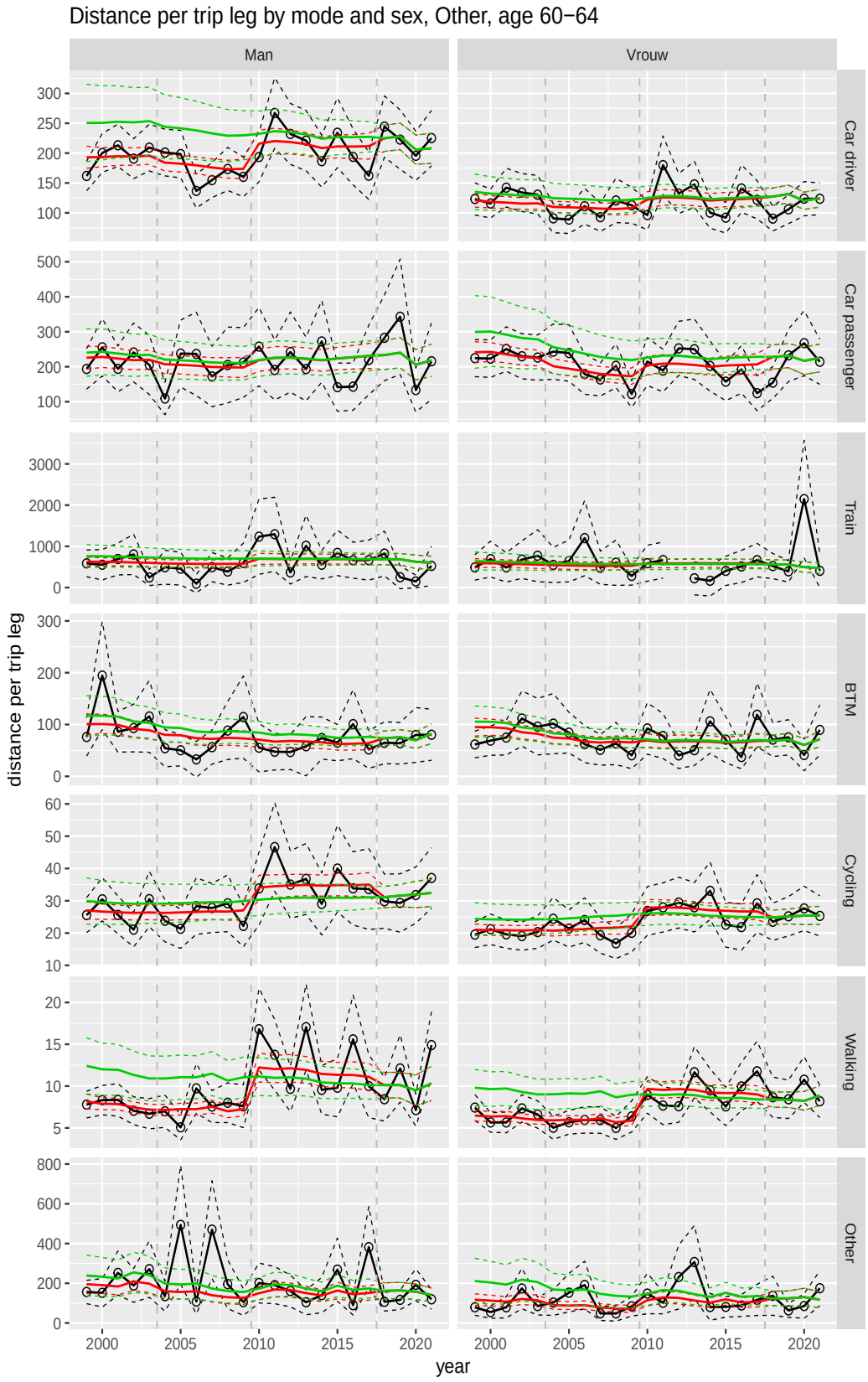

Figure A.166 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

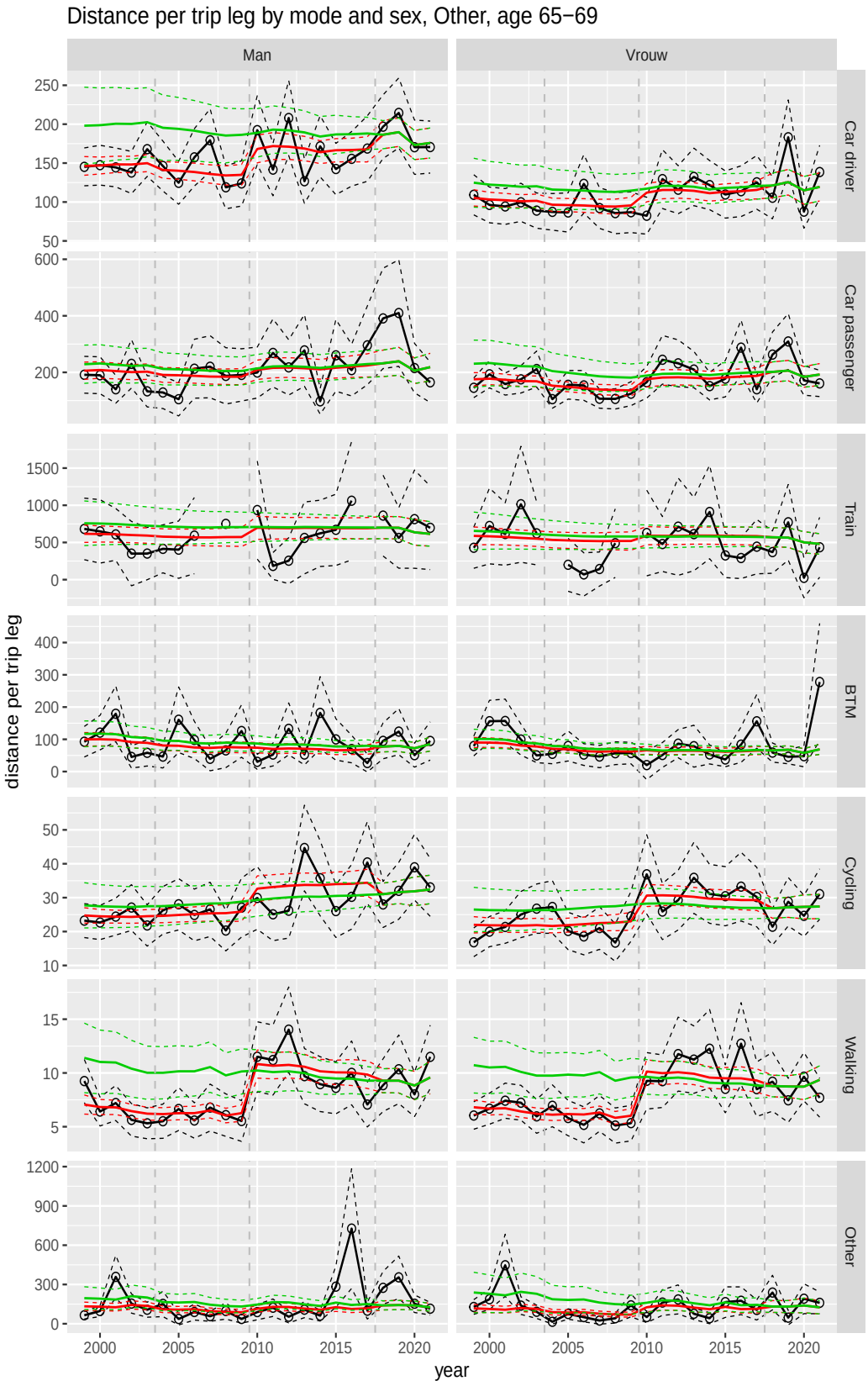

Figure A.167 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

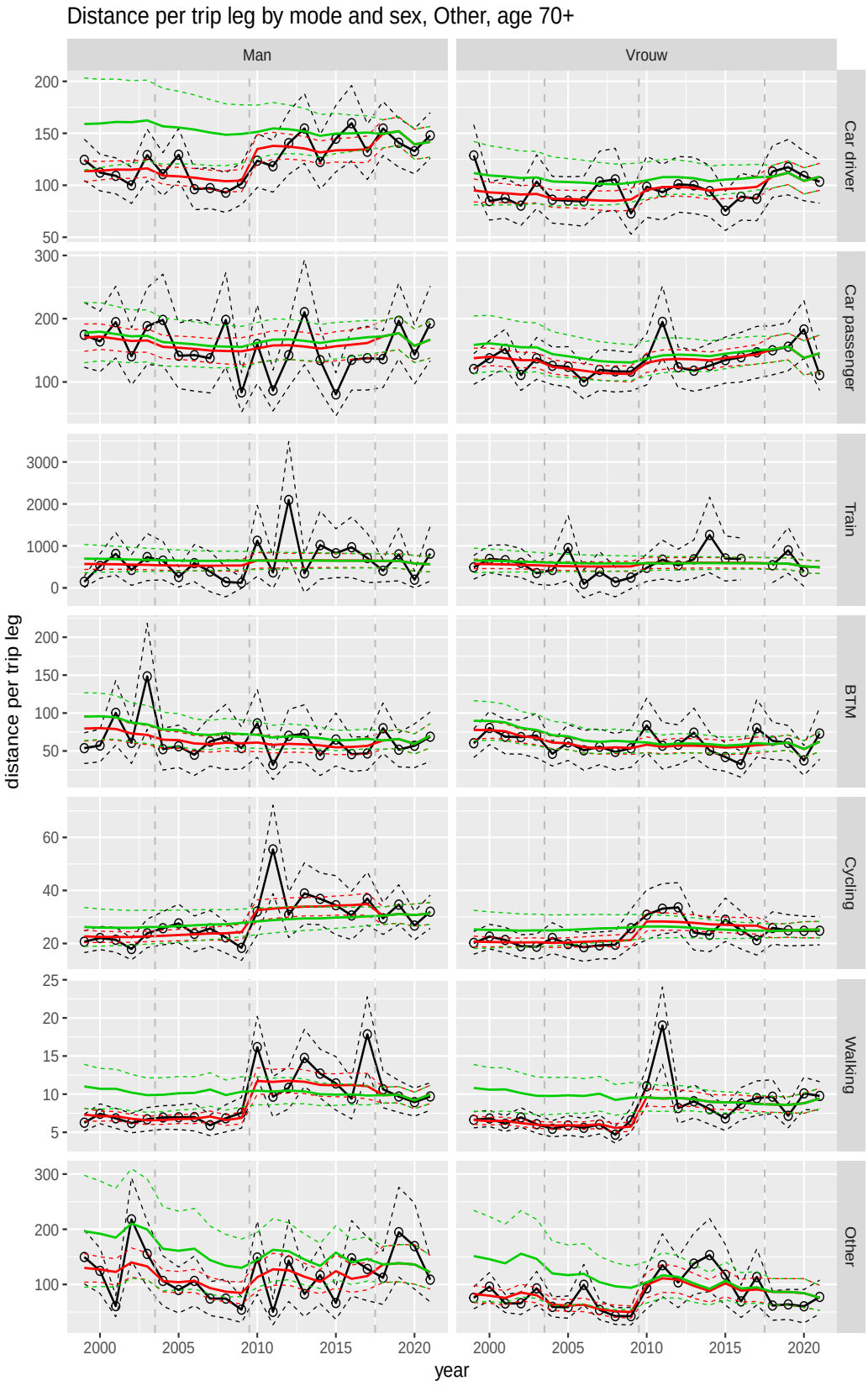

Figure A.168 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

A.5 Average distance per person per day

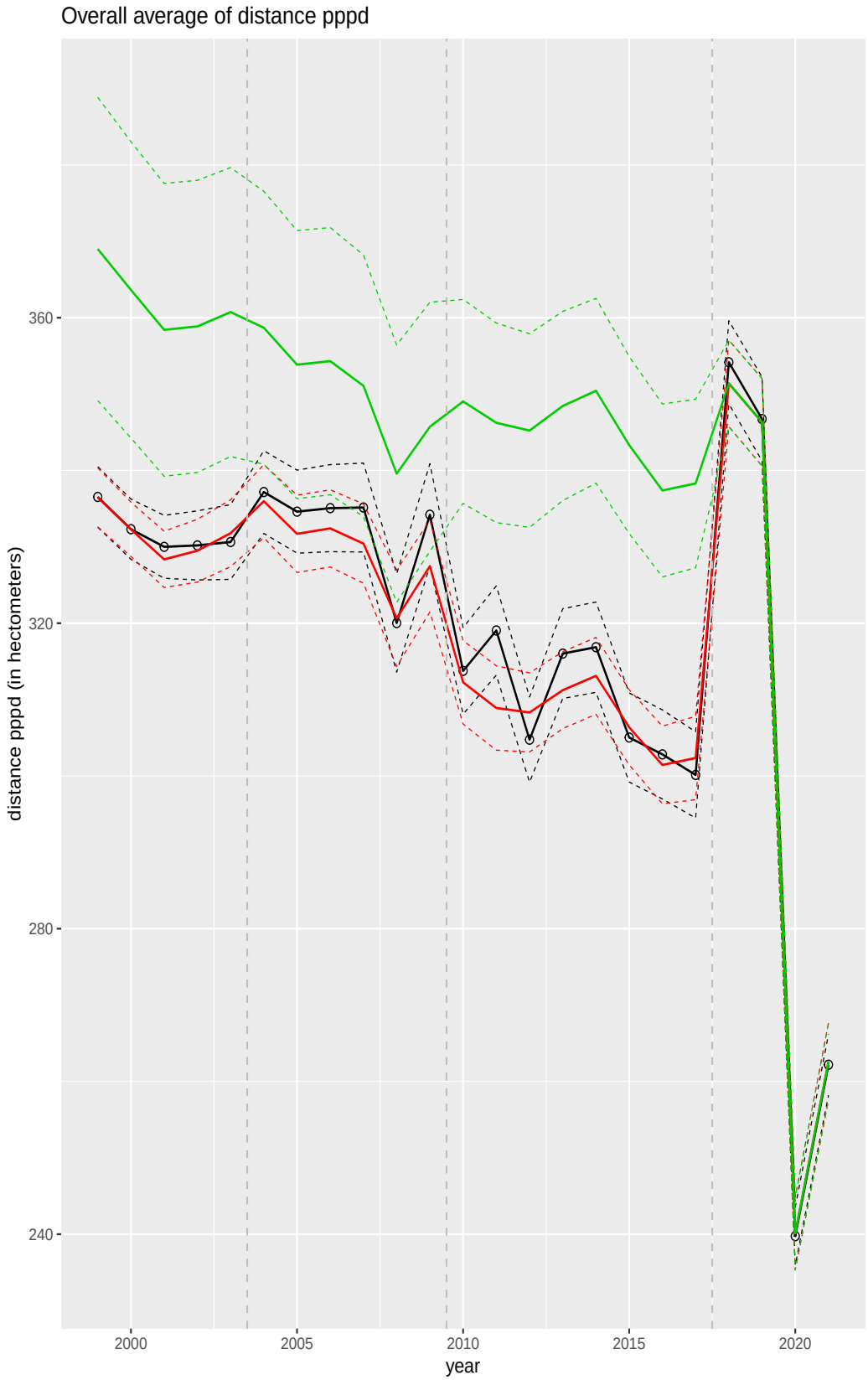

Figure A.169 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

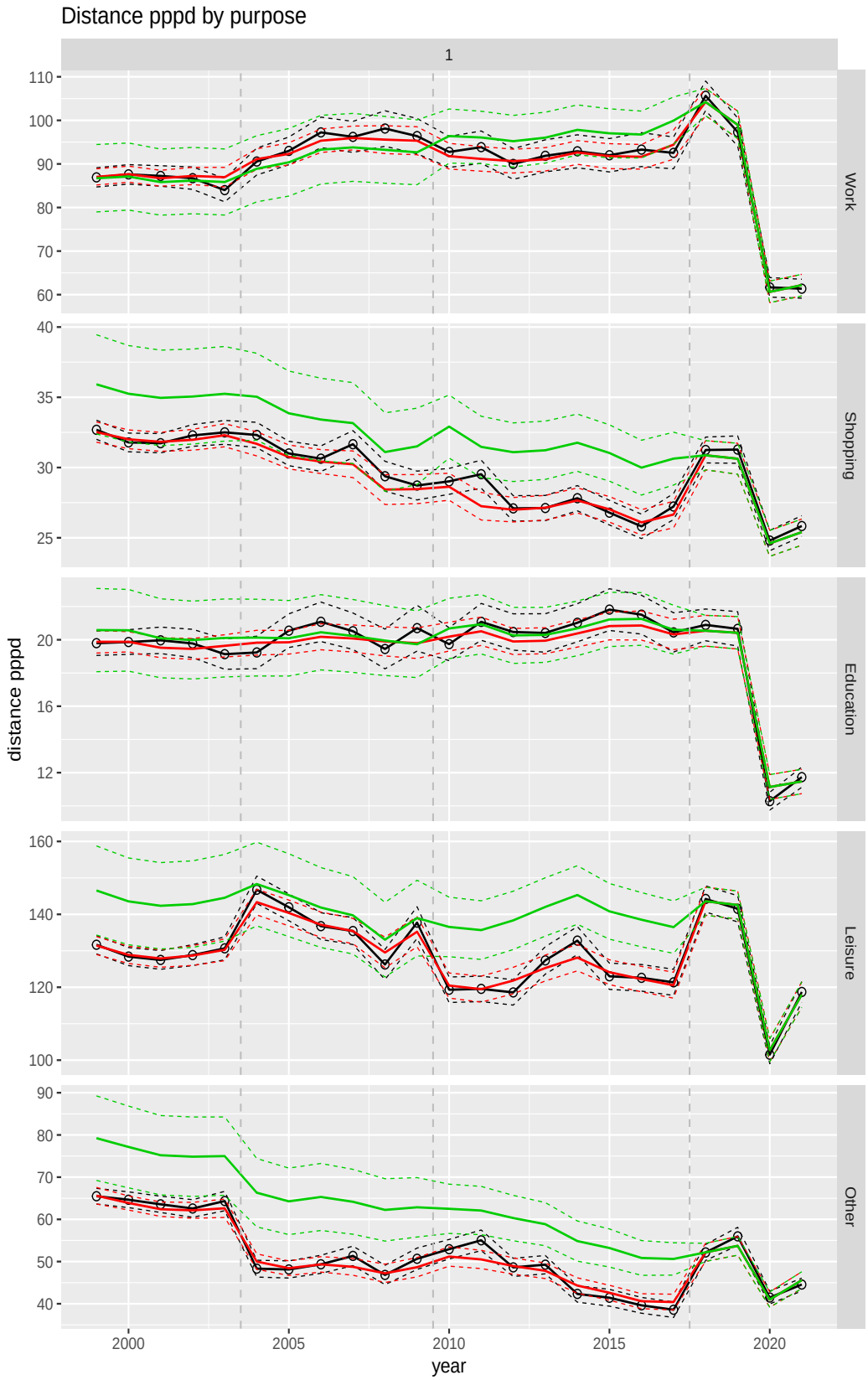

Figure A.170 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

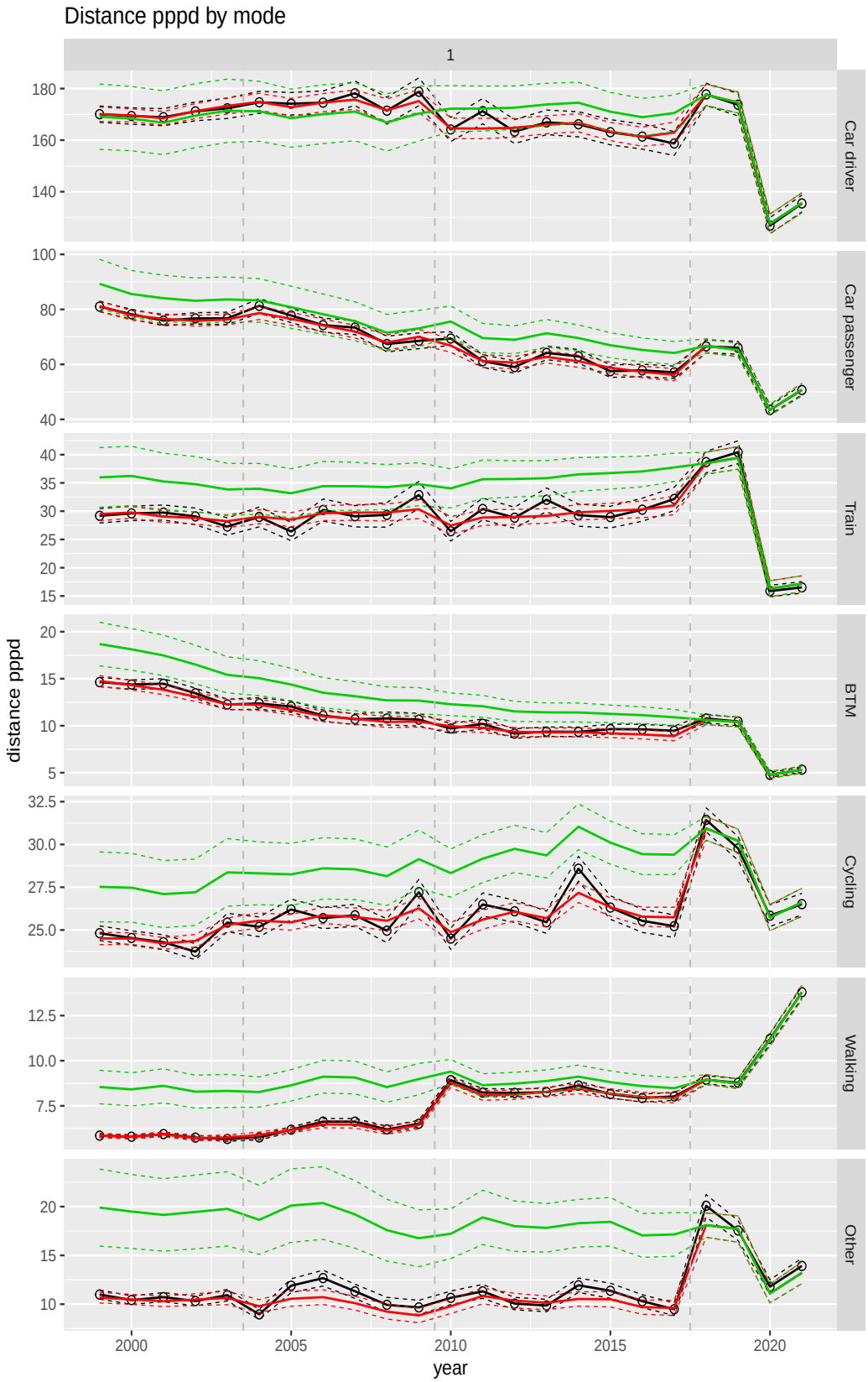

Figure A.171 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

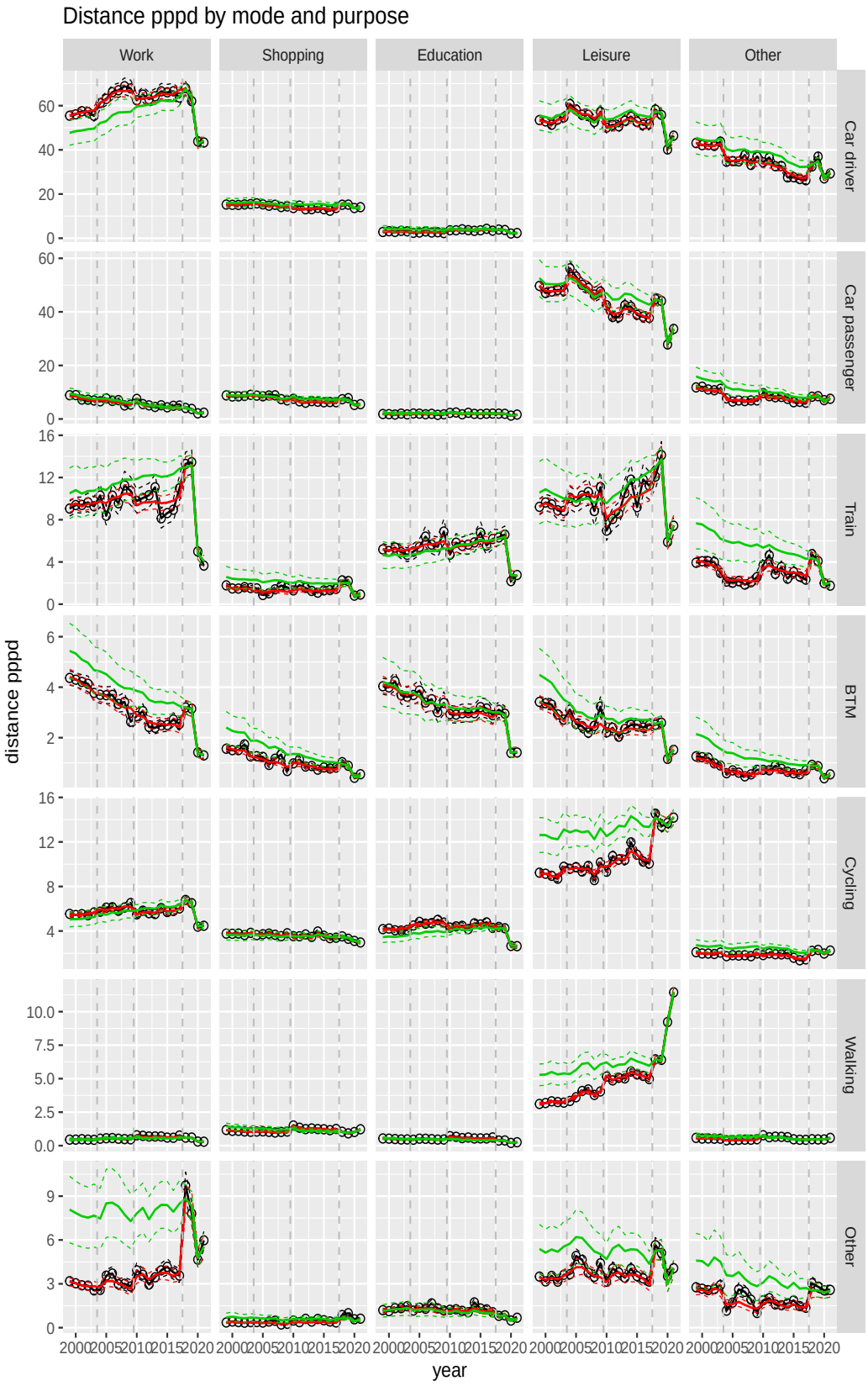

Figure A.172 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

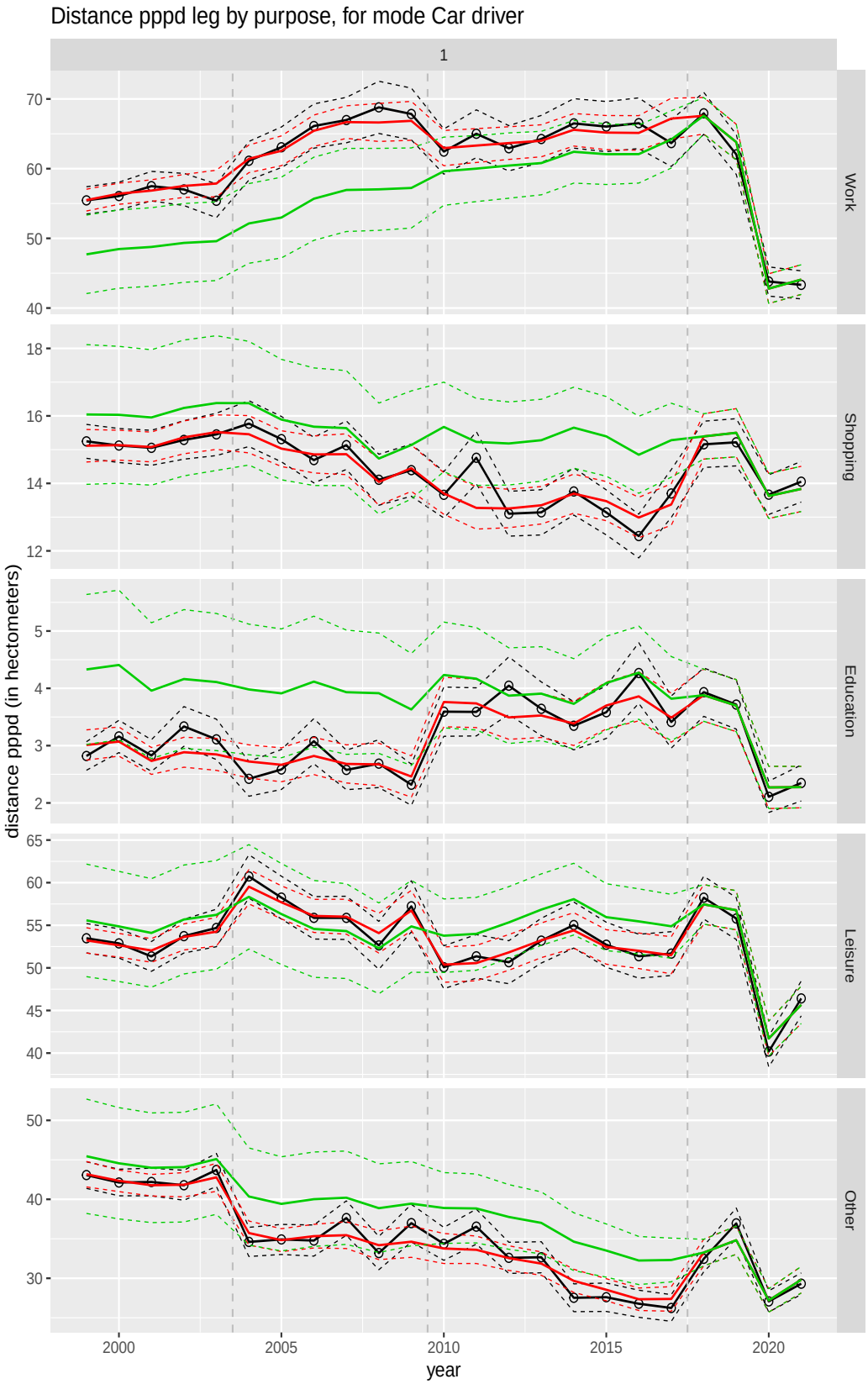

Figure A.173 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

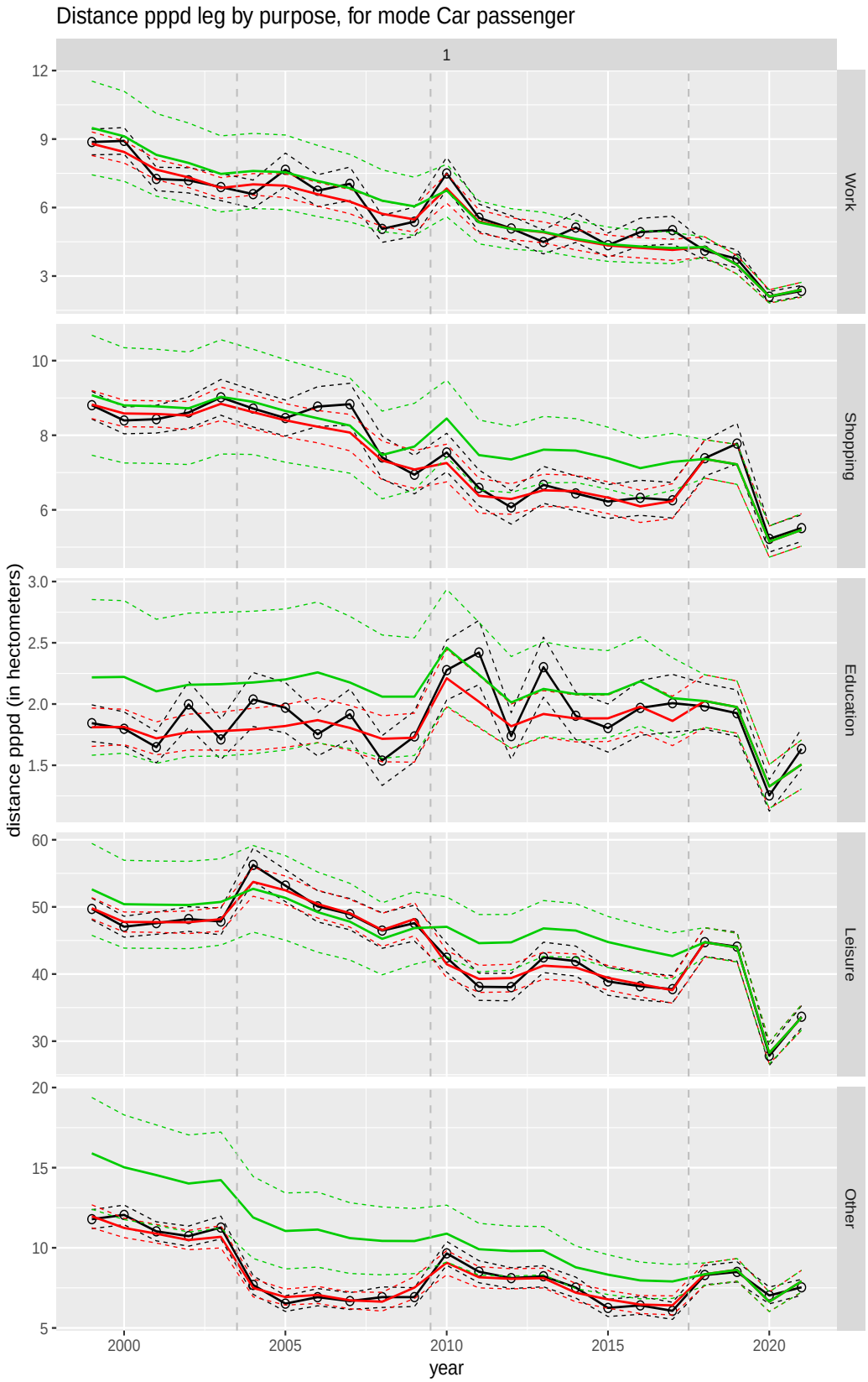

Figure A.174 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

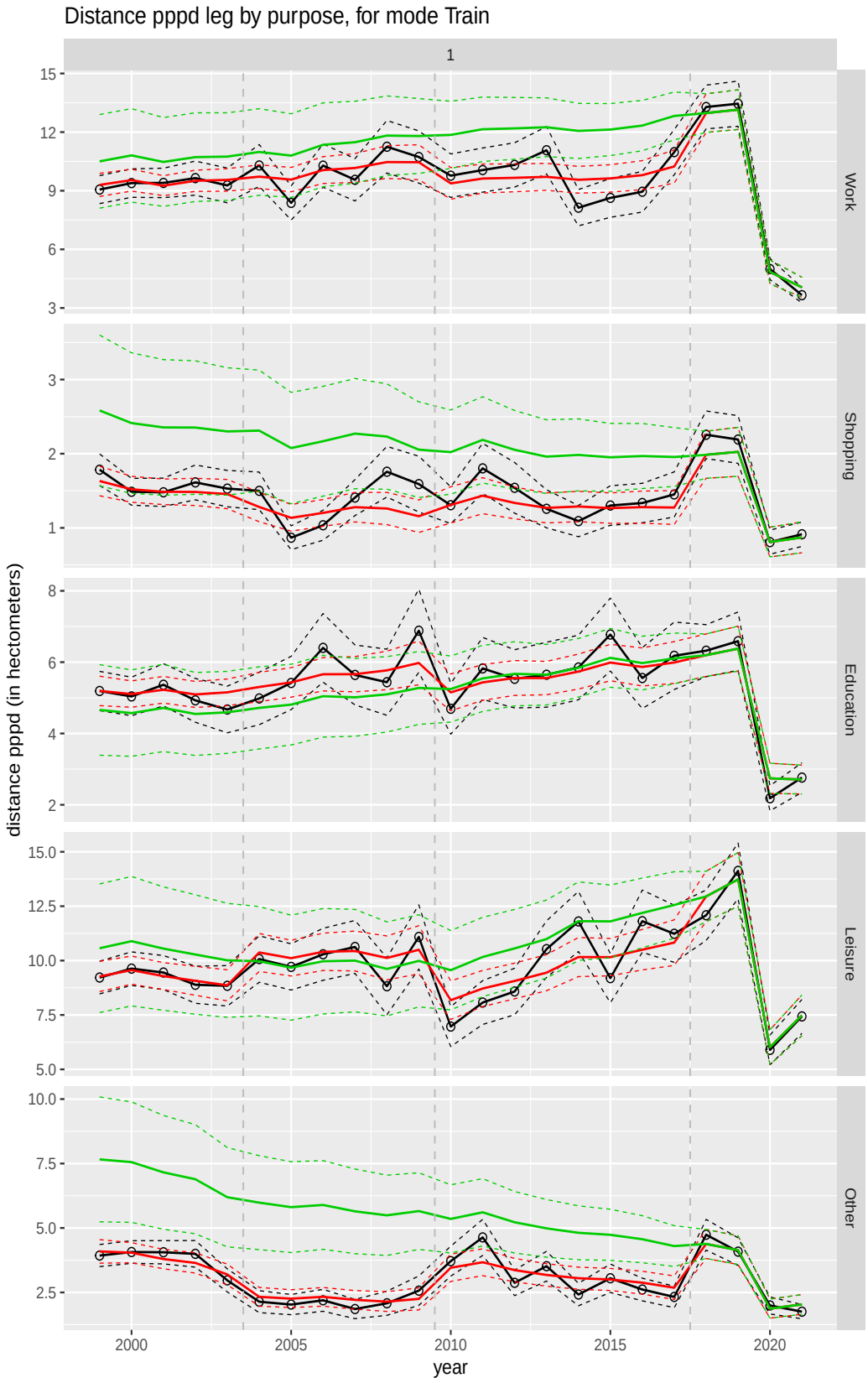

Figure A.175 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

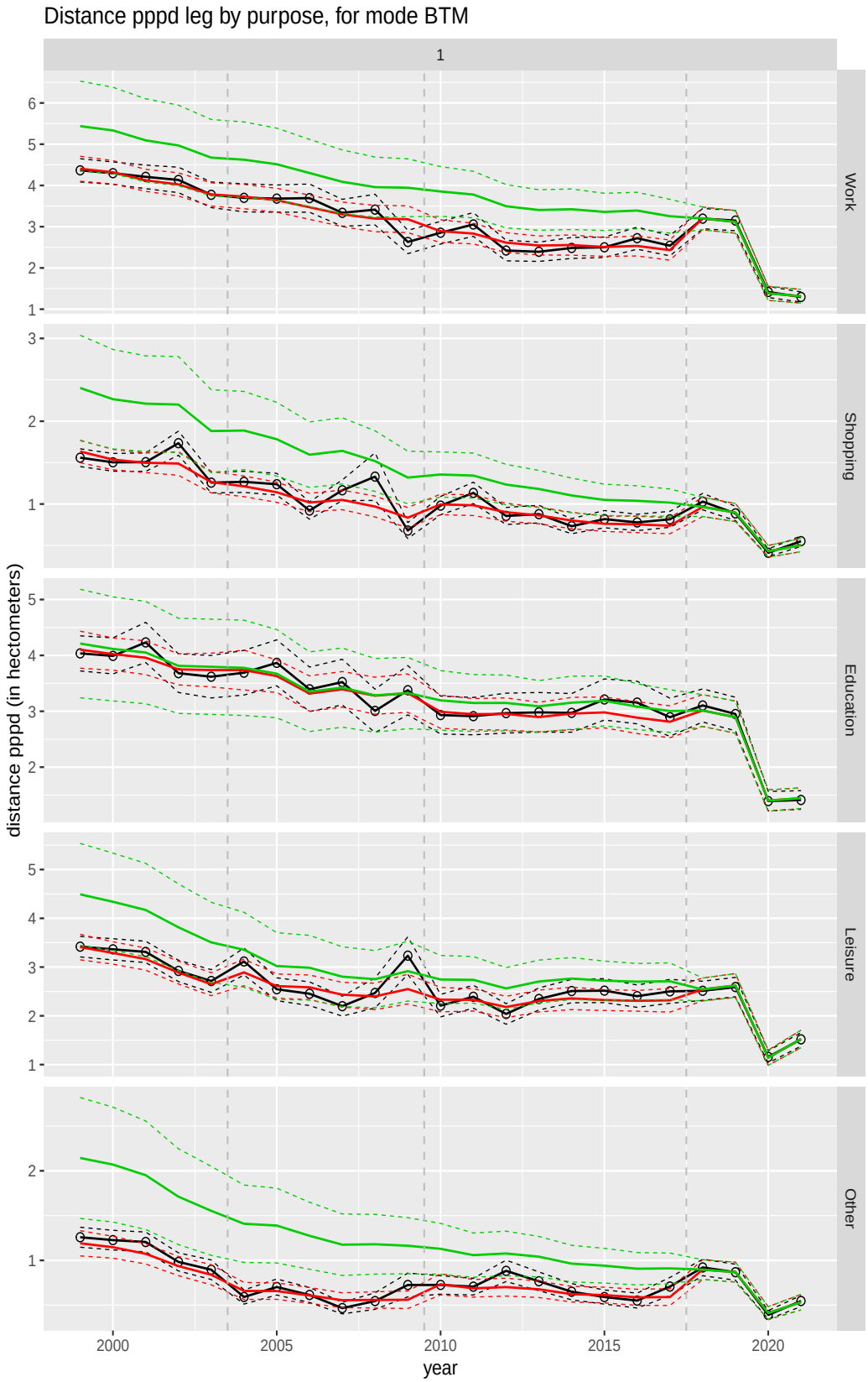

Figure A.176 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

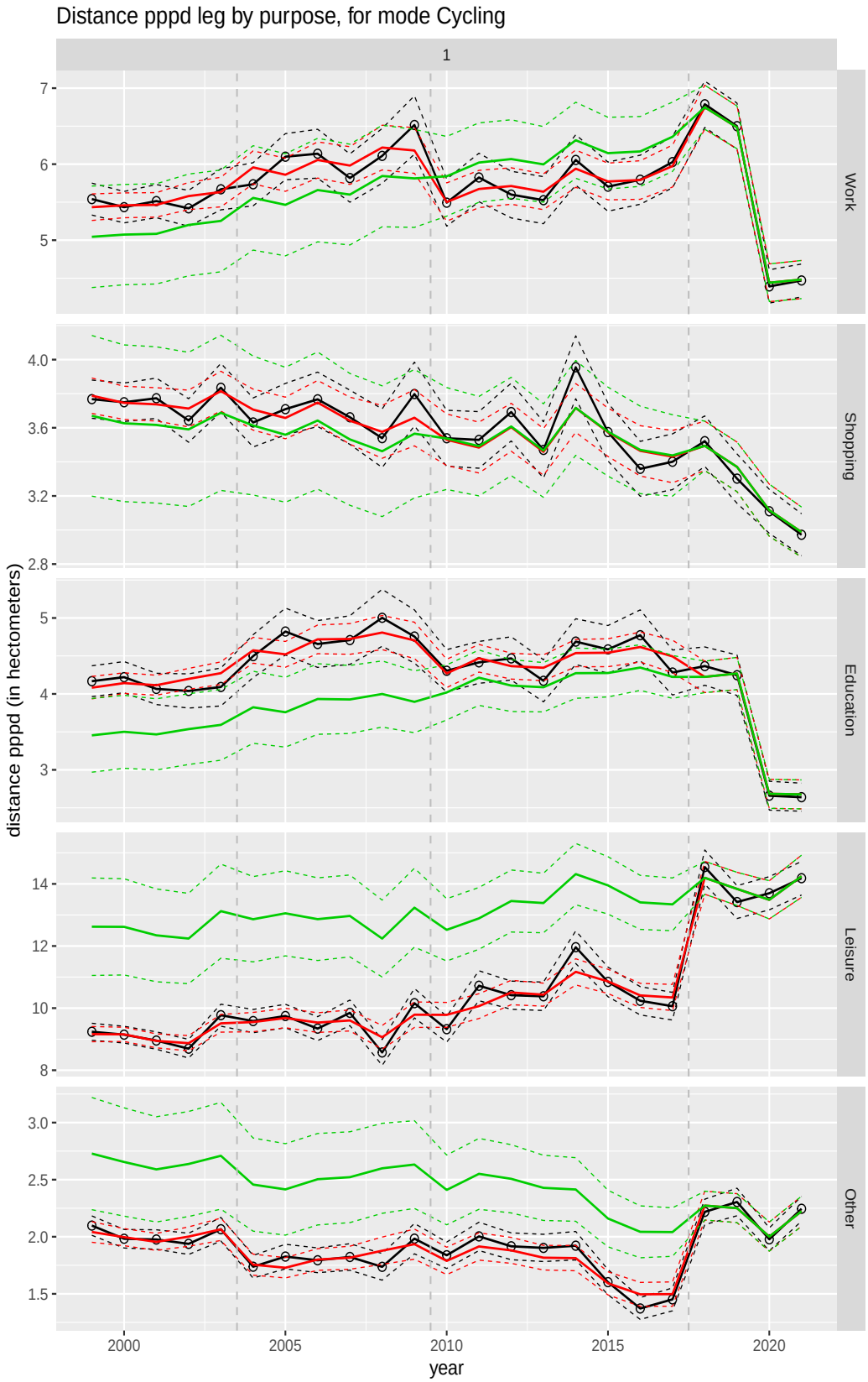

Figure A.177 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

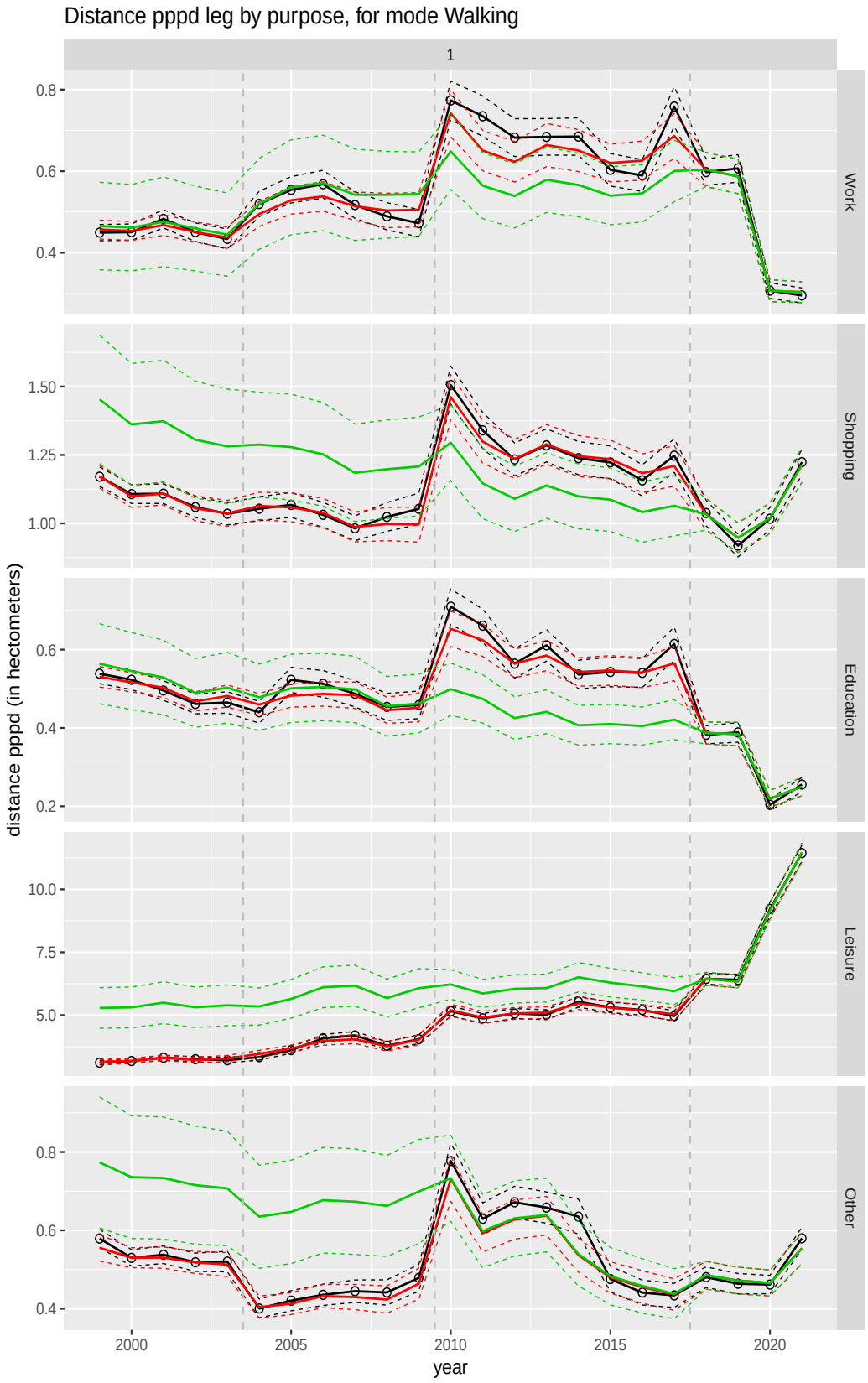

Figure A.178 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

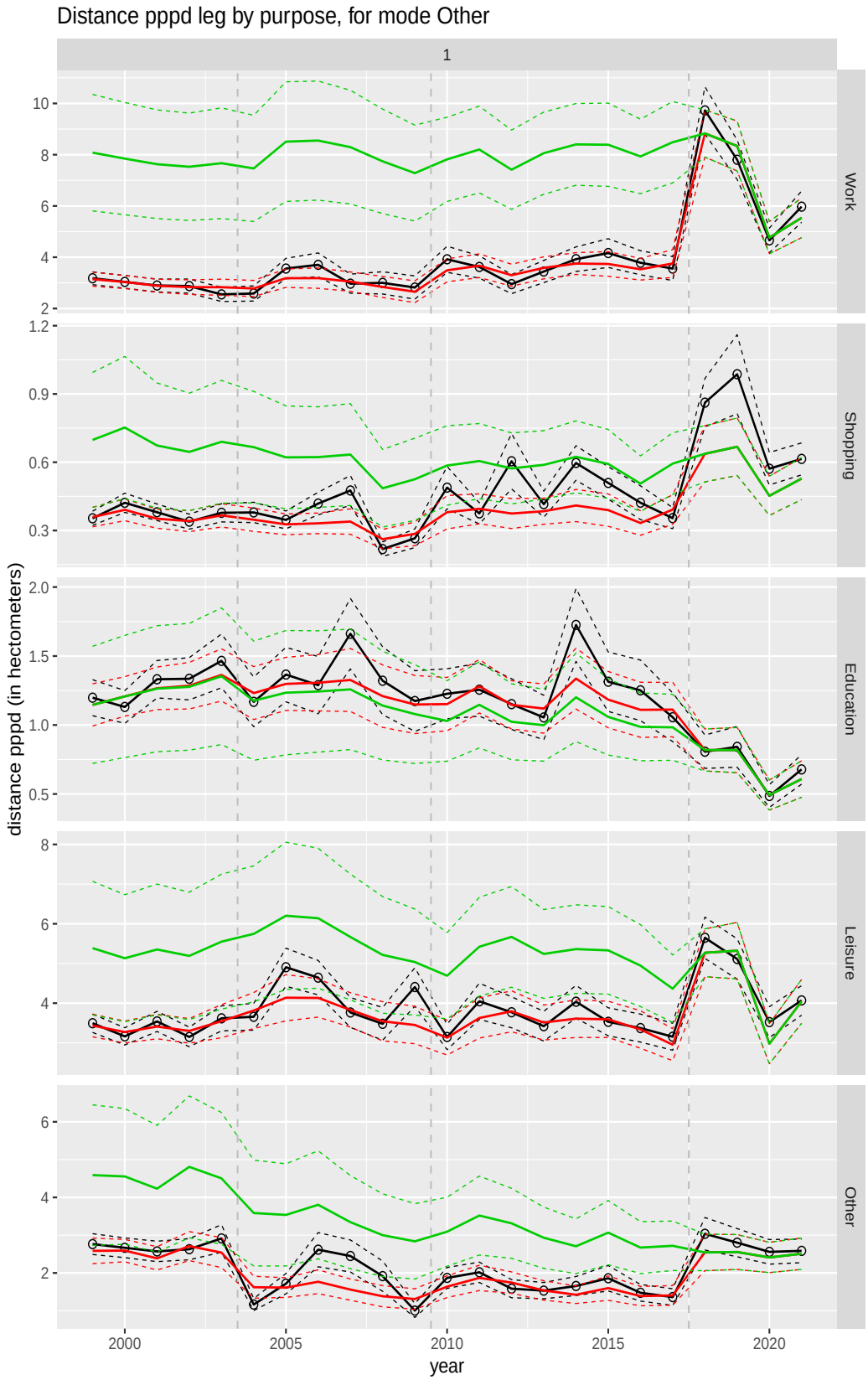

Figure A.179 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

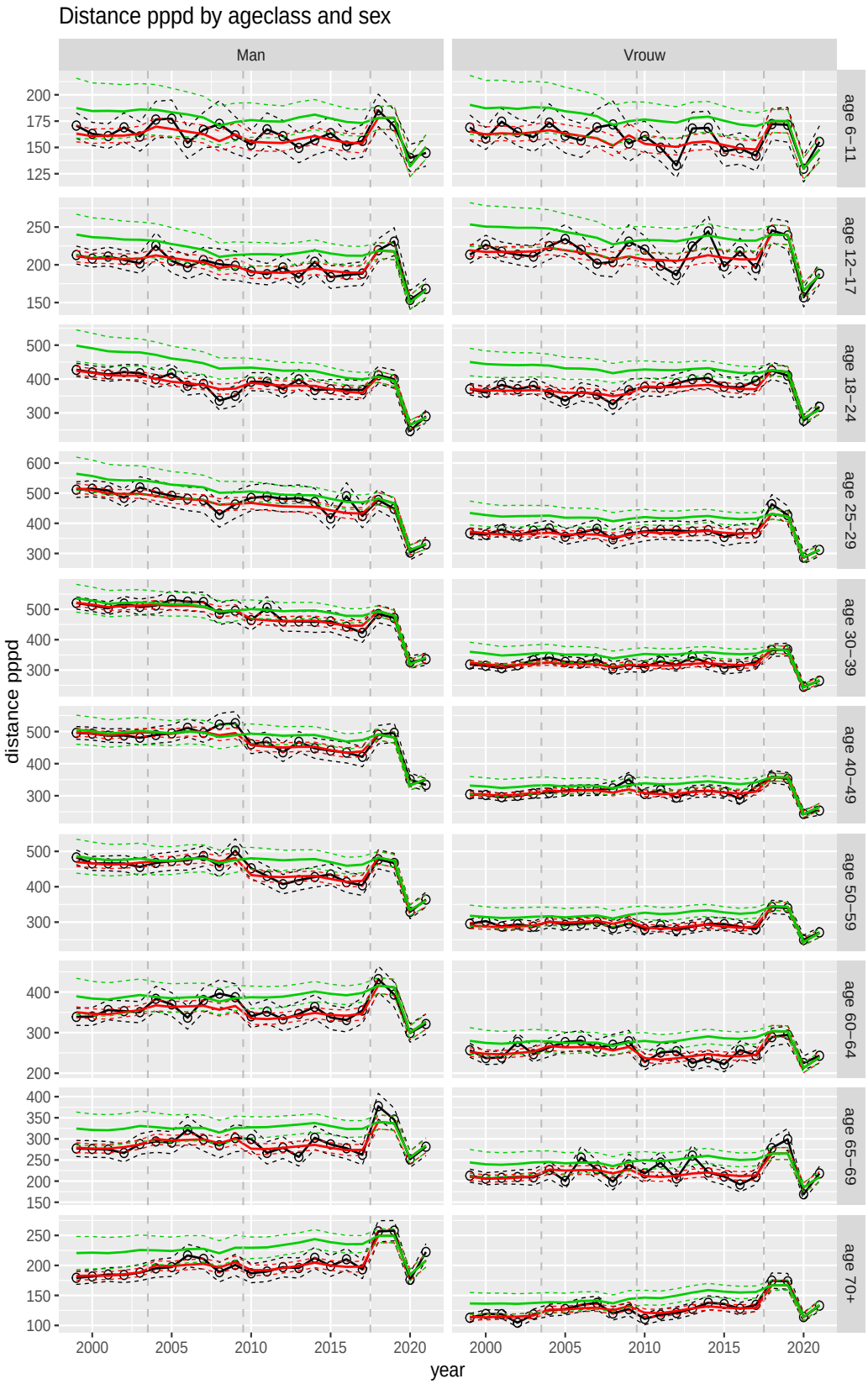

Figure A.180 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

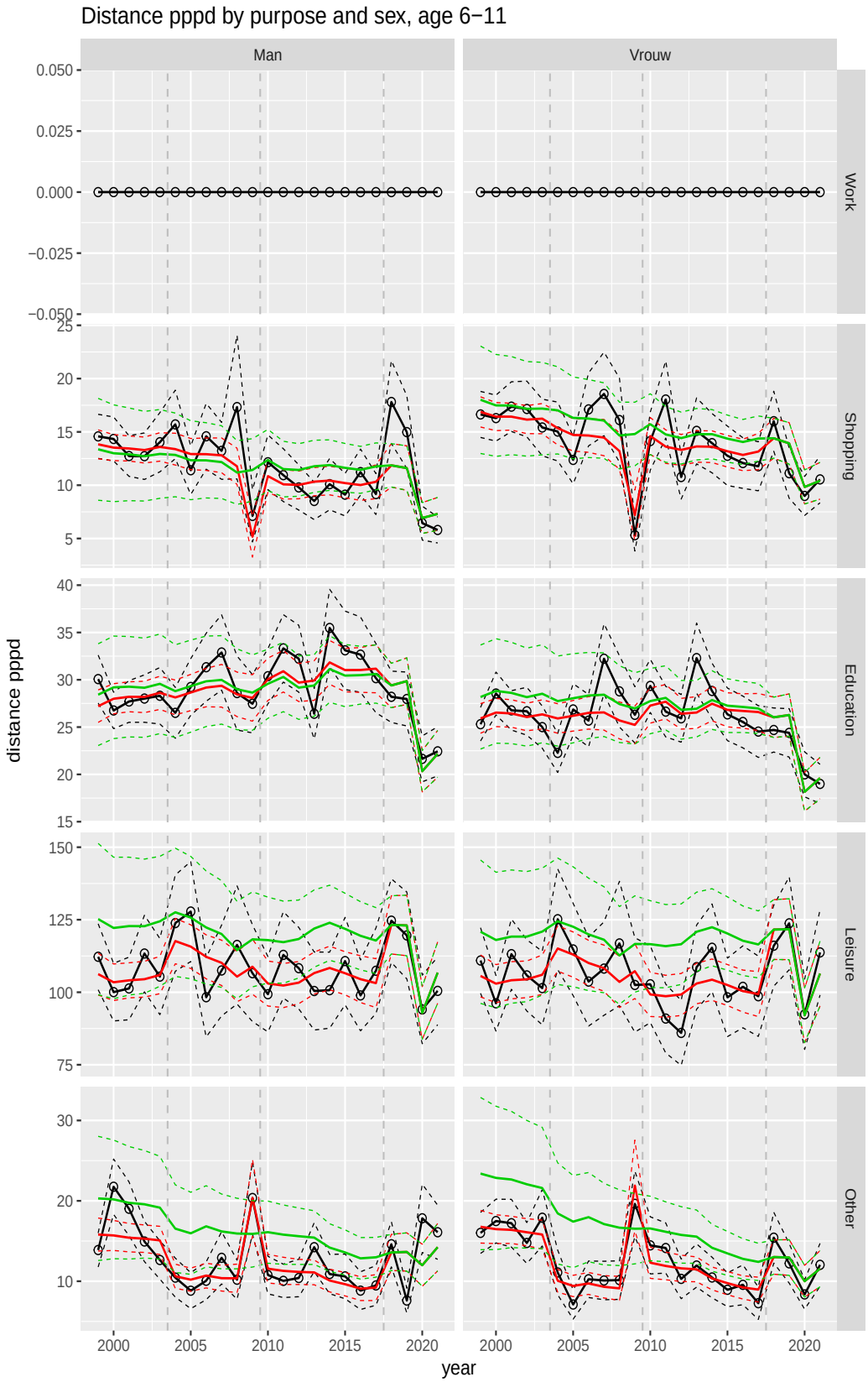

Figure A.181 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

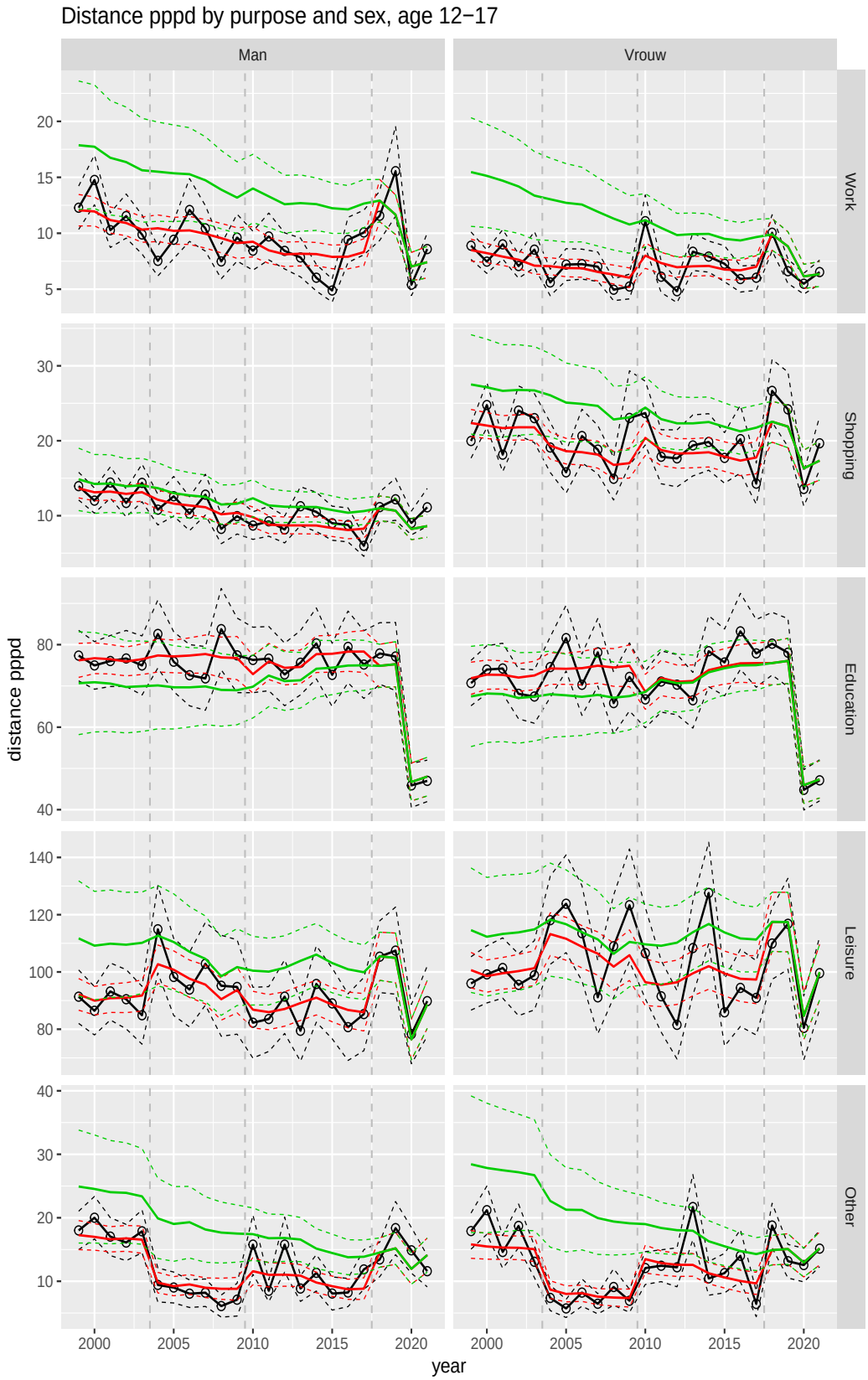

Figure A.182 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

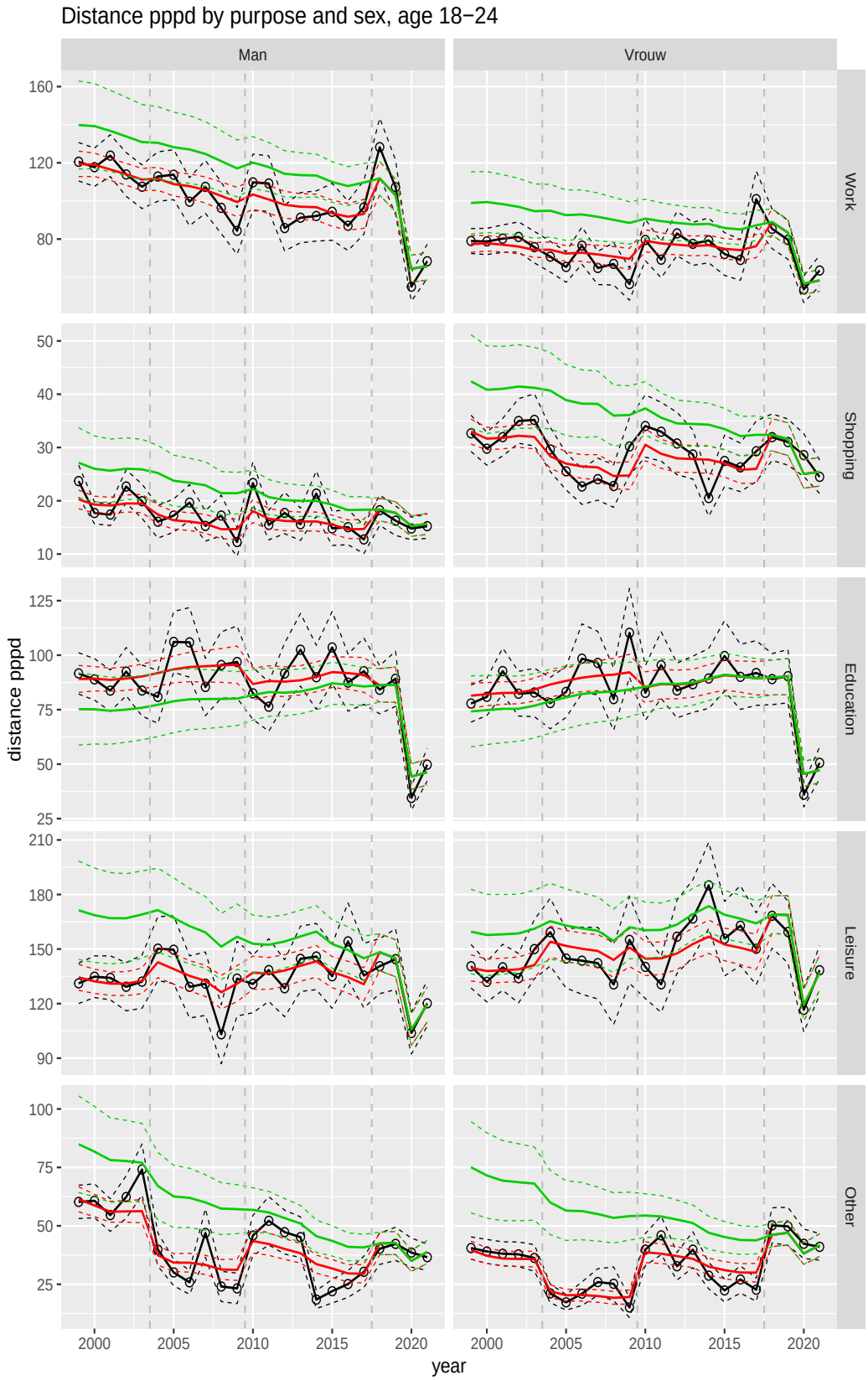

Figure A.183 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.184 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

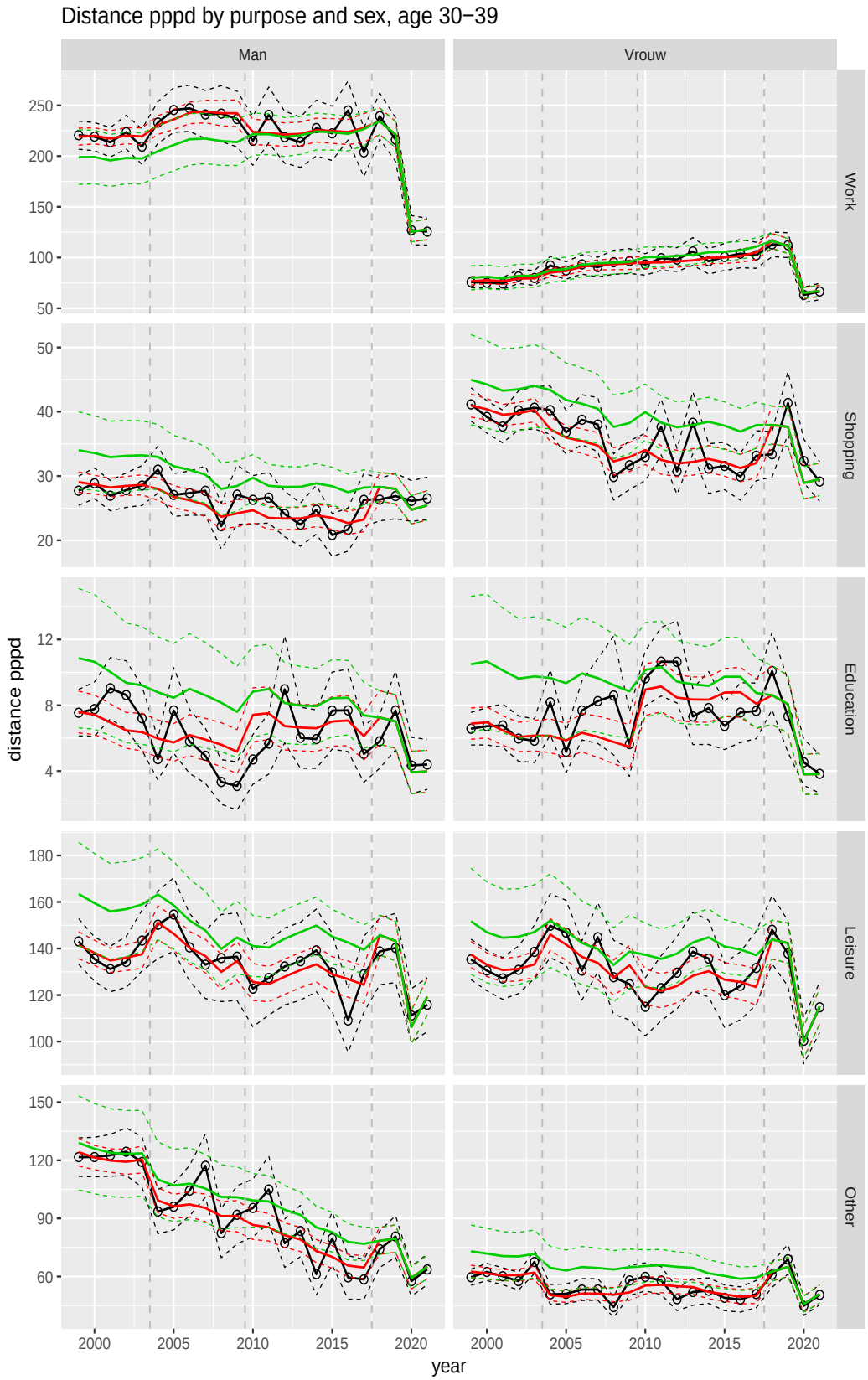

Figure A.185 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.186 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.187 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

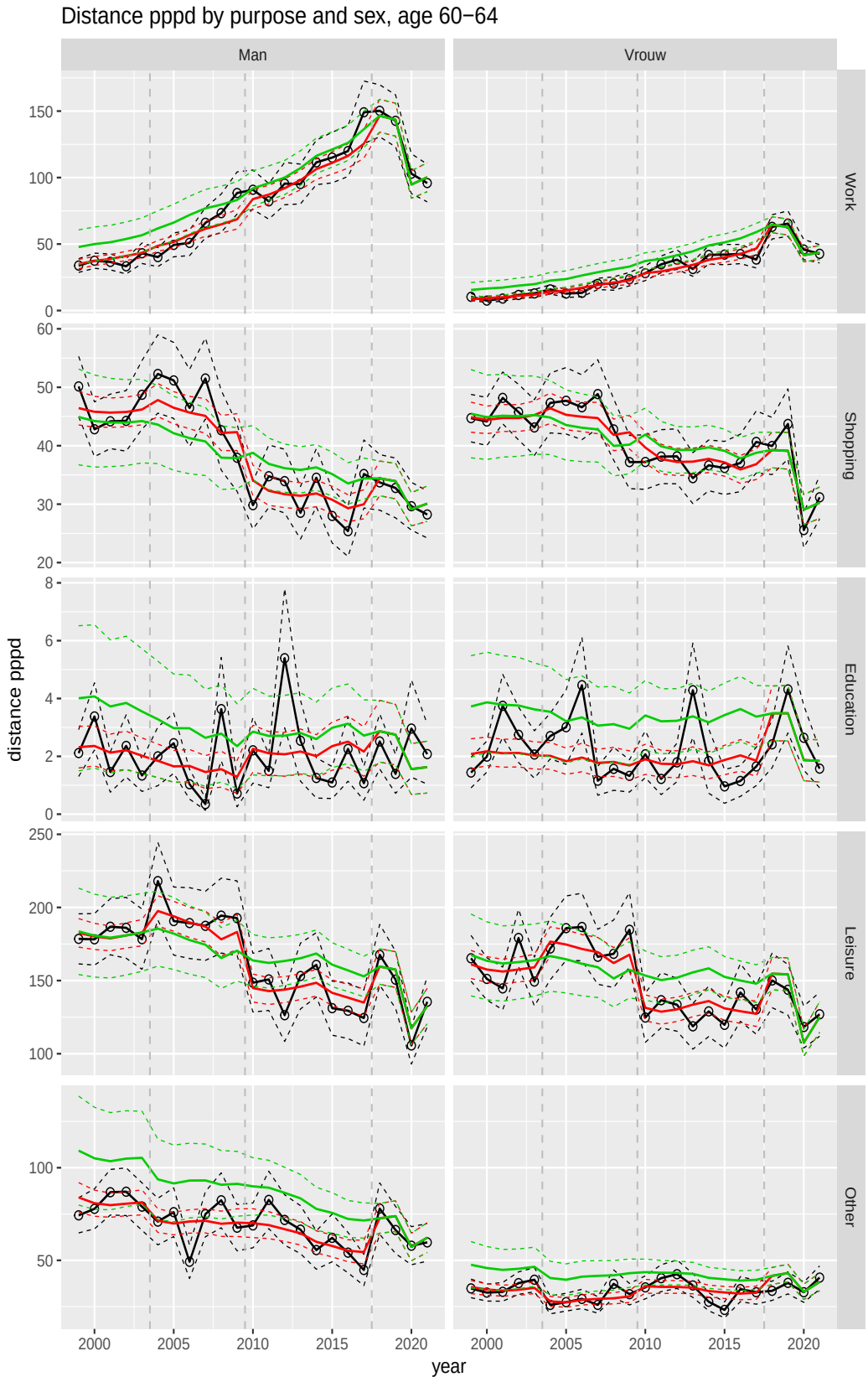

Figure A.188 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

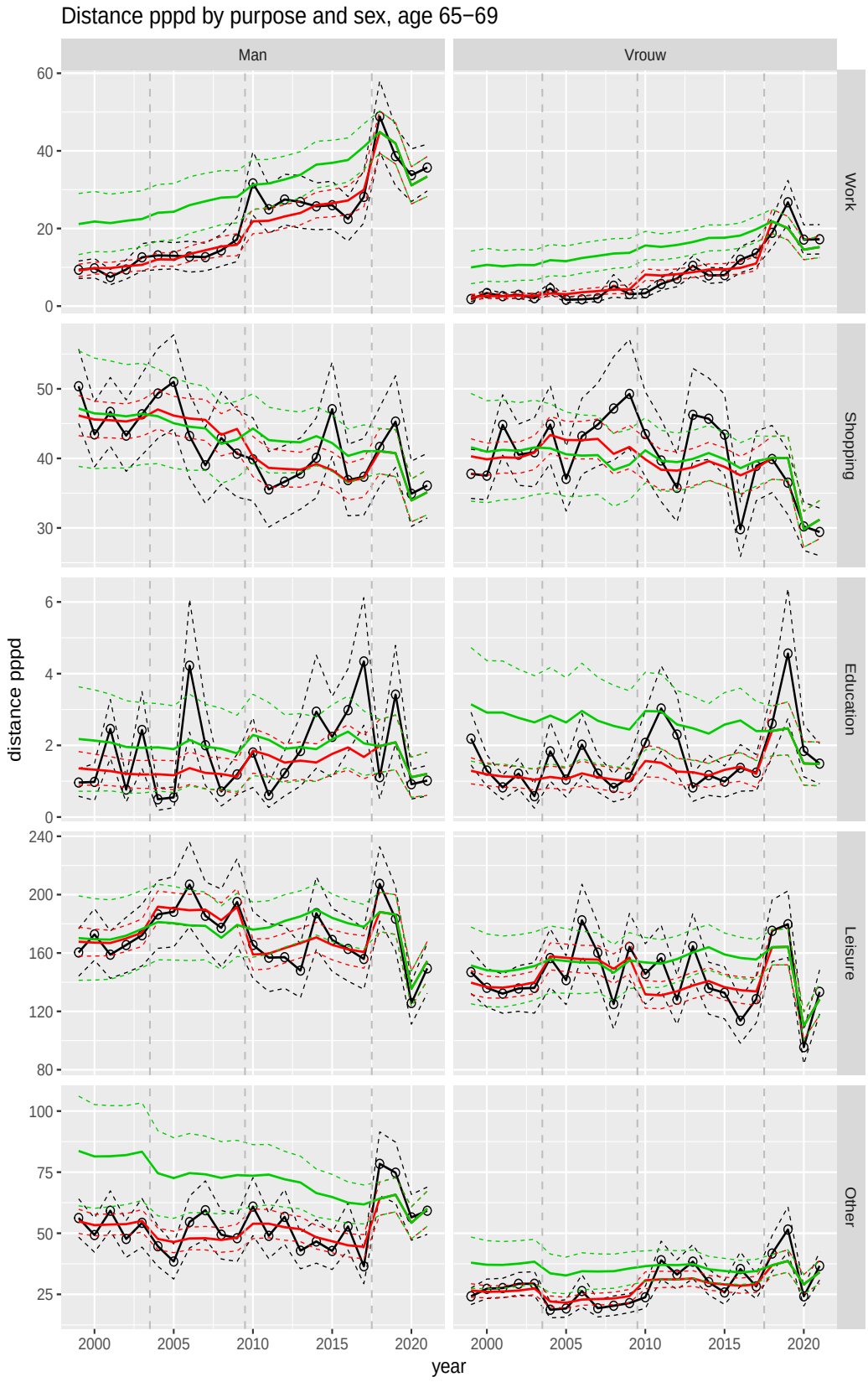

Figure A.189 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.190 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

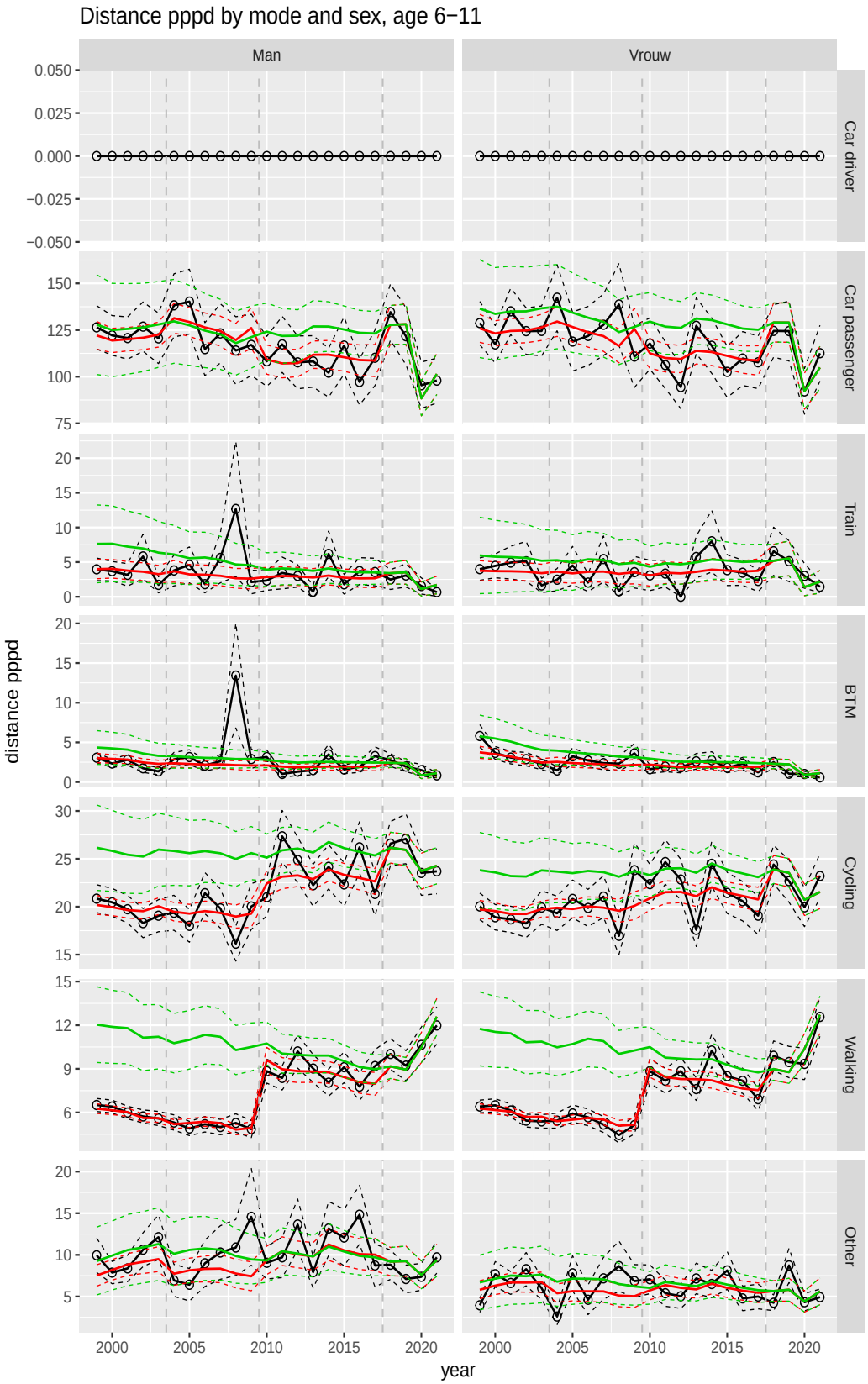

Figure A.191 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

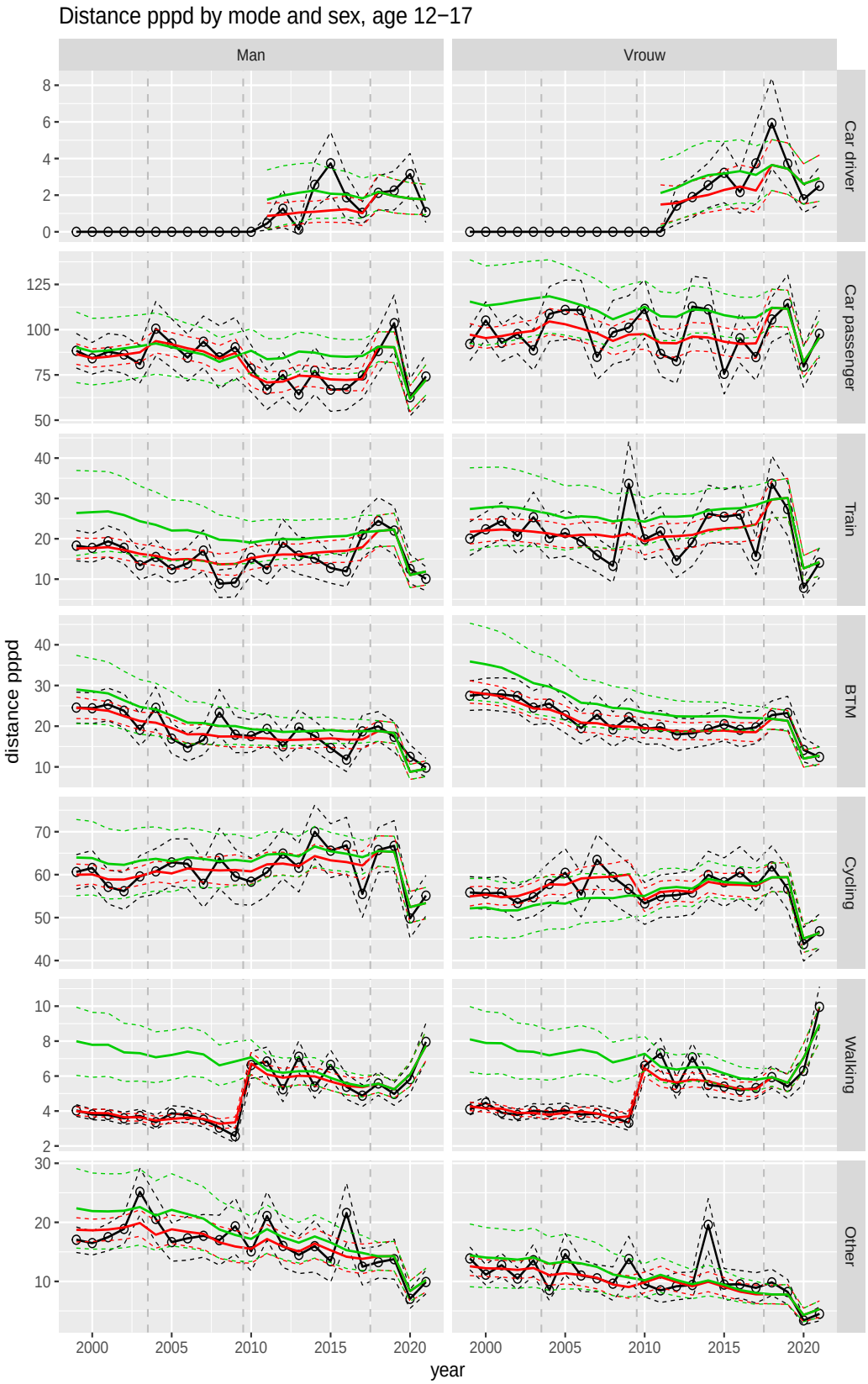

Figure A.192 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.193 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

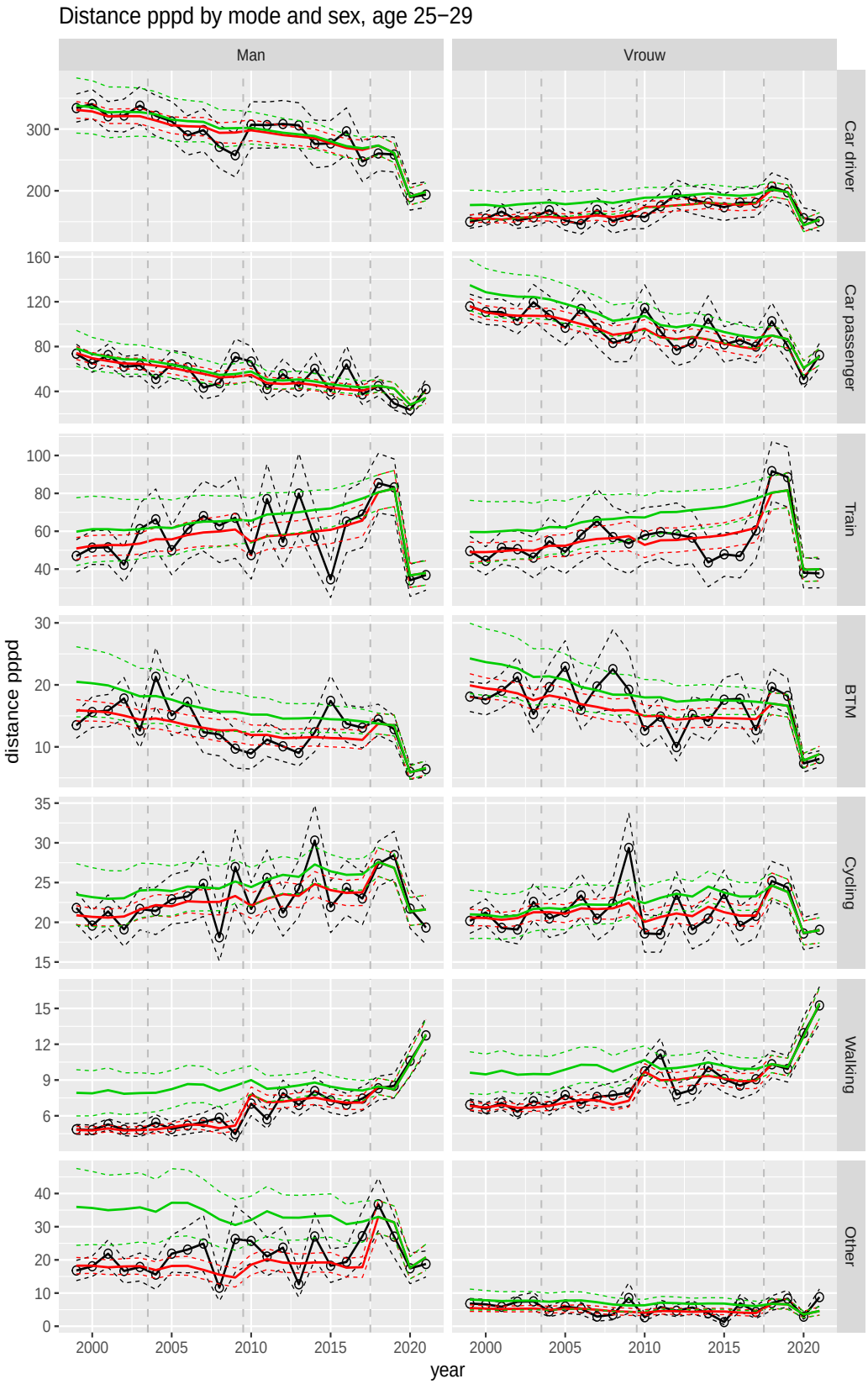

Figure A.194 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

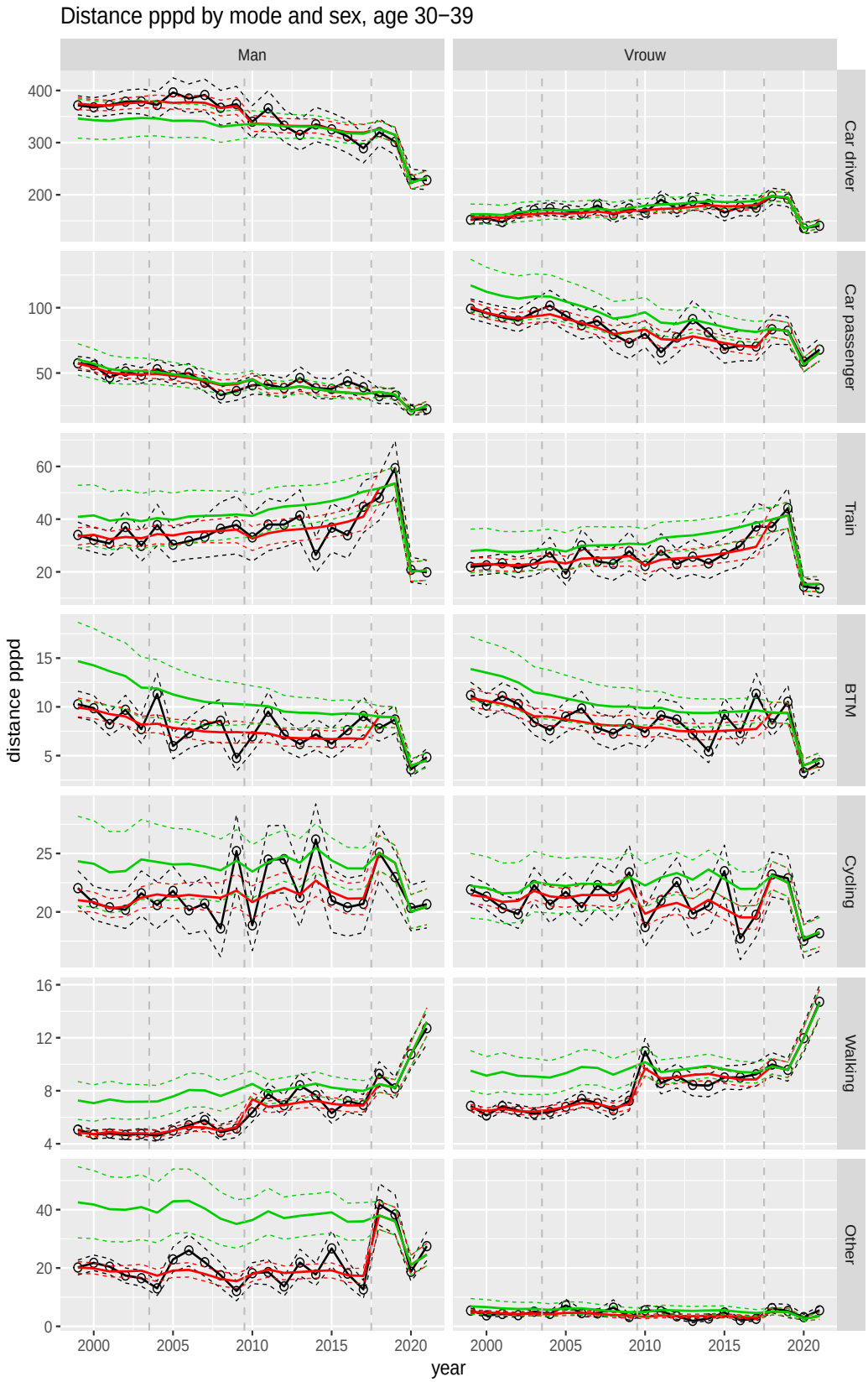

Figure A.195 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

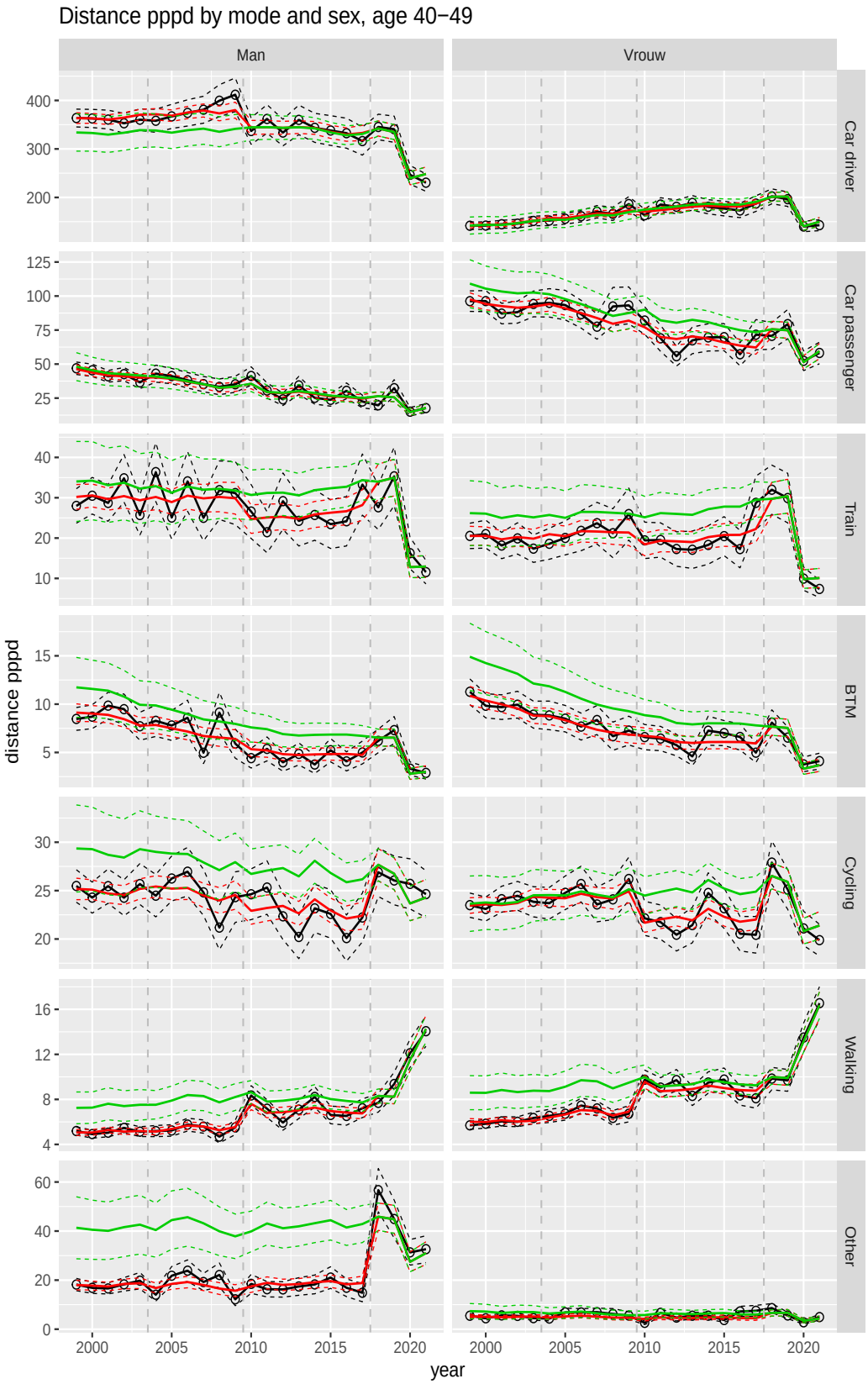

Figure A.196 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

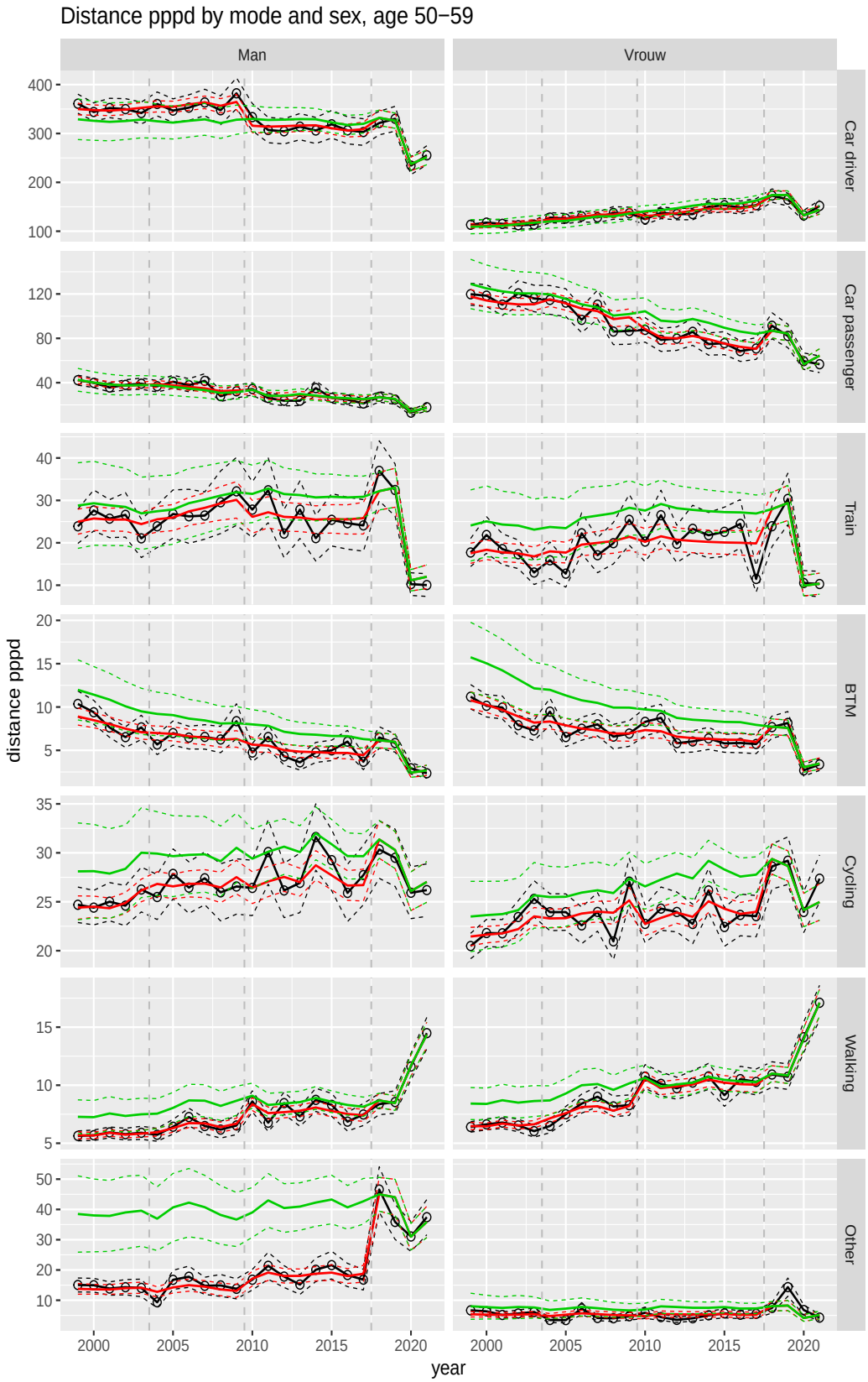

Figure A.197 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

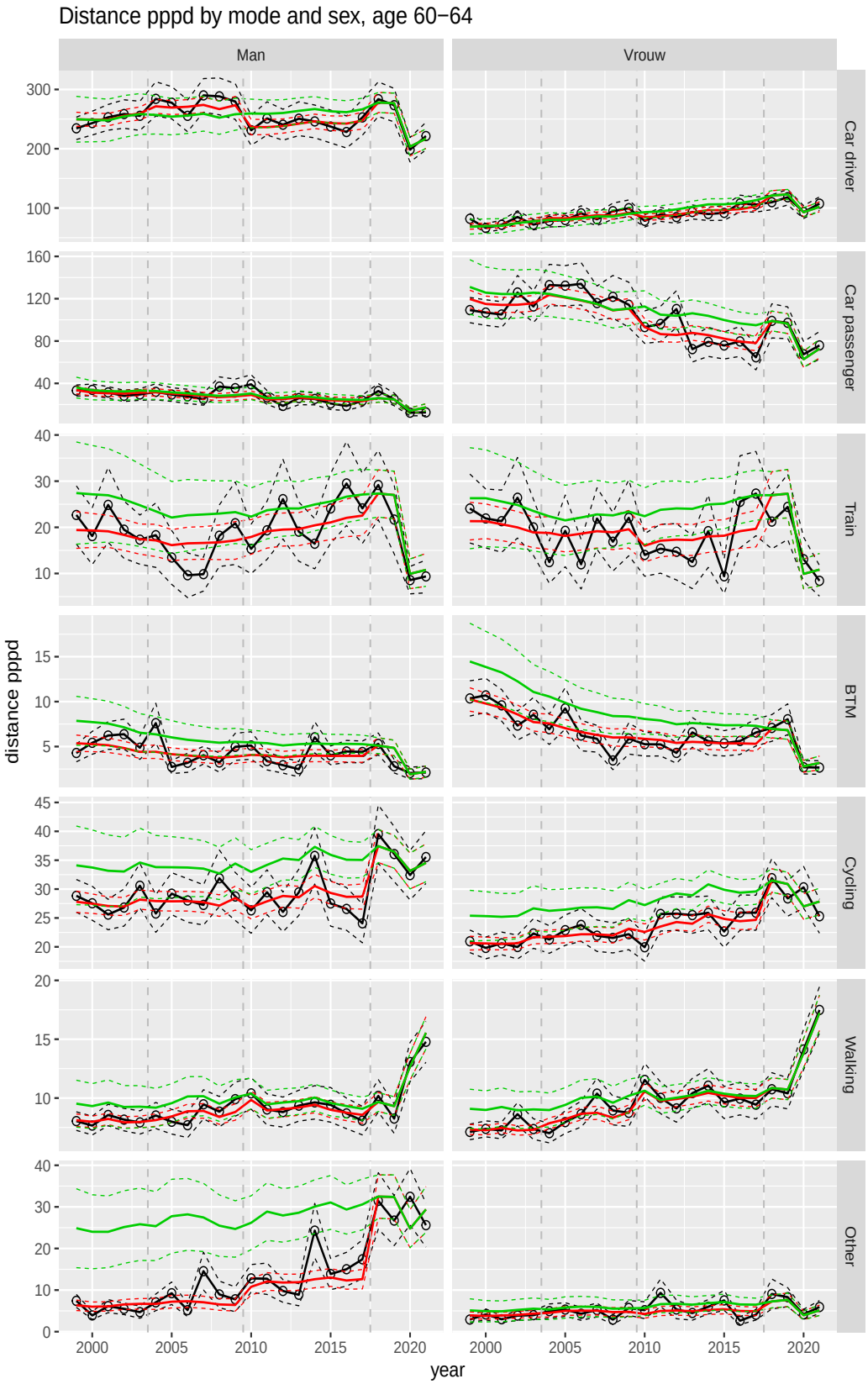

Figure A.198 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

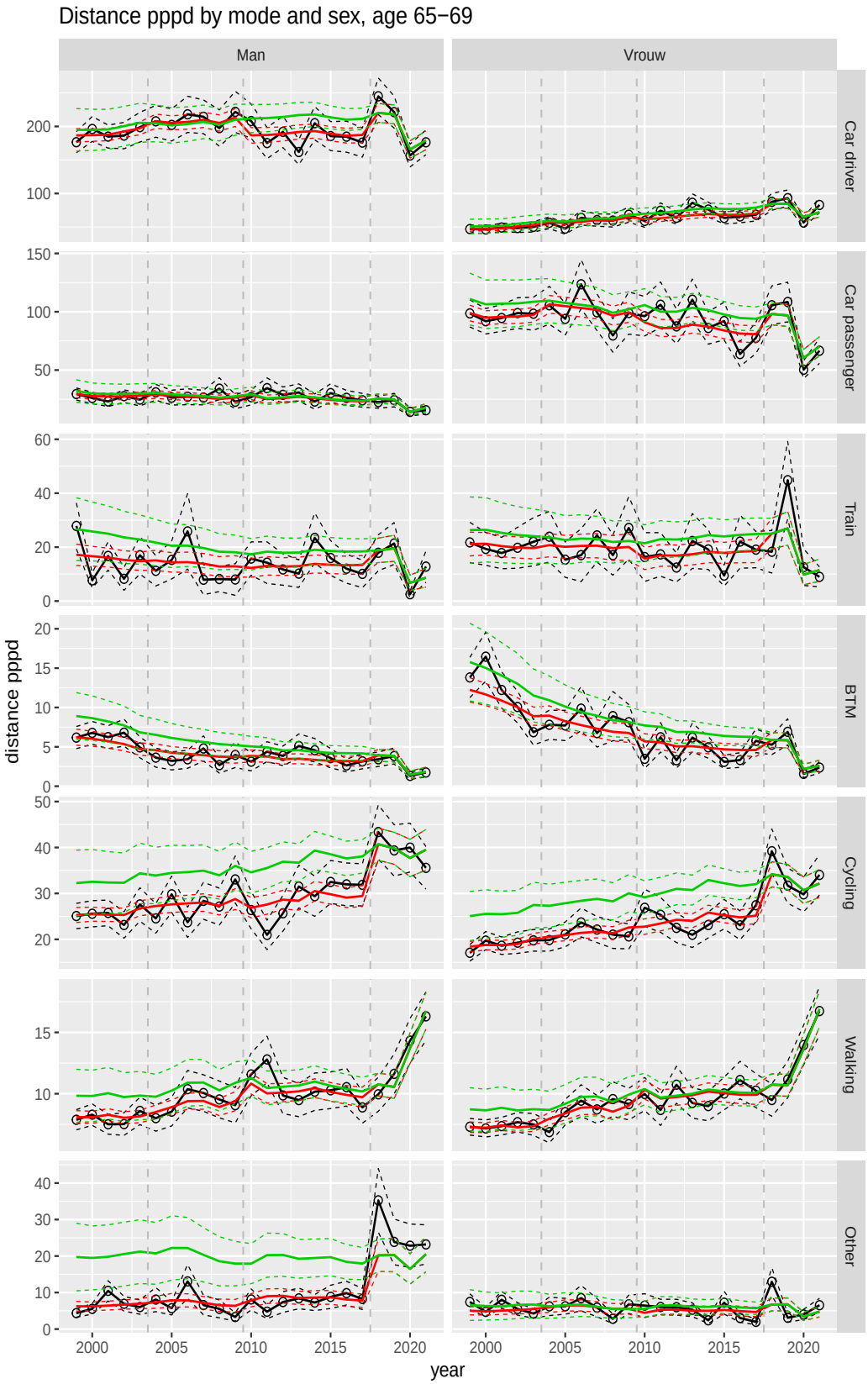

Figure A.199 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

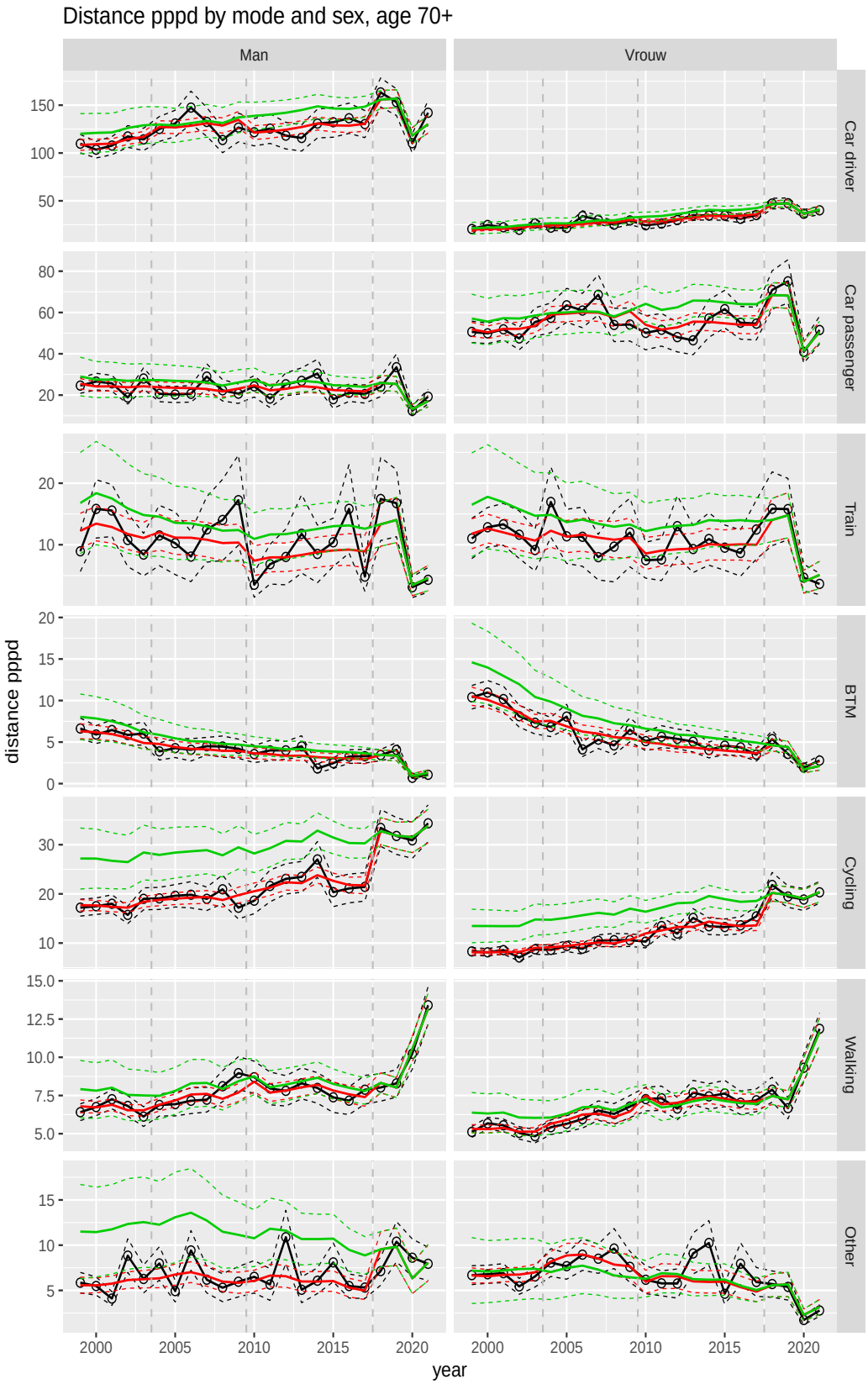

Figure A.200 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

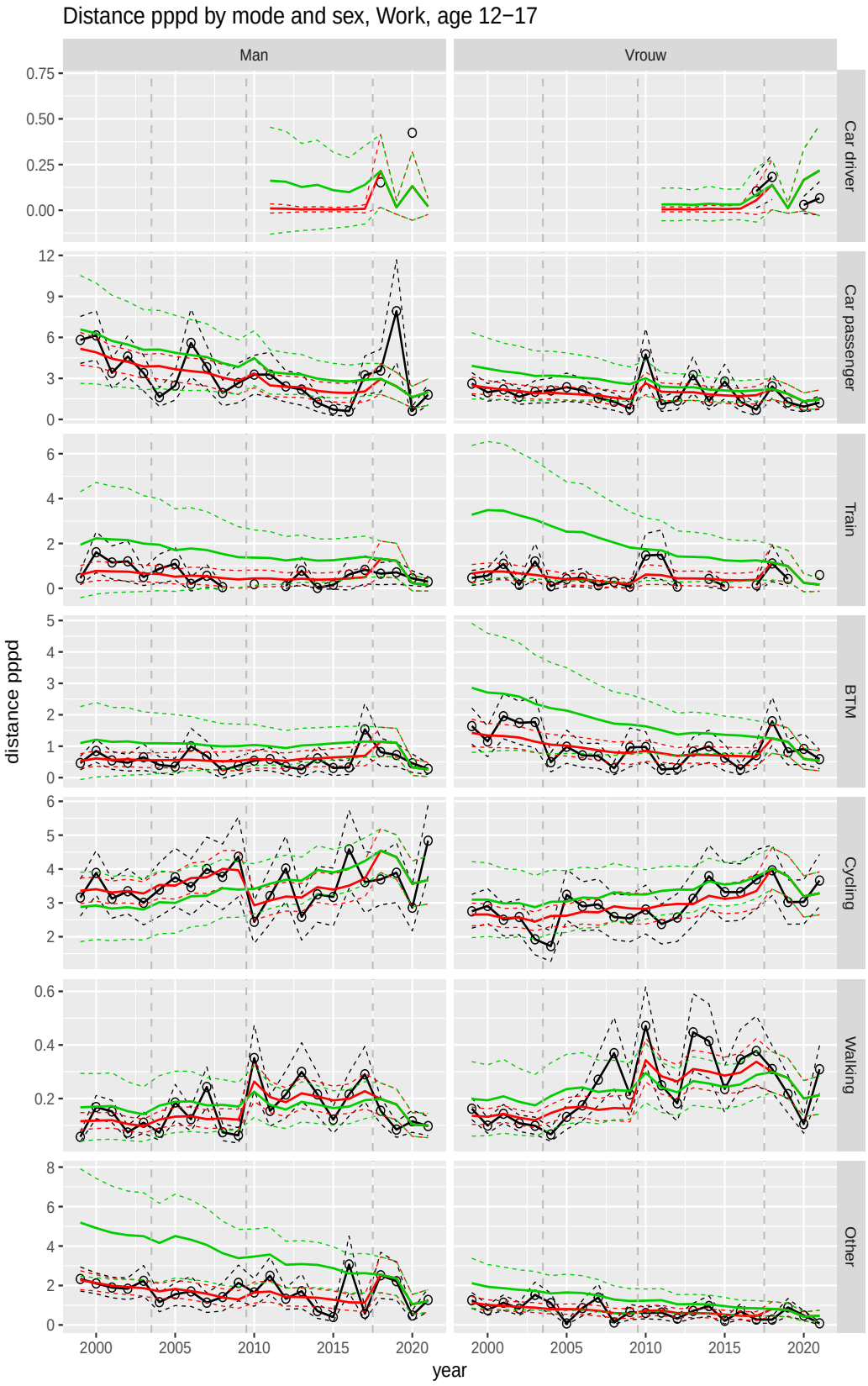

Figure A.201 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

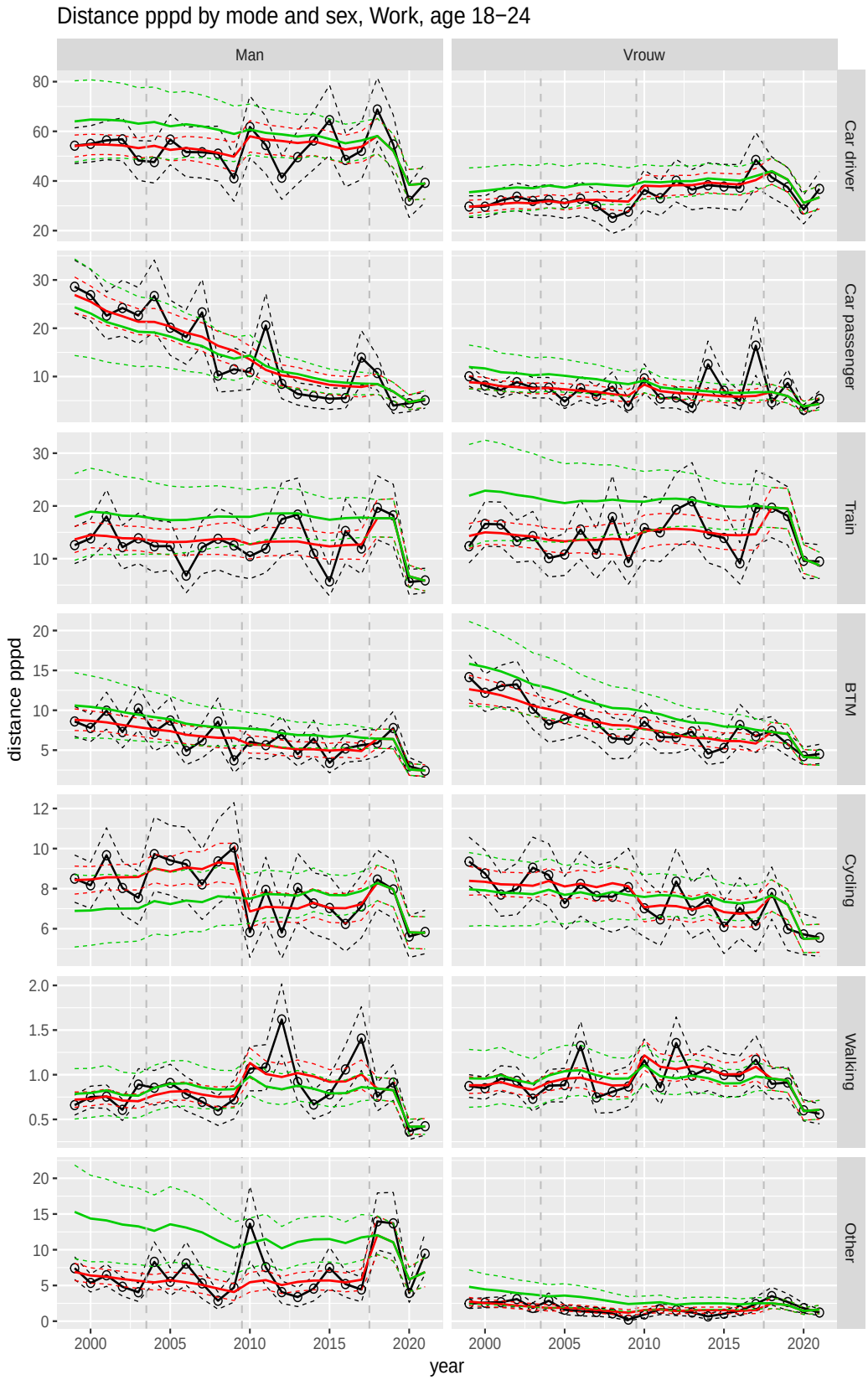

Figure A.202 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

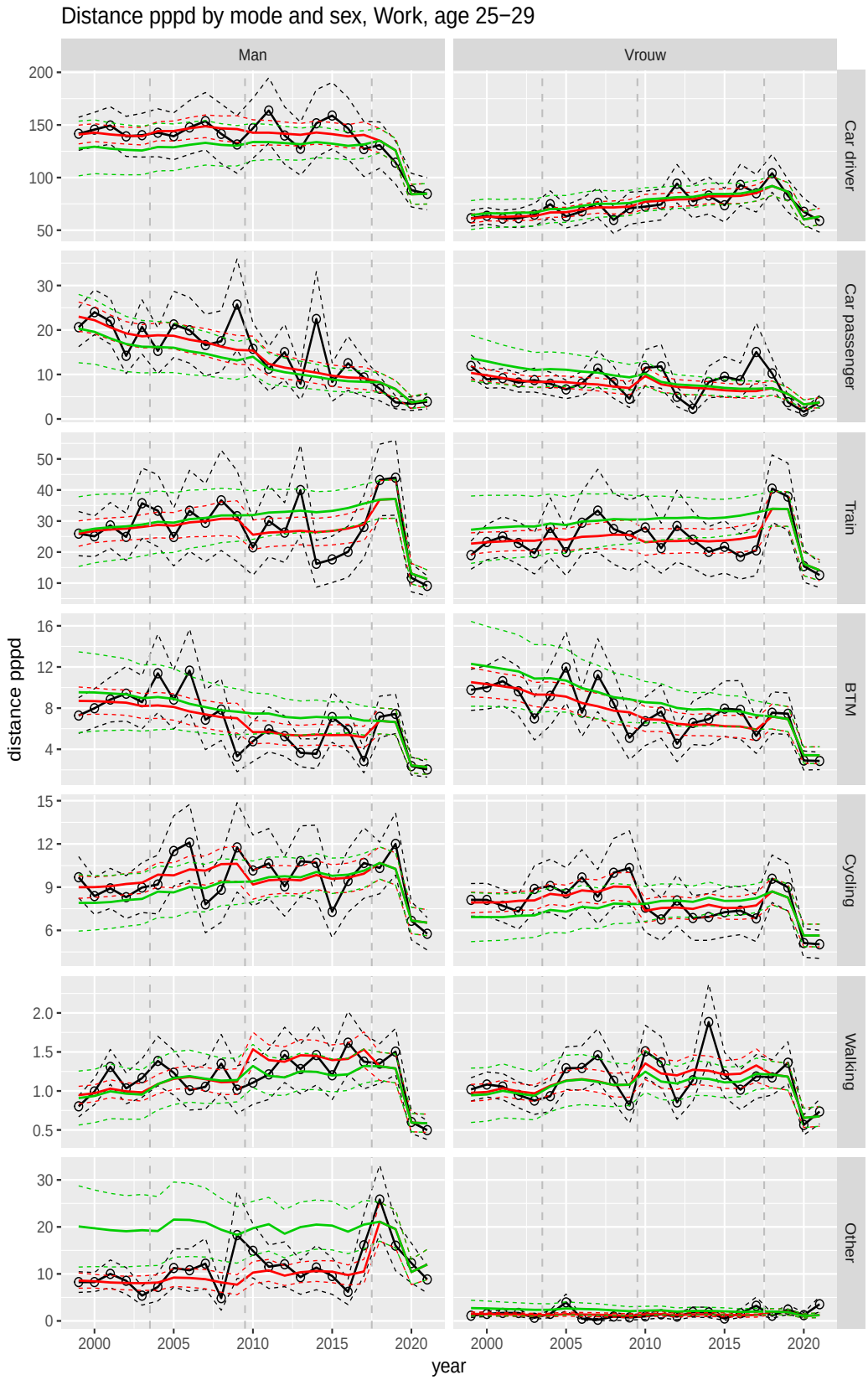

Figure A.203 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

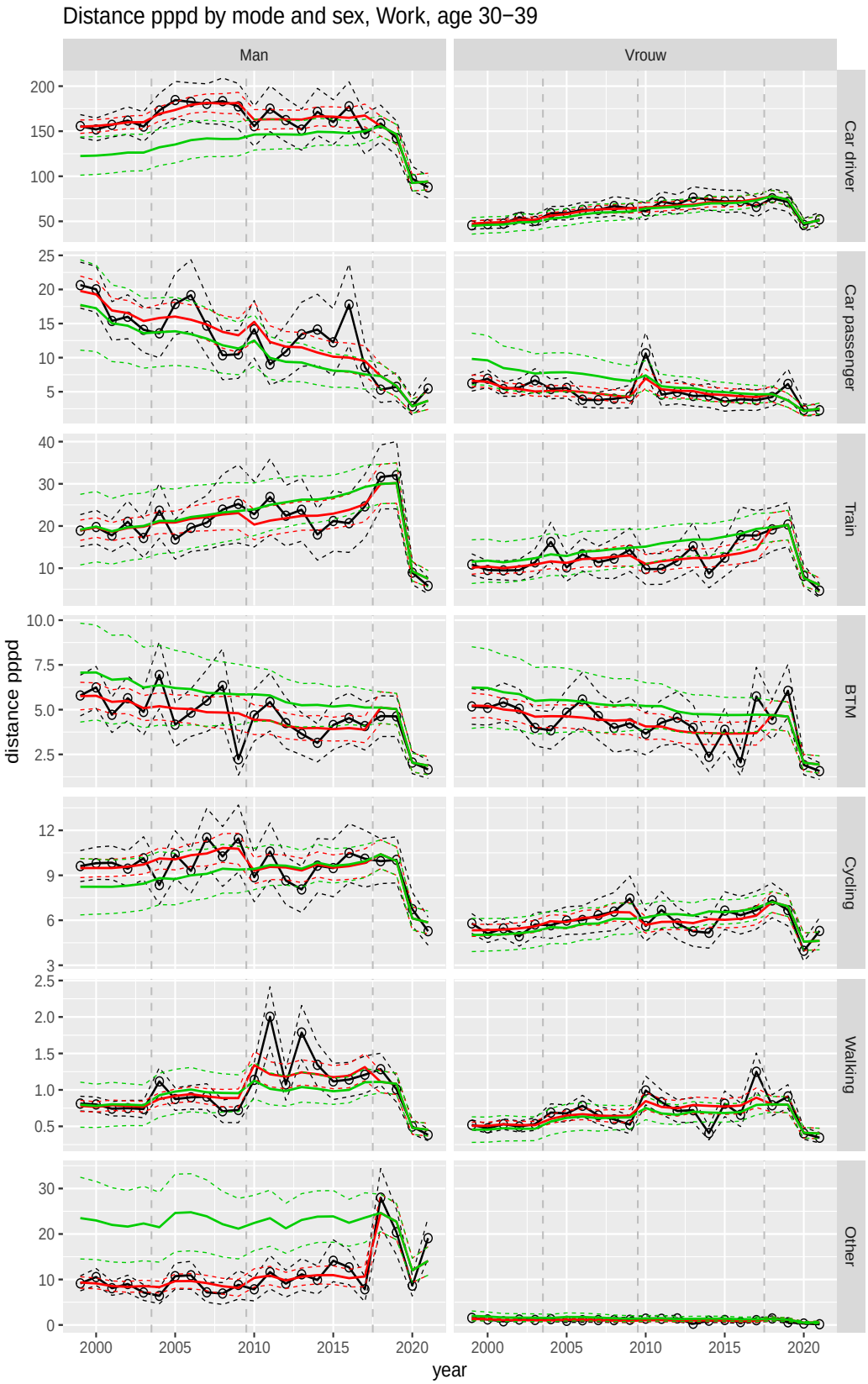

Figure A.204 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

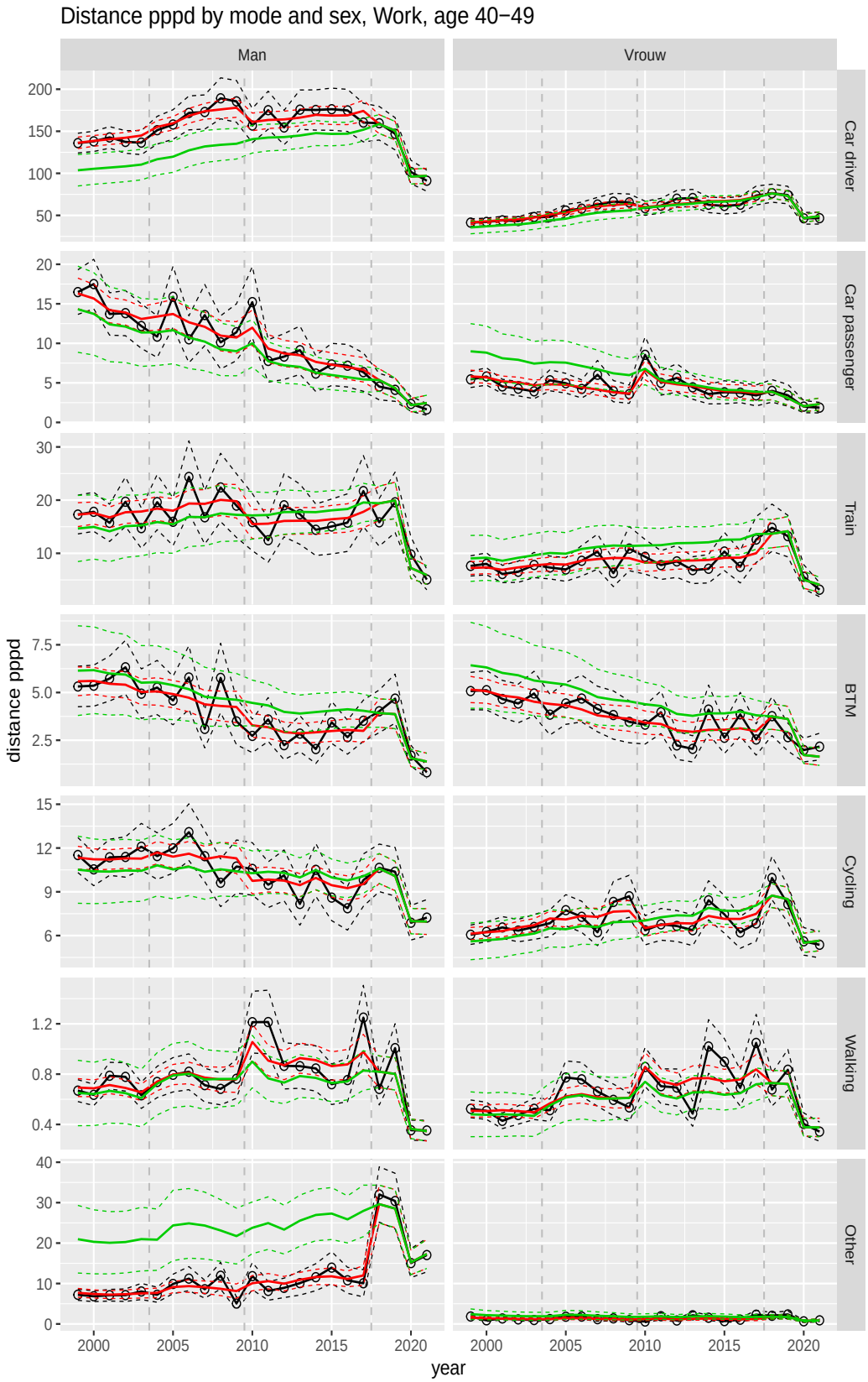

Figure A.205 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

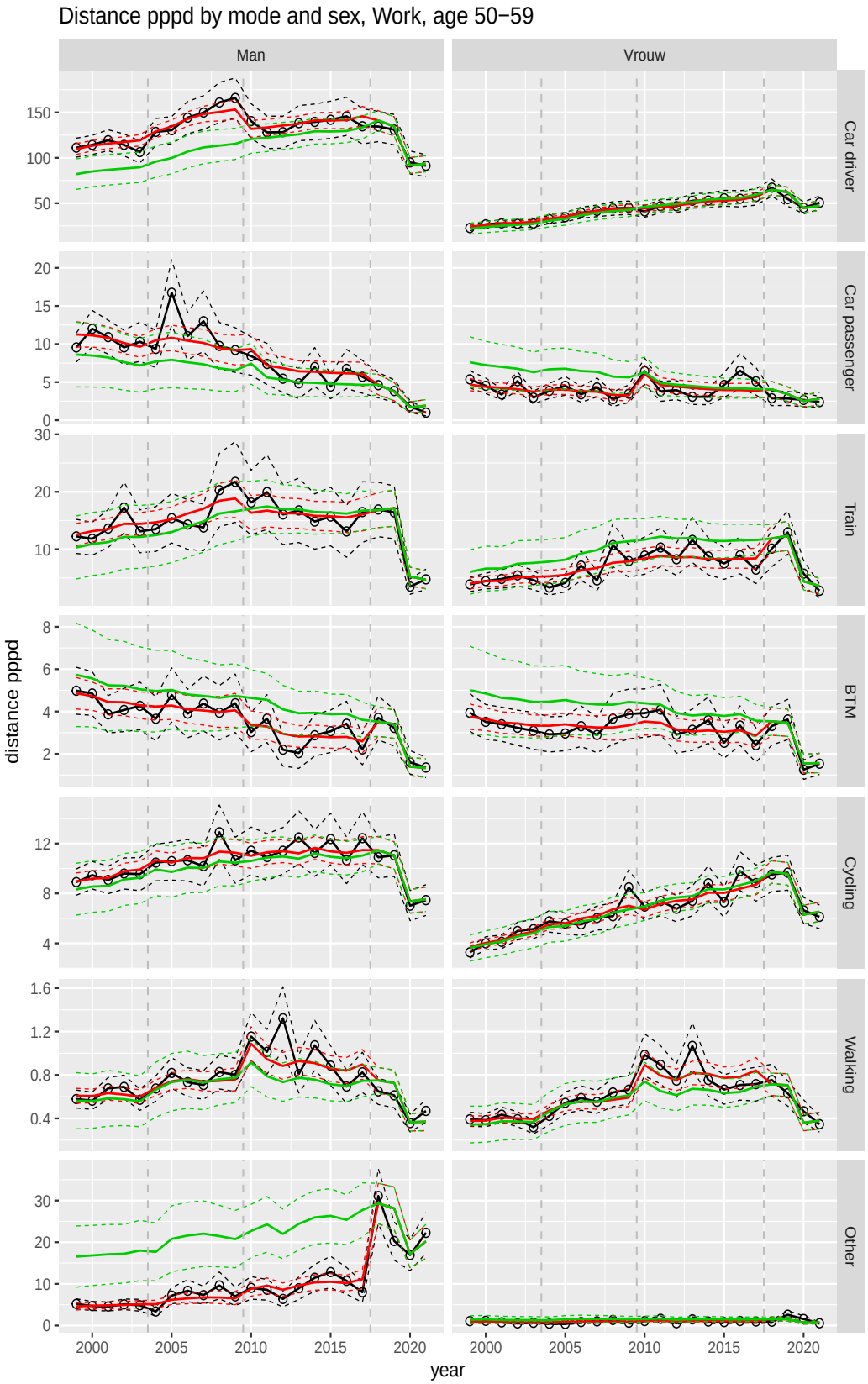

Figure A.206 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

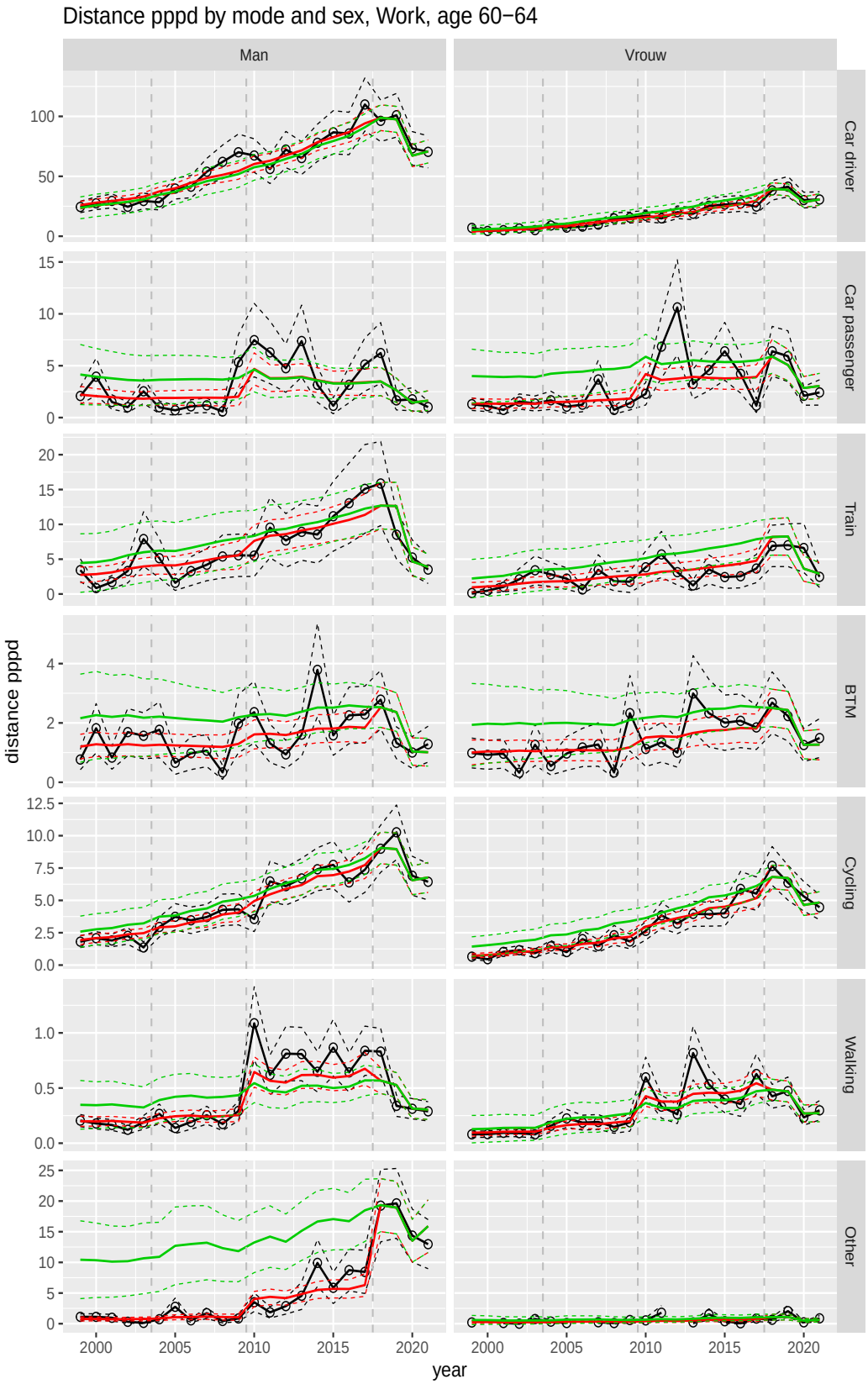

Figure A.207 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

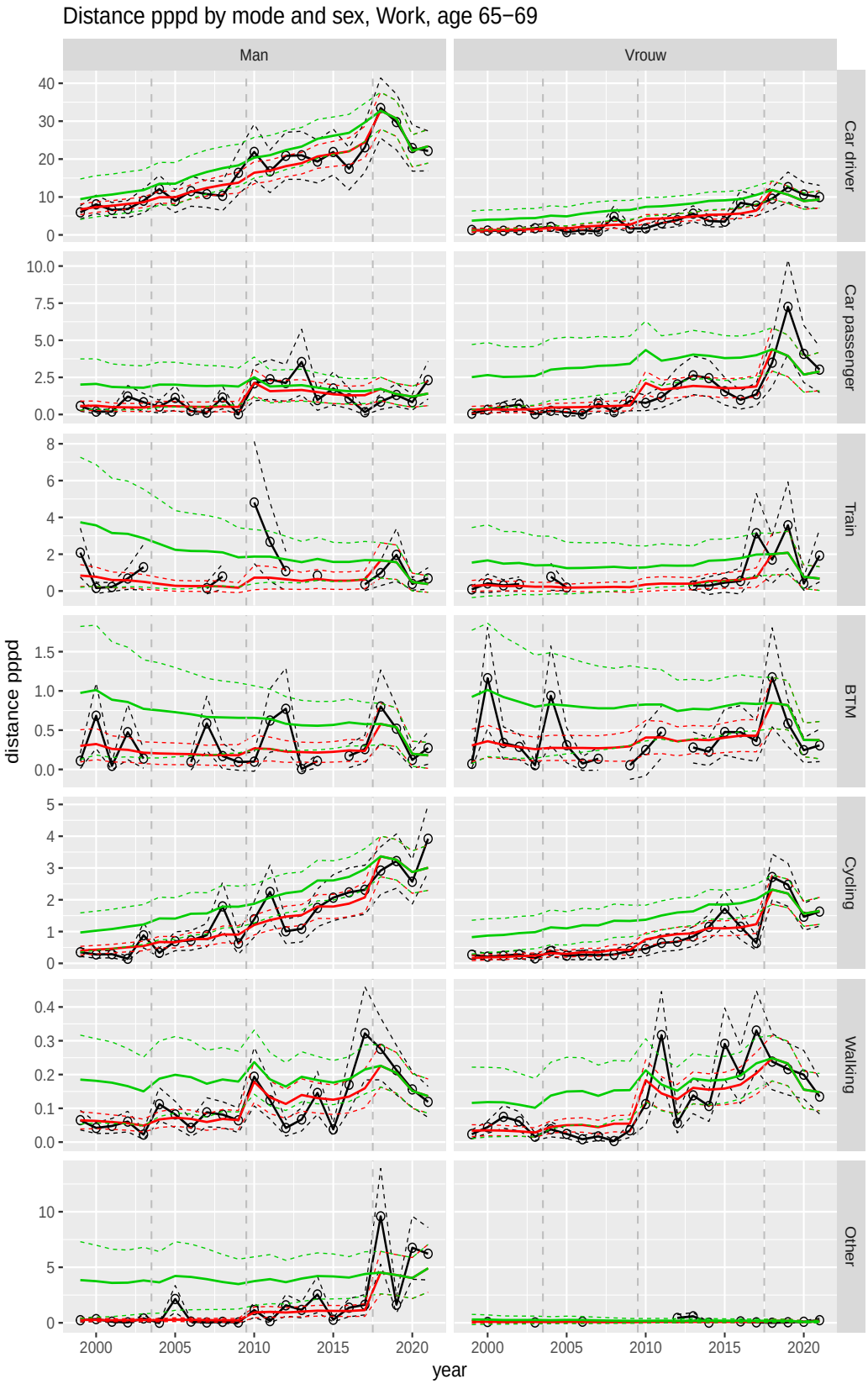

Figure A.208 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

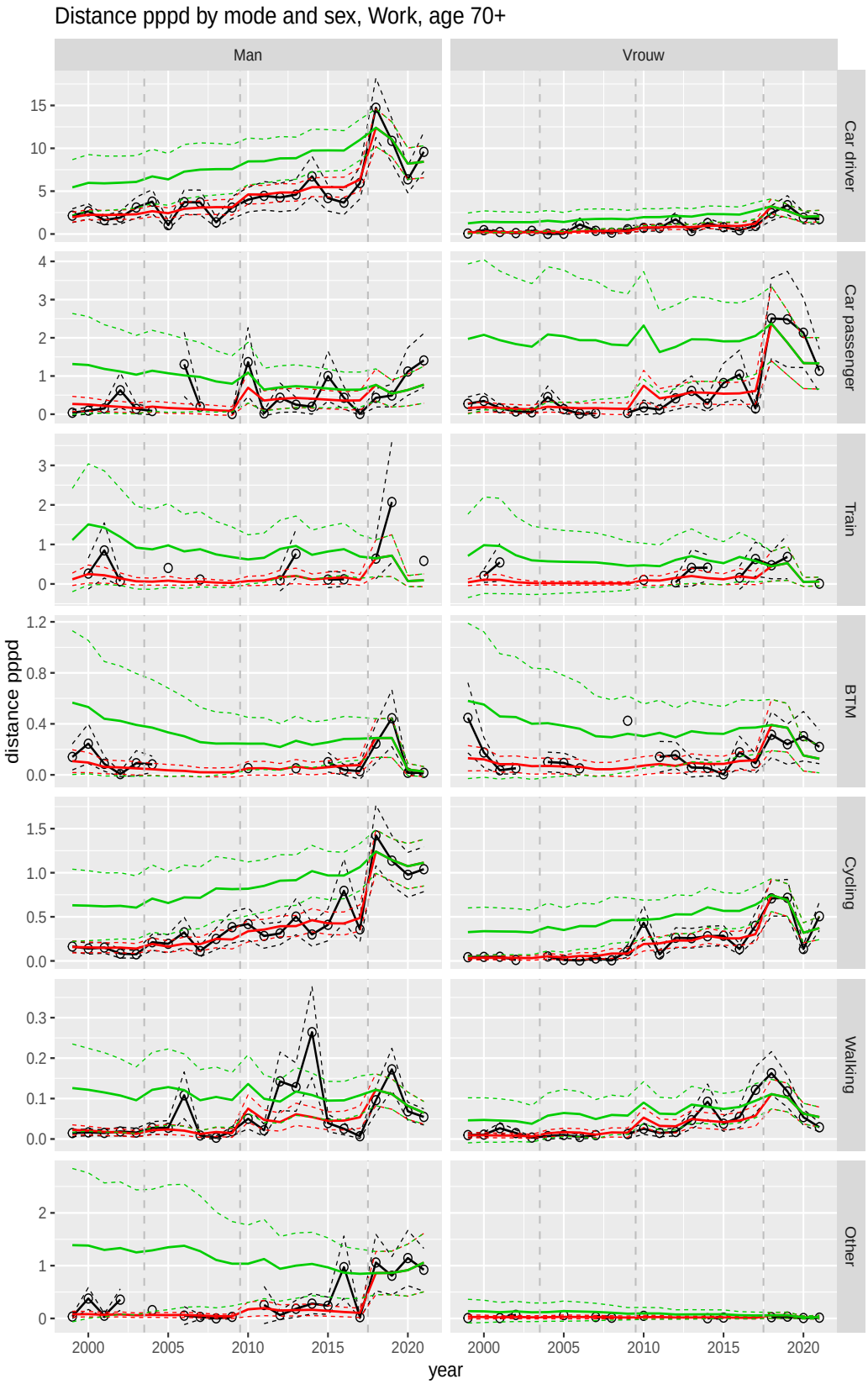

Figure A.209 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

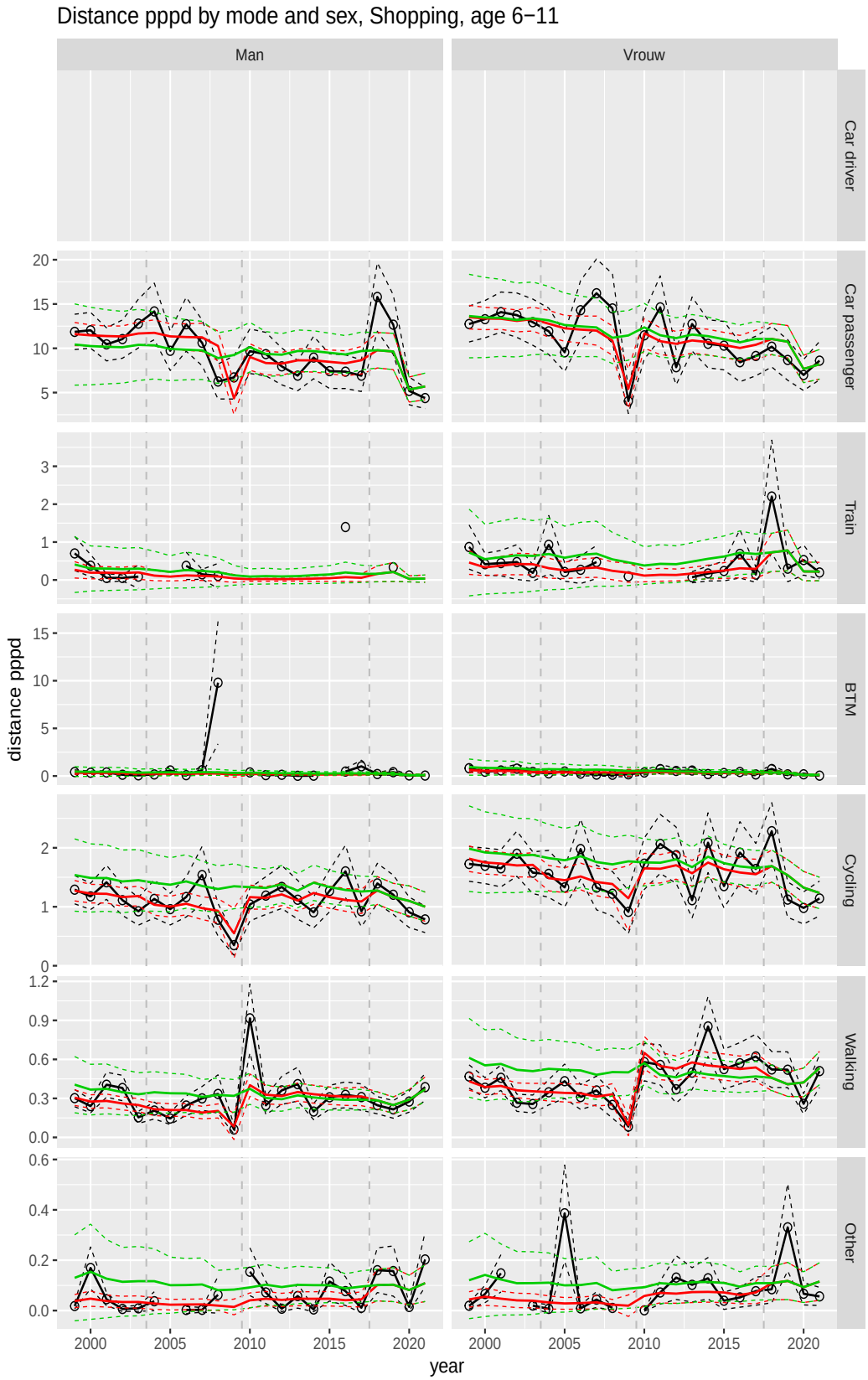

Figure A.210 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

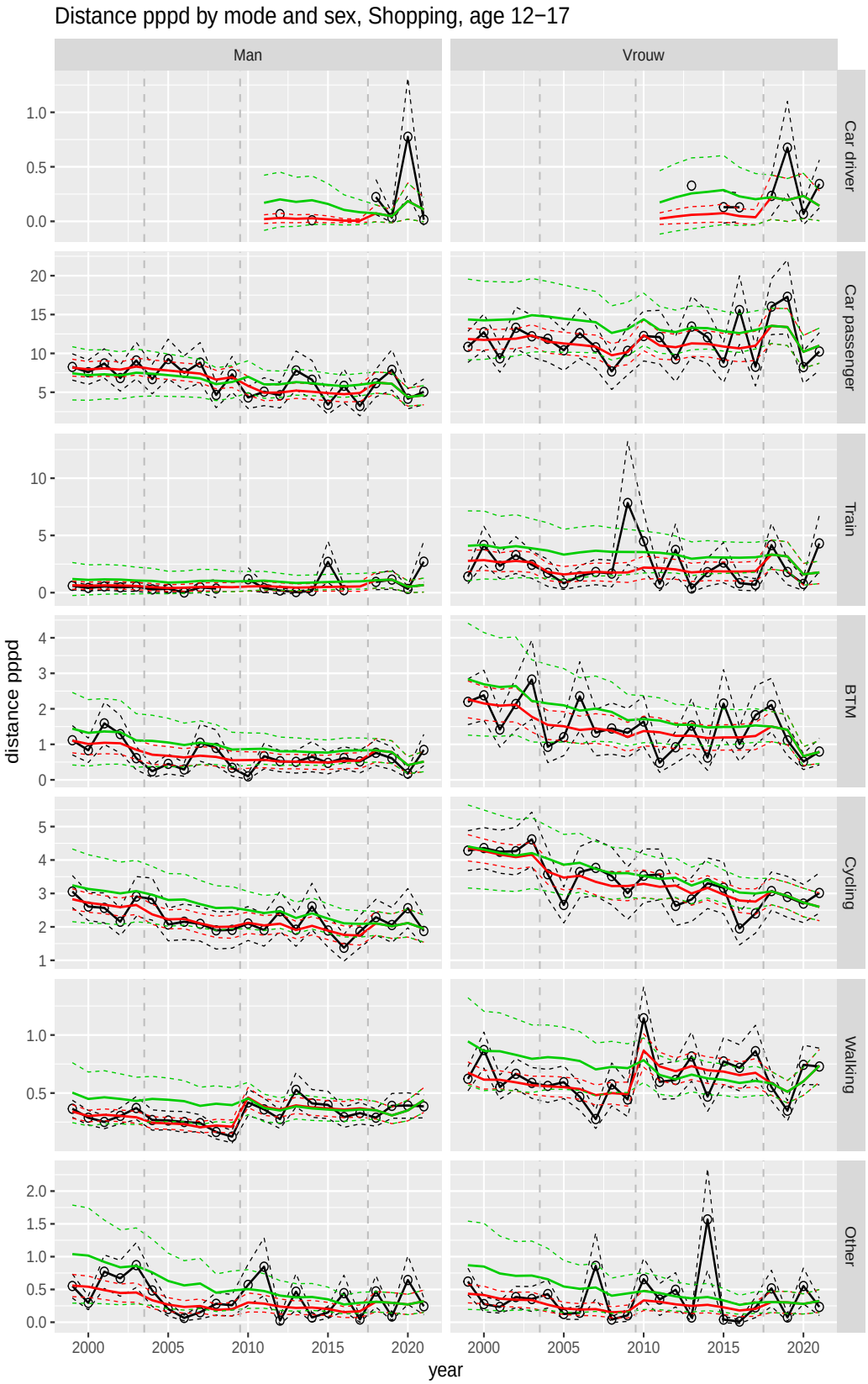

Figure A.211 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.212 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

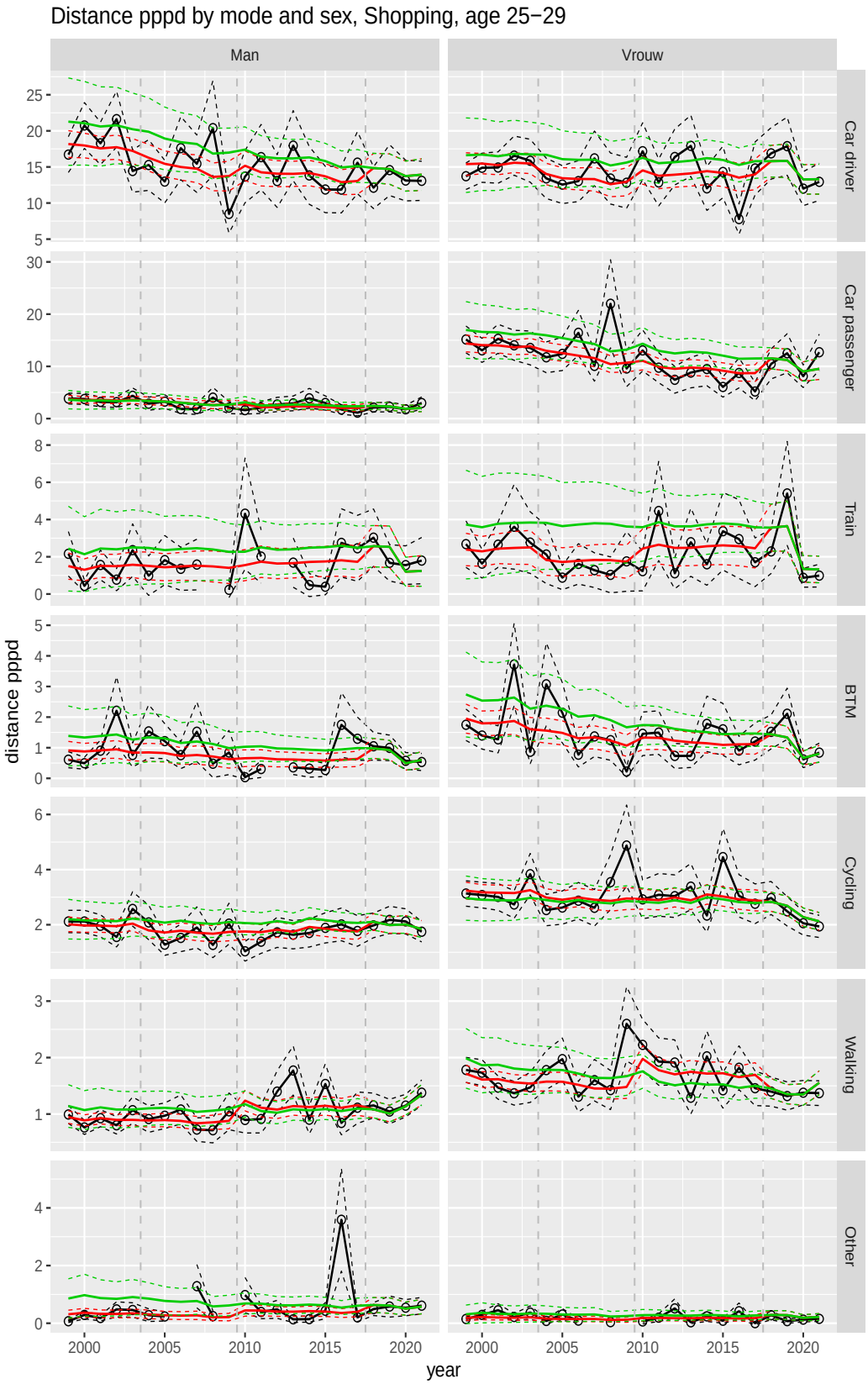

Figure A.213 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

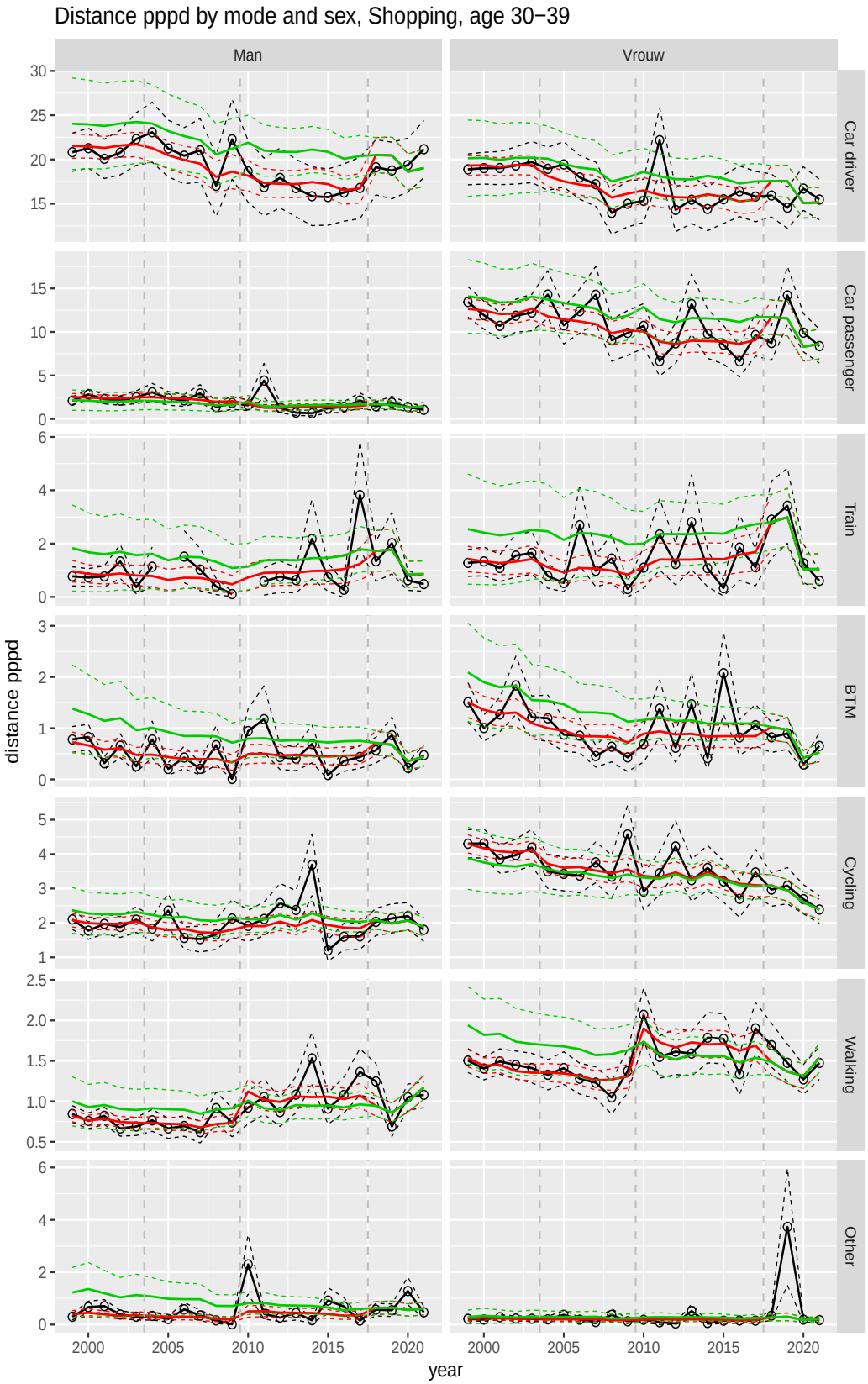

Figure A.214 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

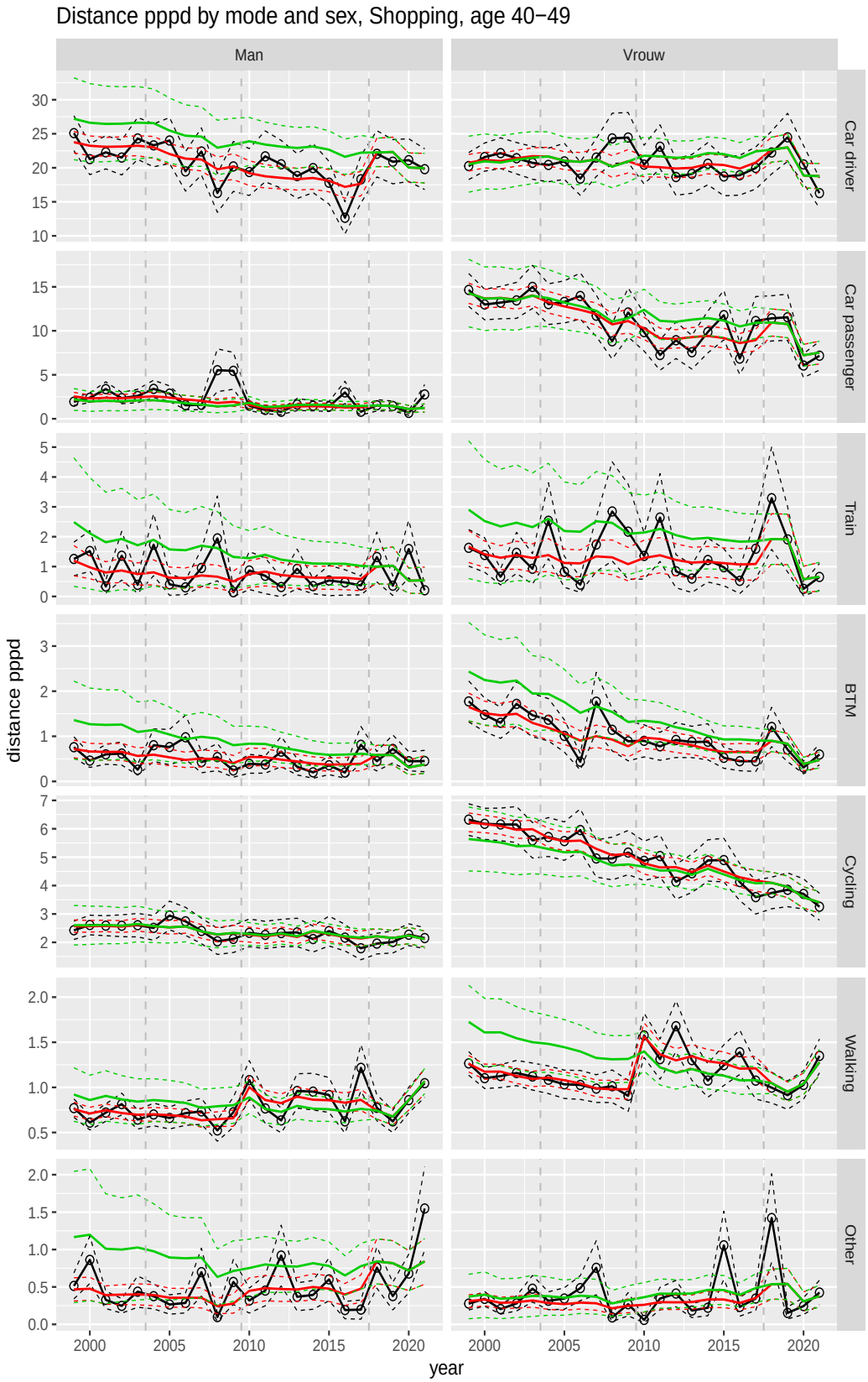

Figure A.215 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

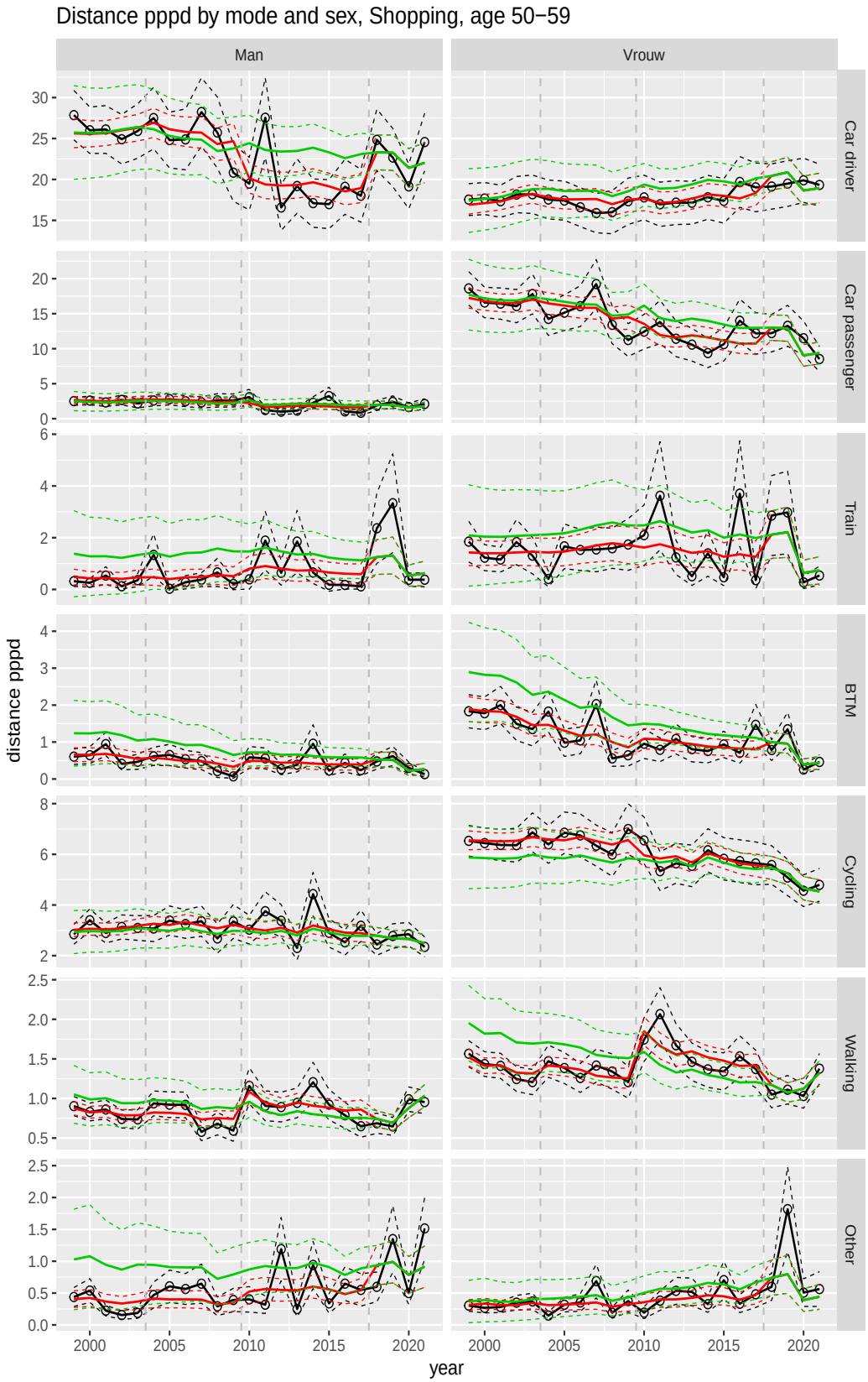

Figure A.216 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

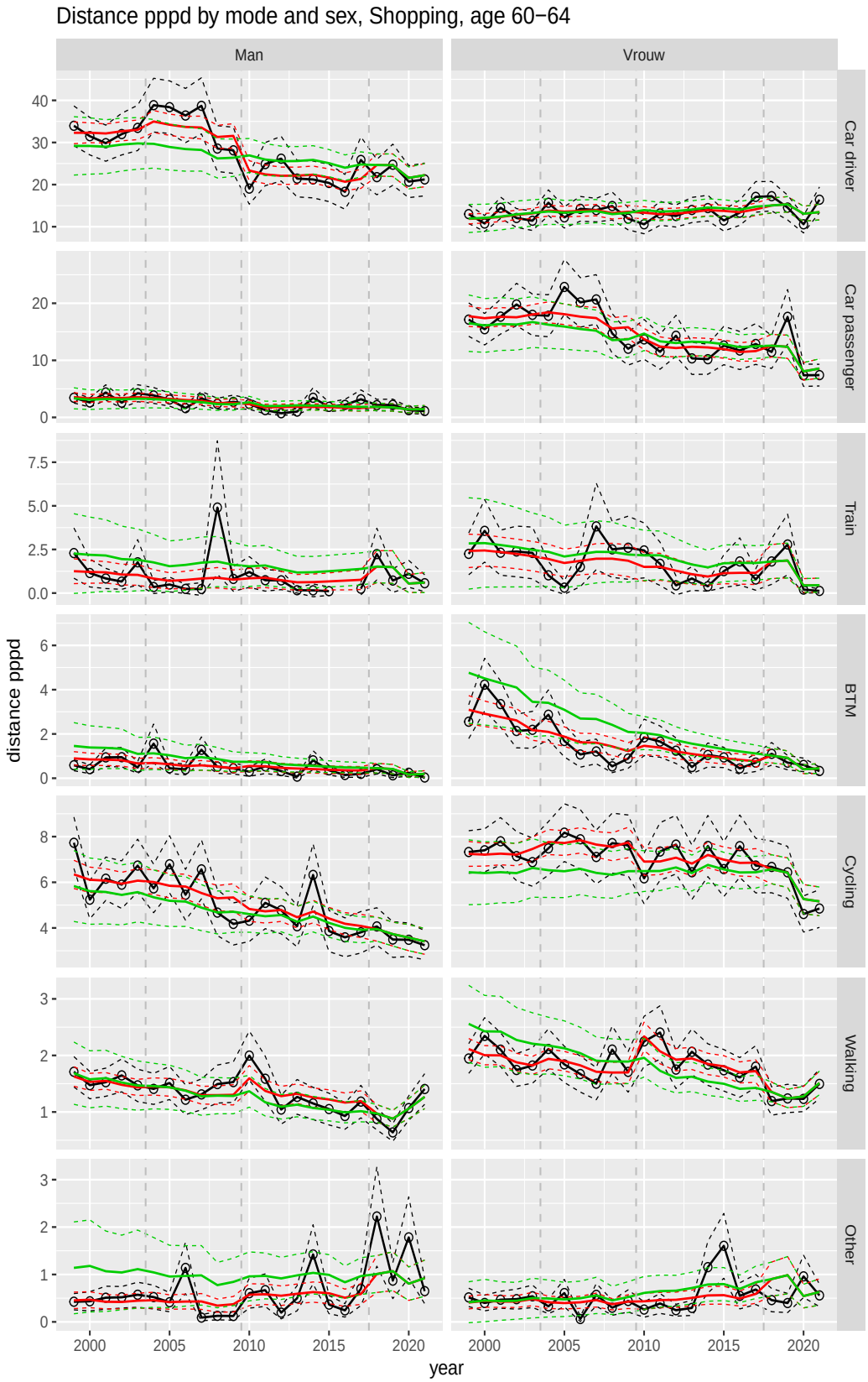

Figure A.217 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

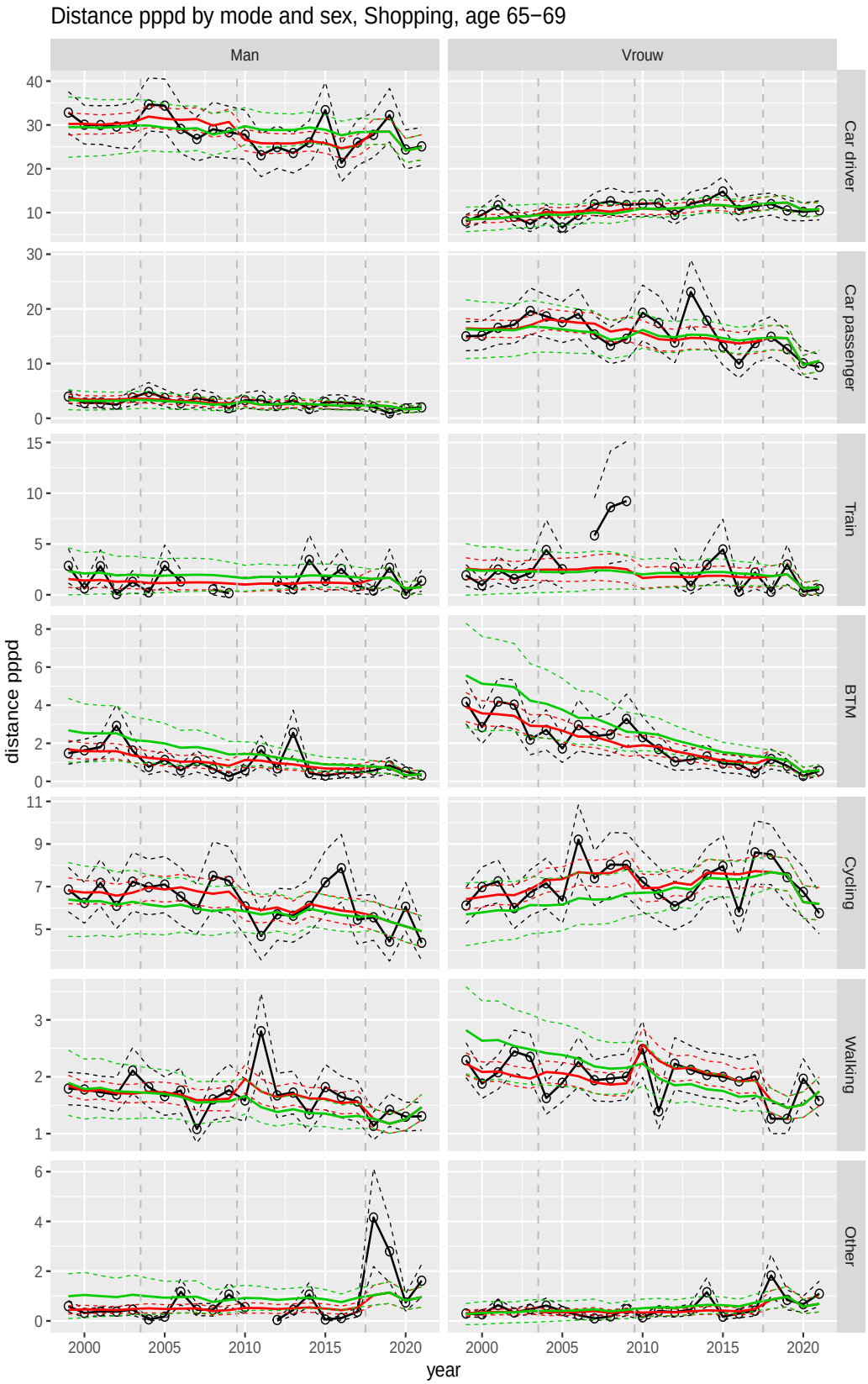

Figure A.218 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

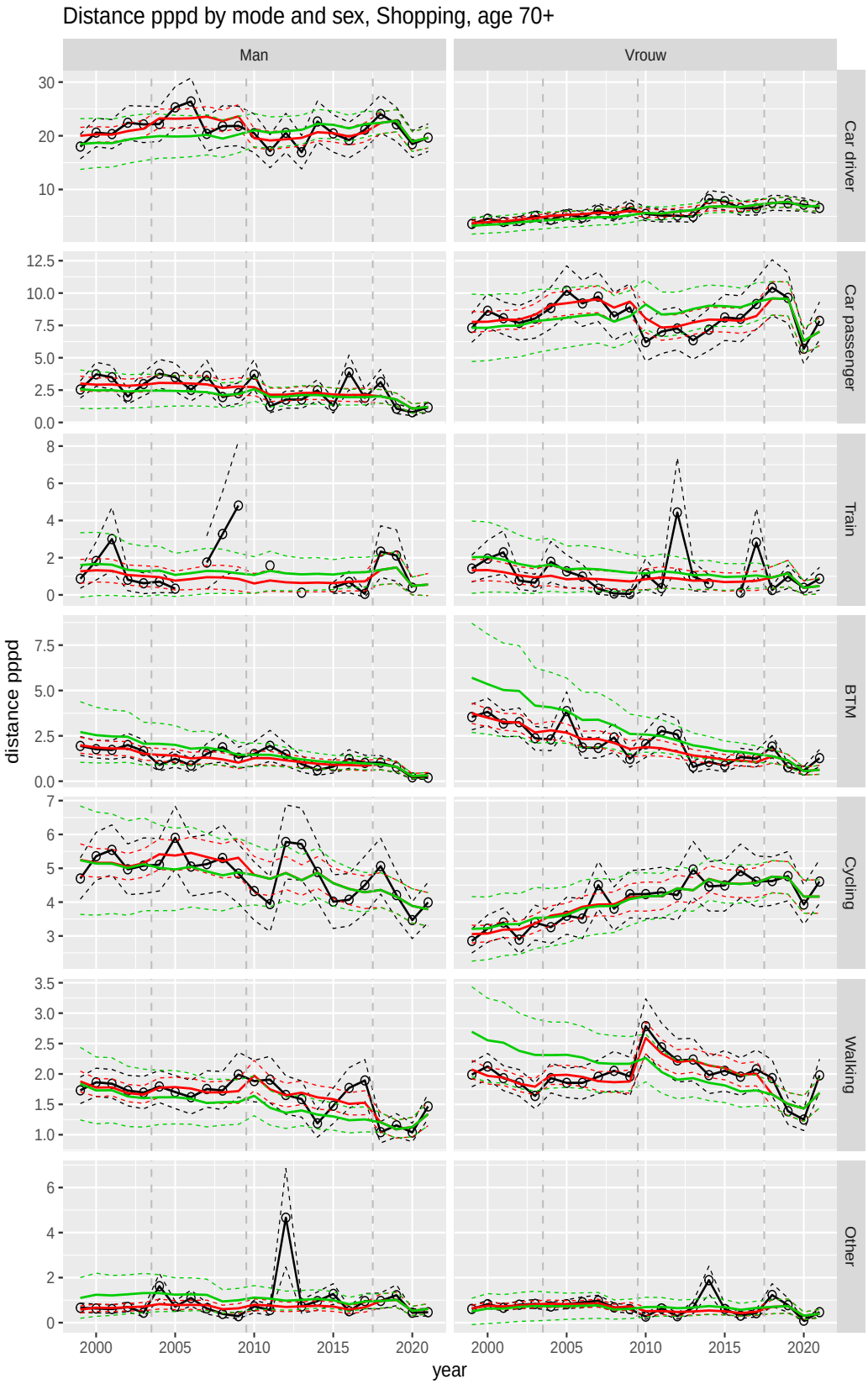

Figure A.219 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

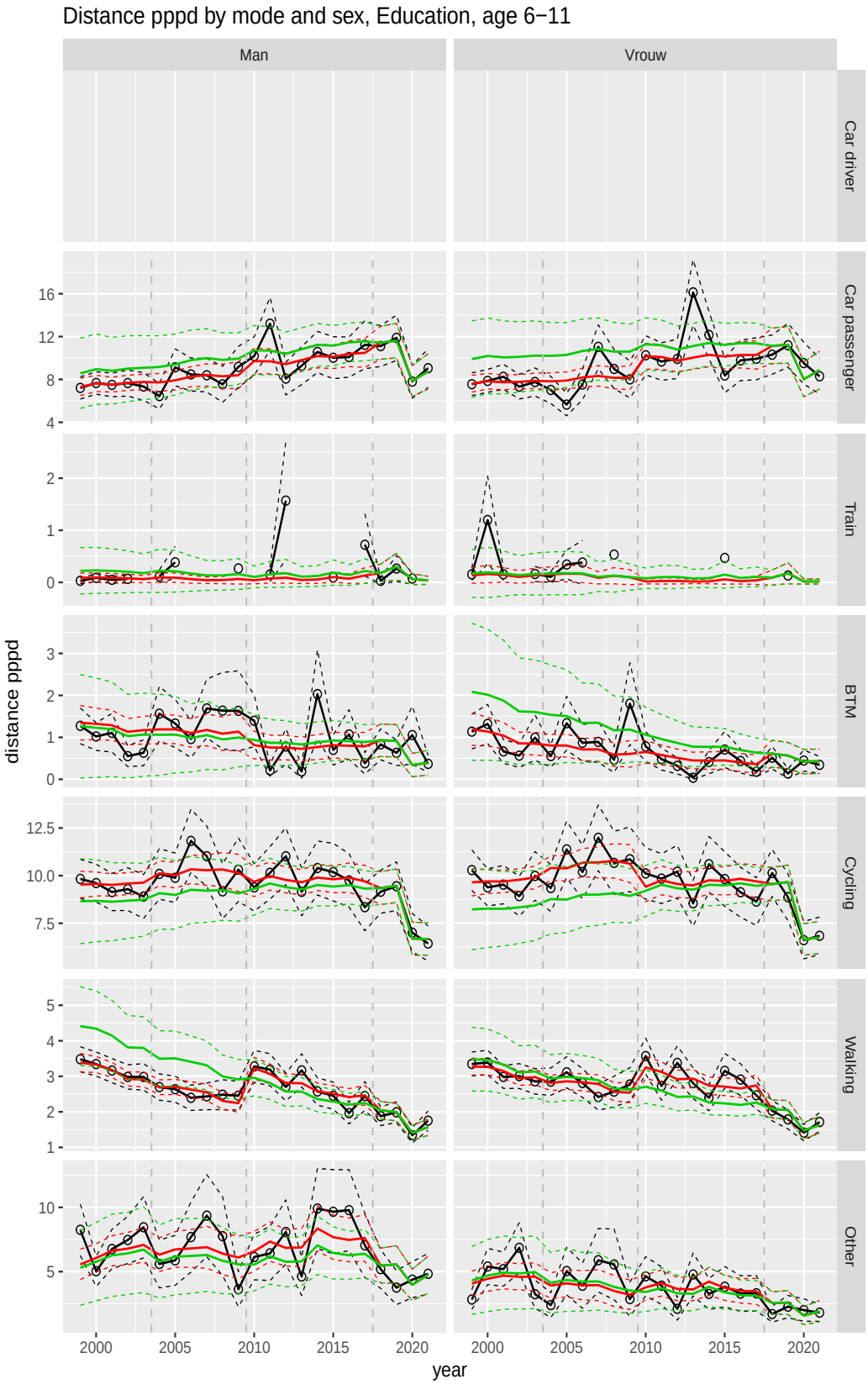

Figure A.220 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

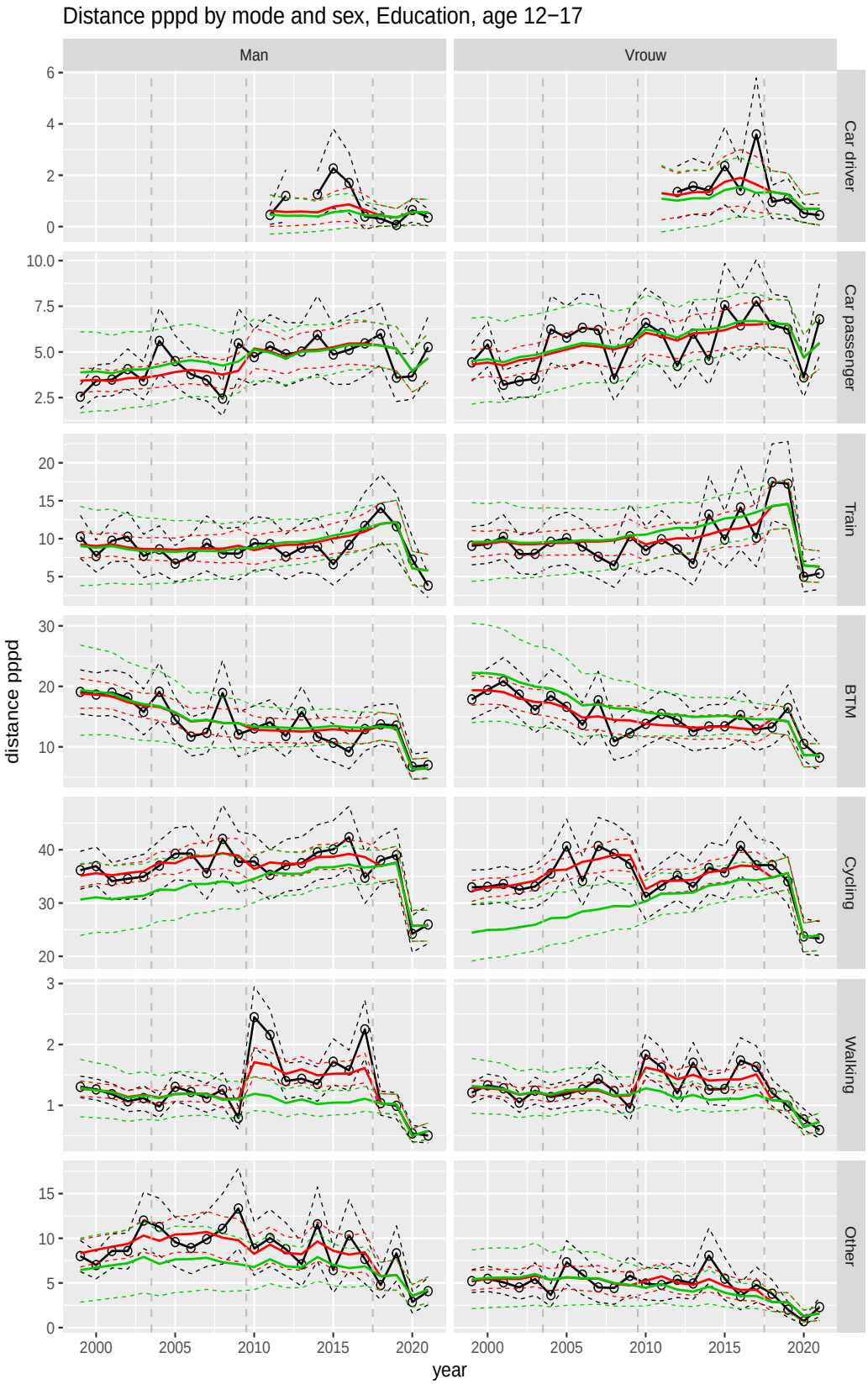

Figure A.221 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

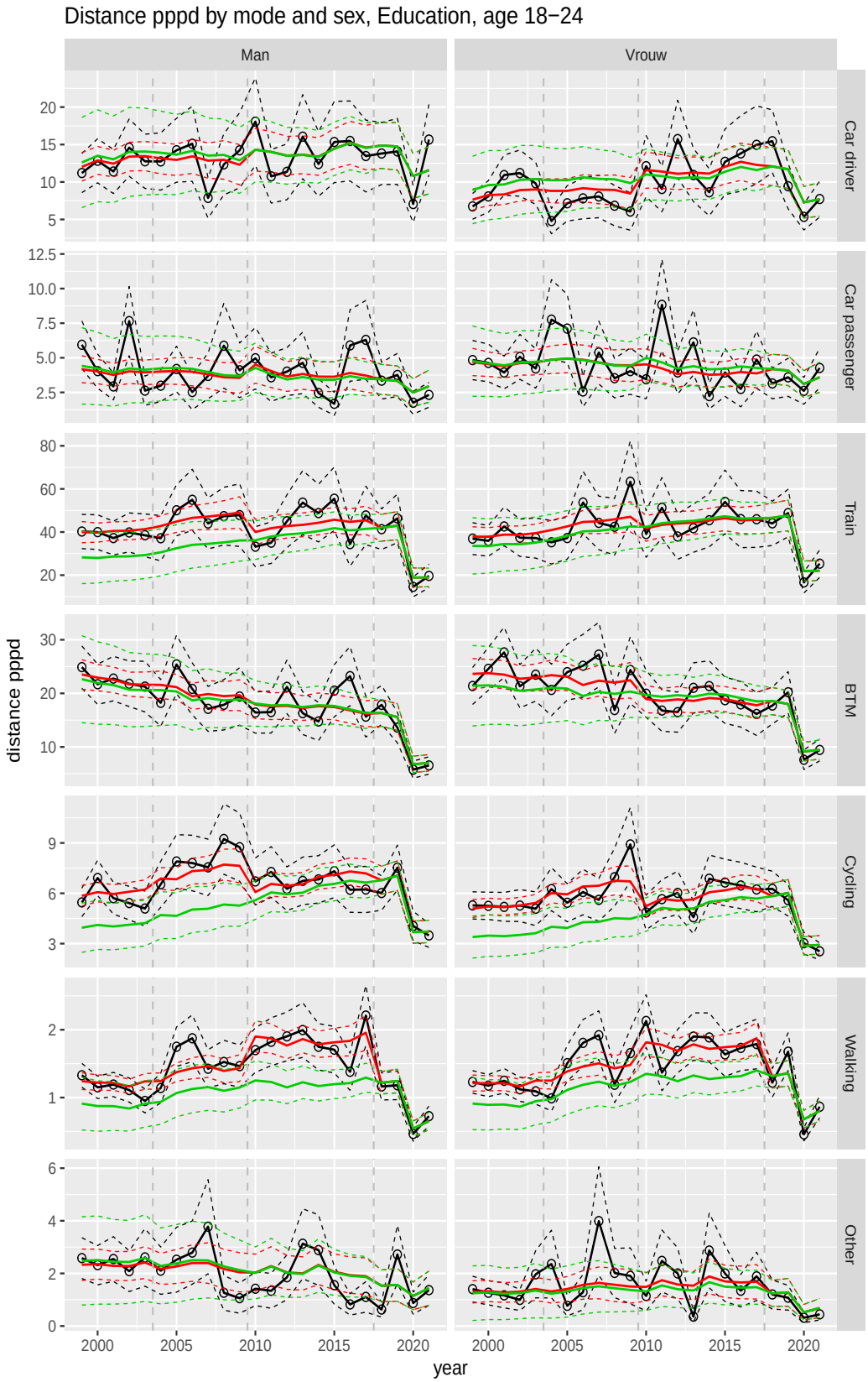

Figure A.222 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

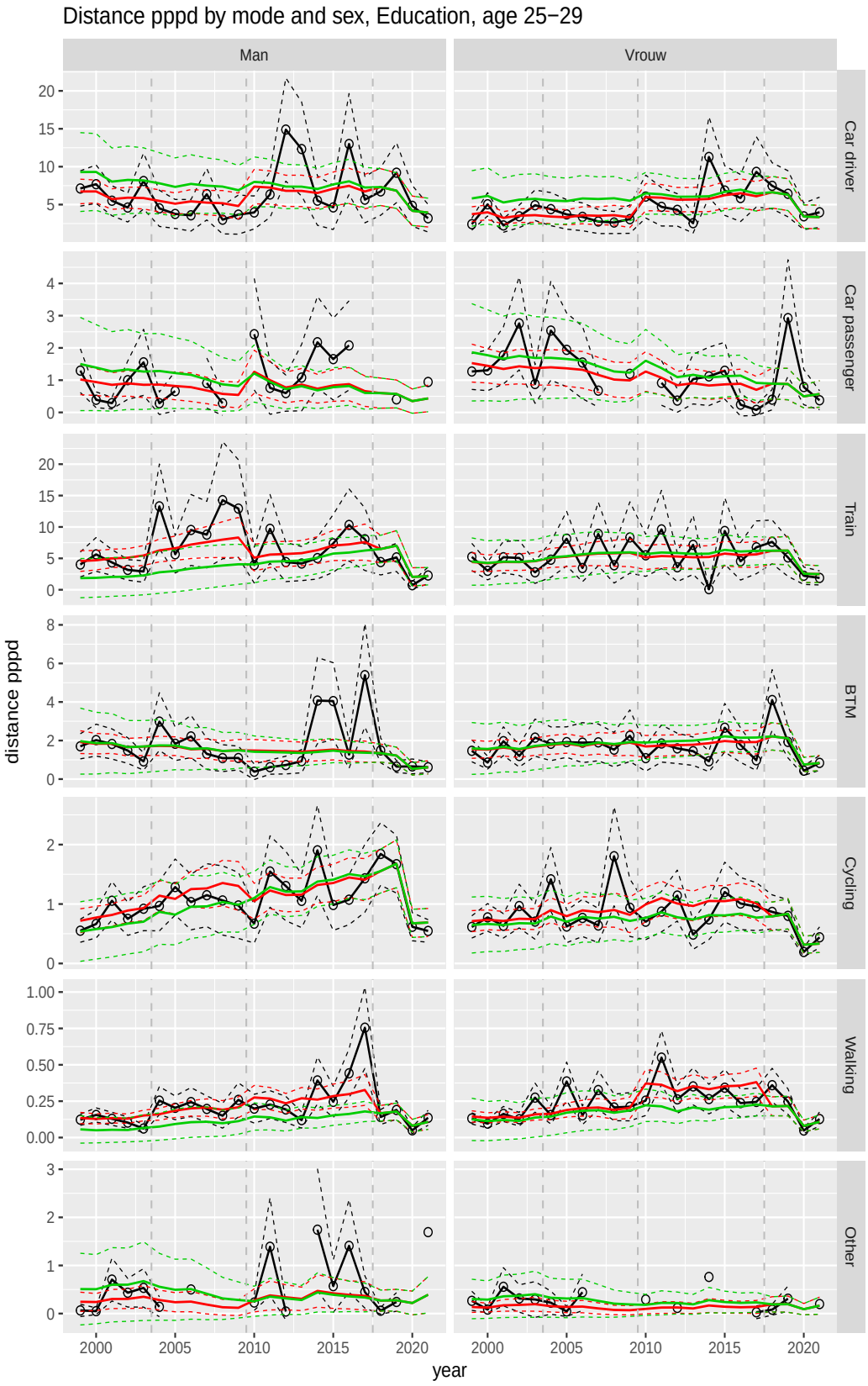

Figure A.223 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

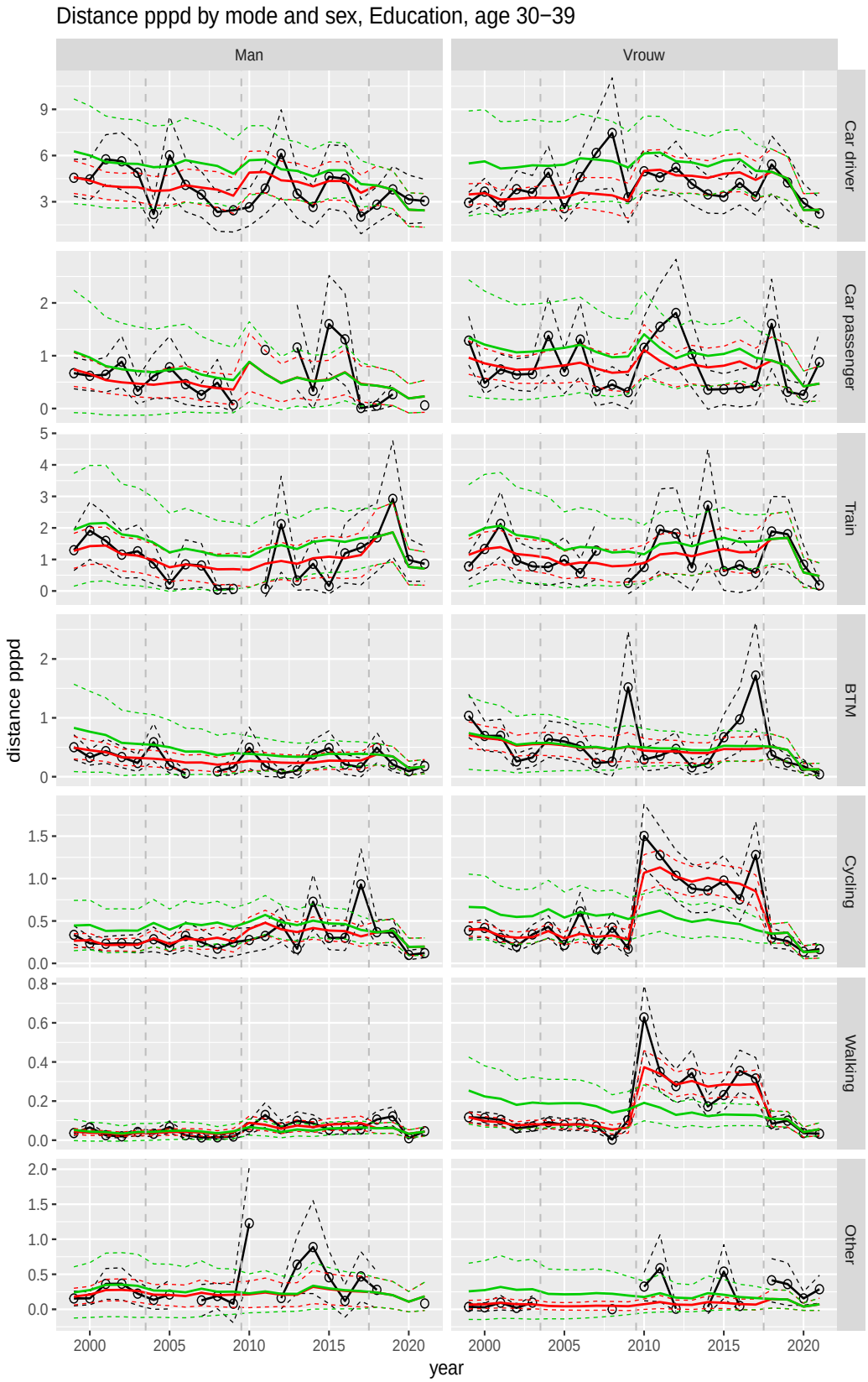

Figure A.224 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

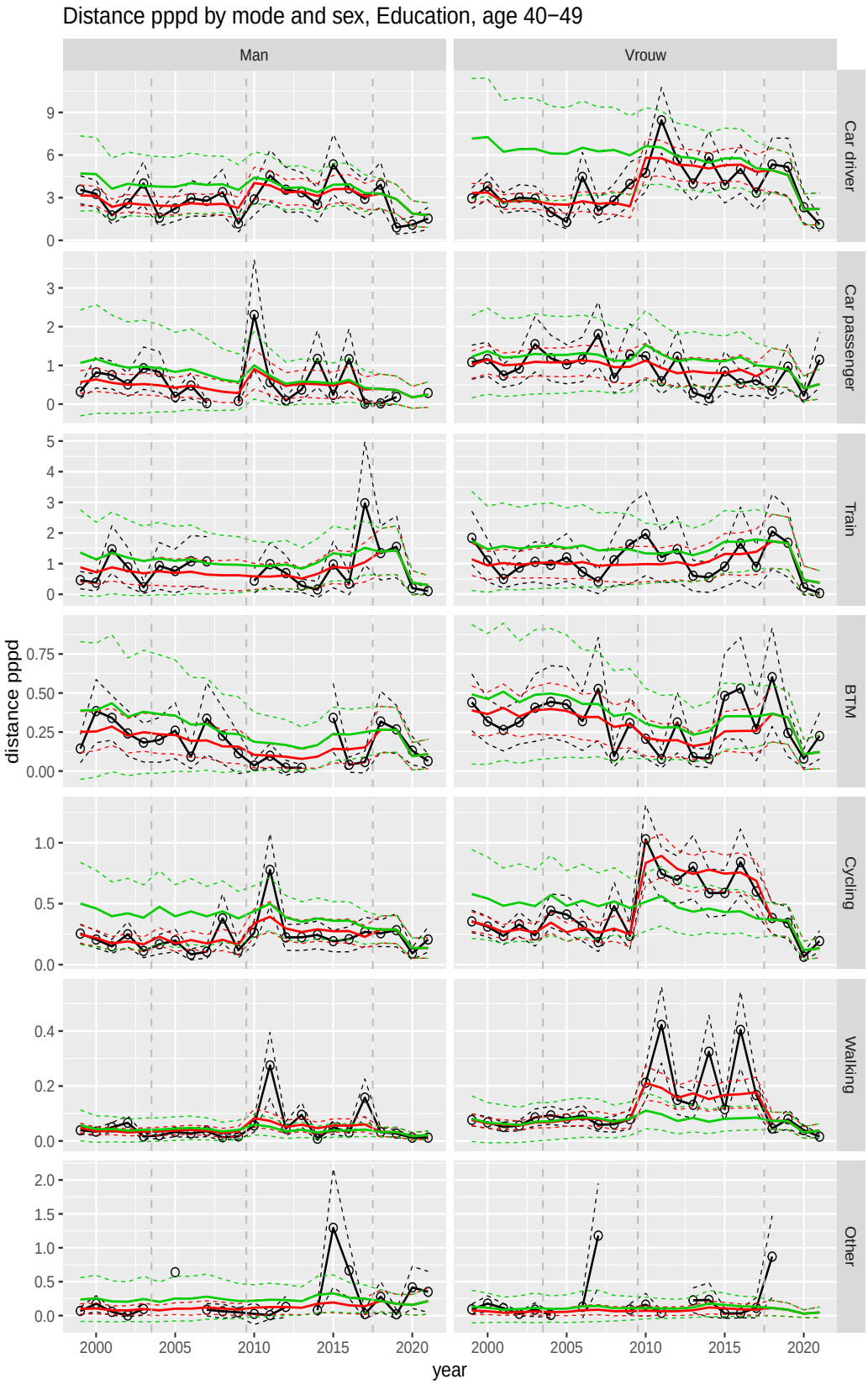

Figure A.225 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

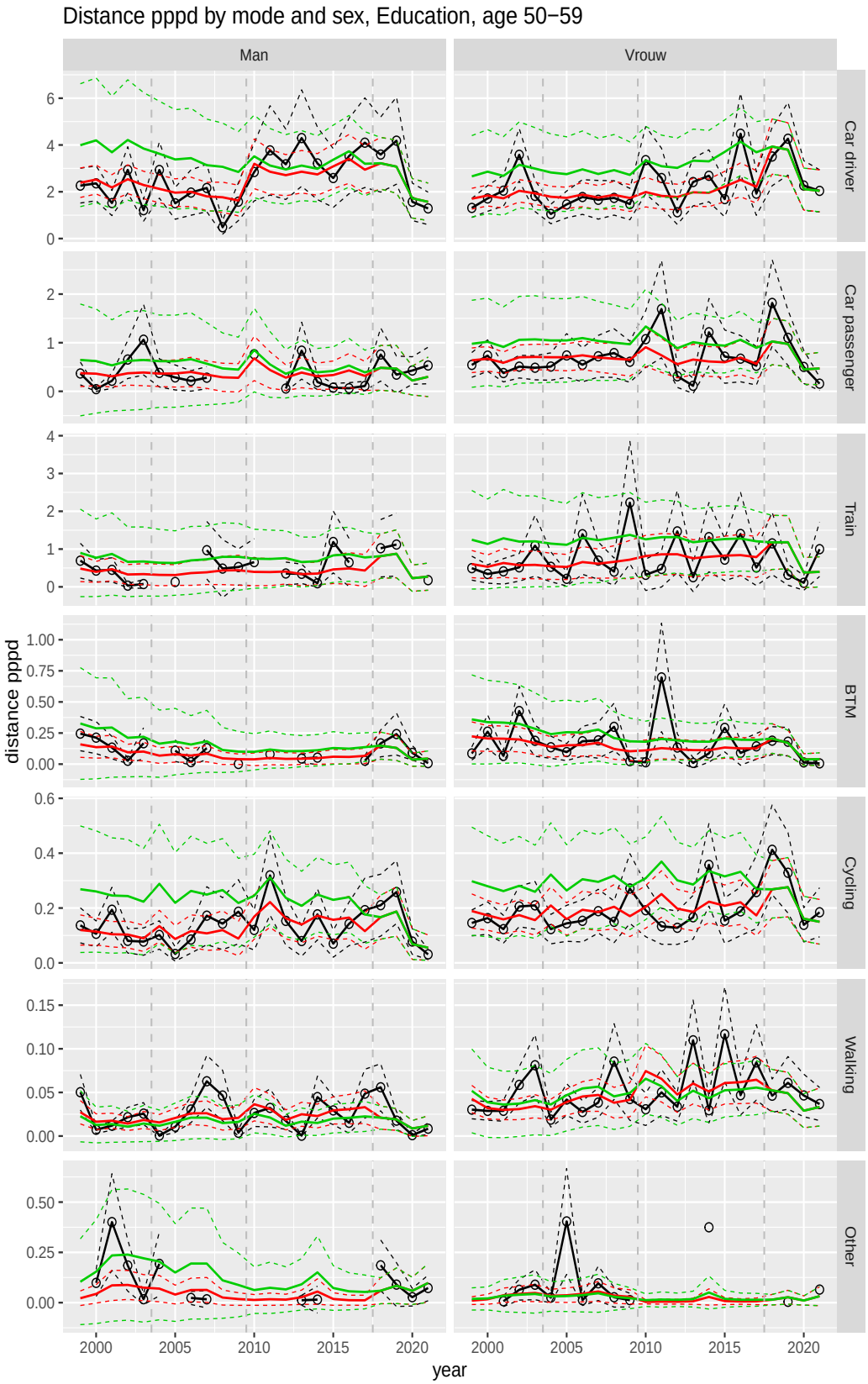

Figure A.226 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

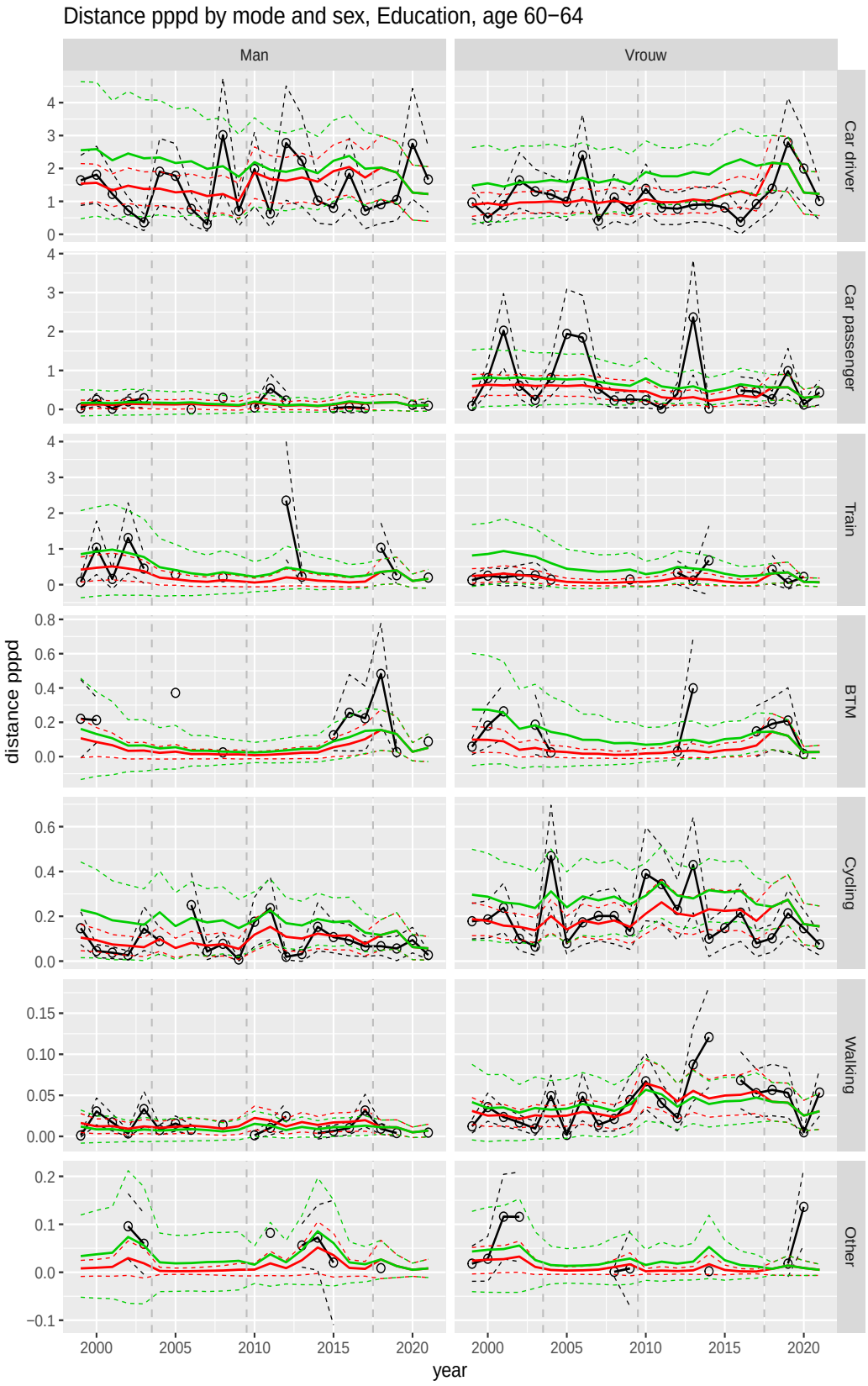

Figure A.227 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

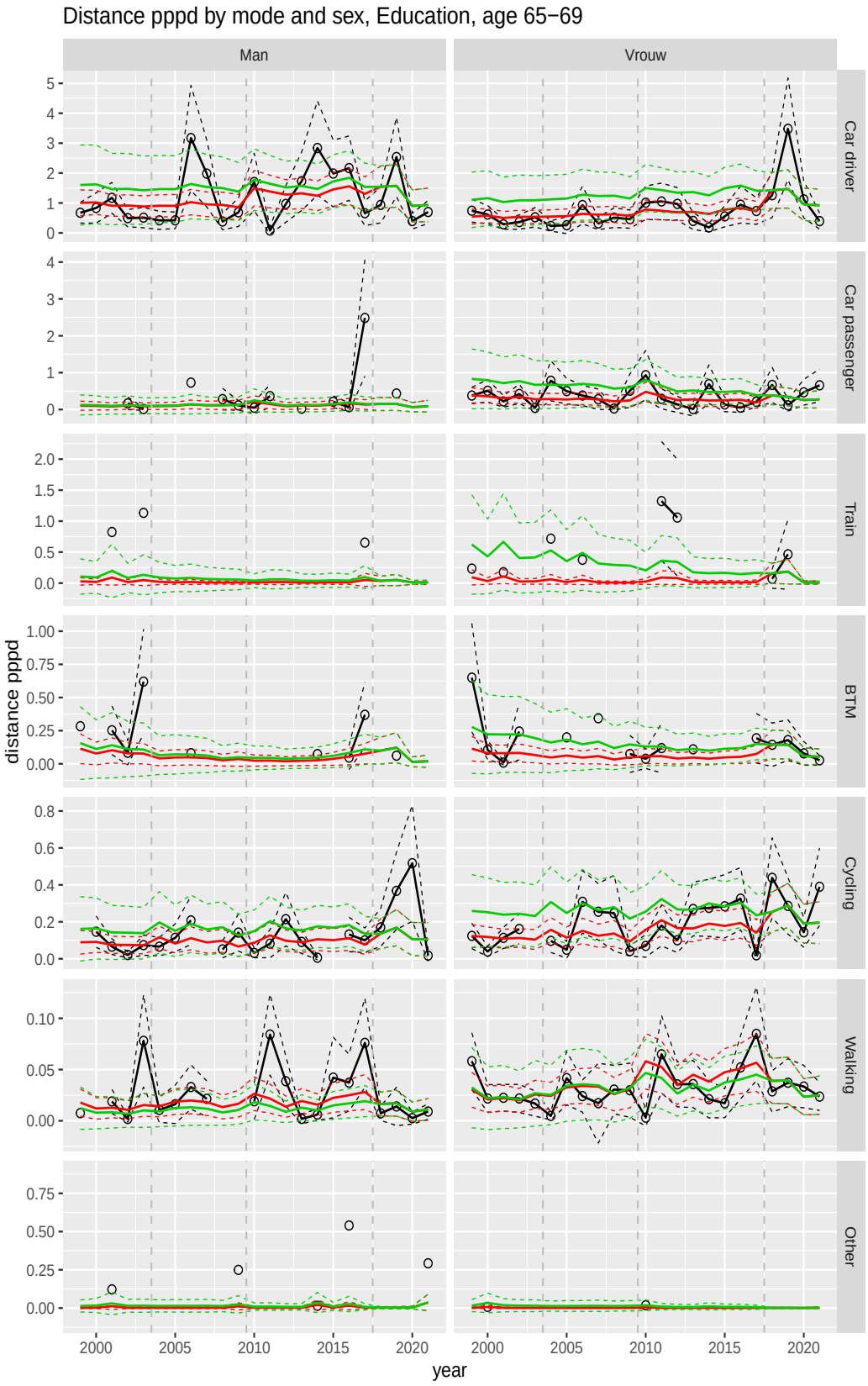

Figure A.228 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

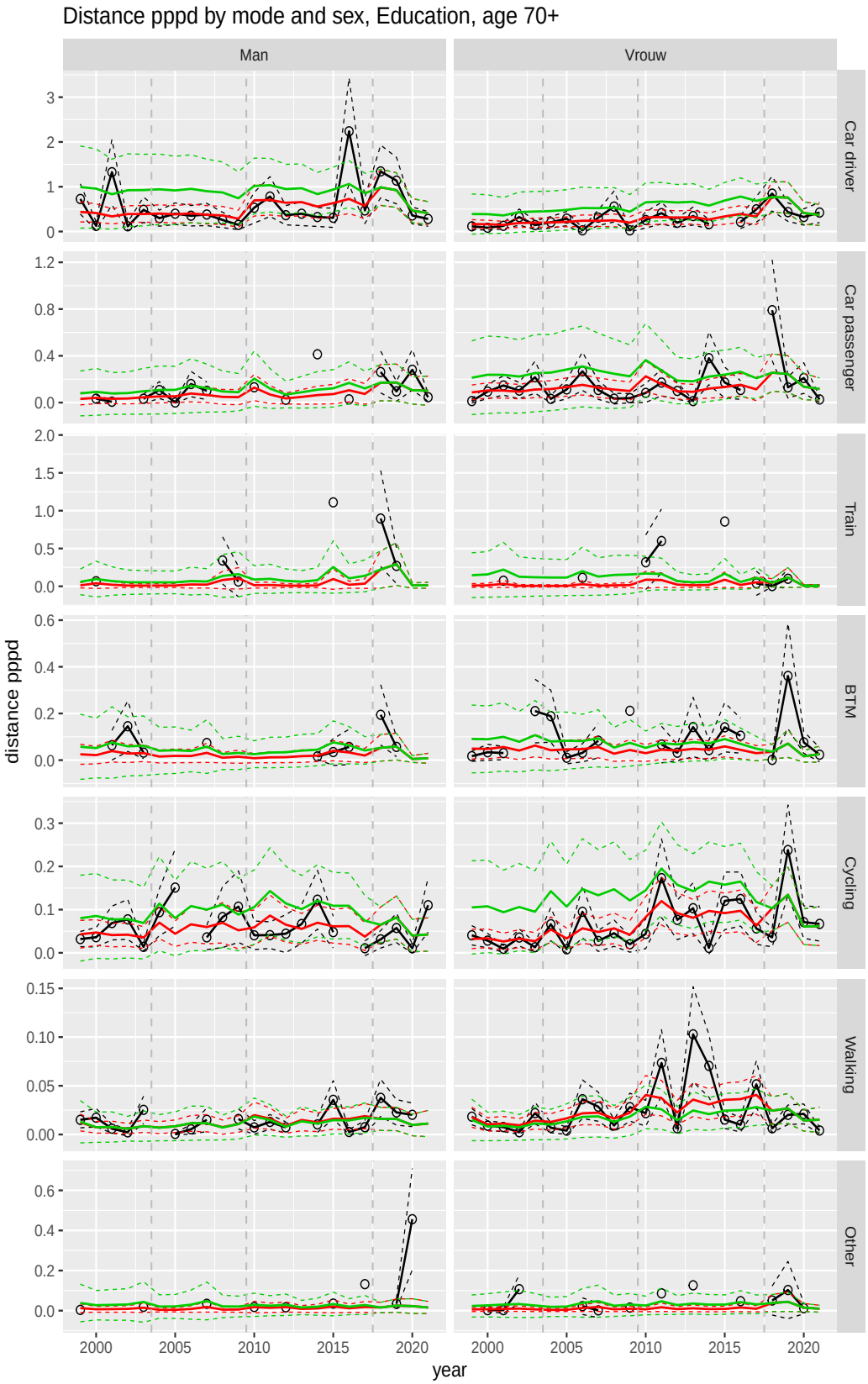

Figure A.229 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

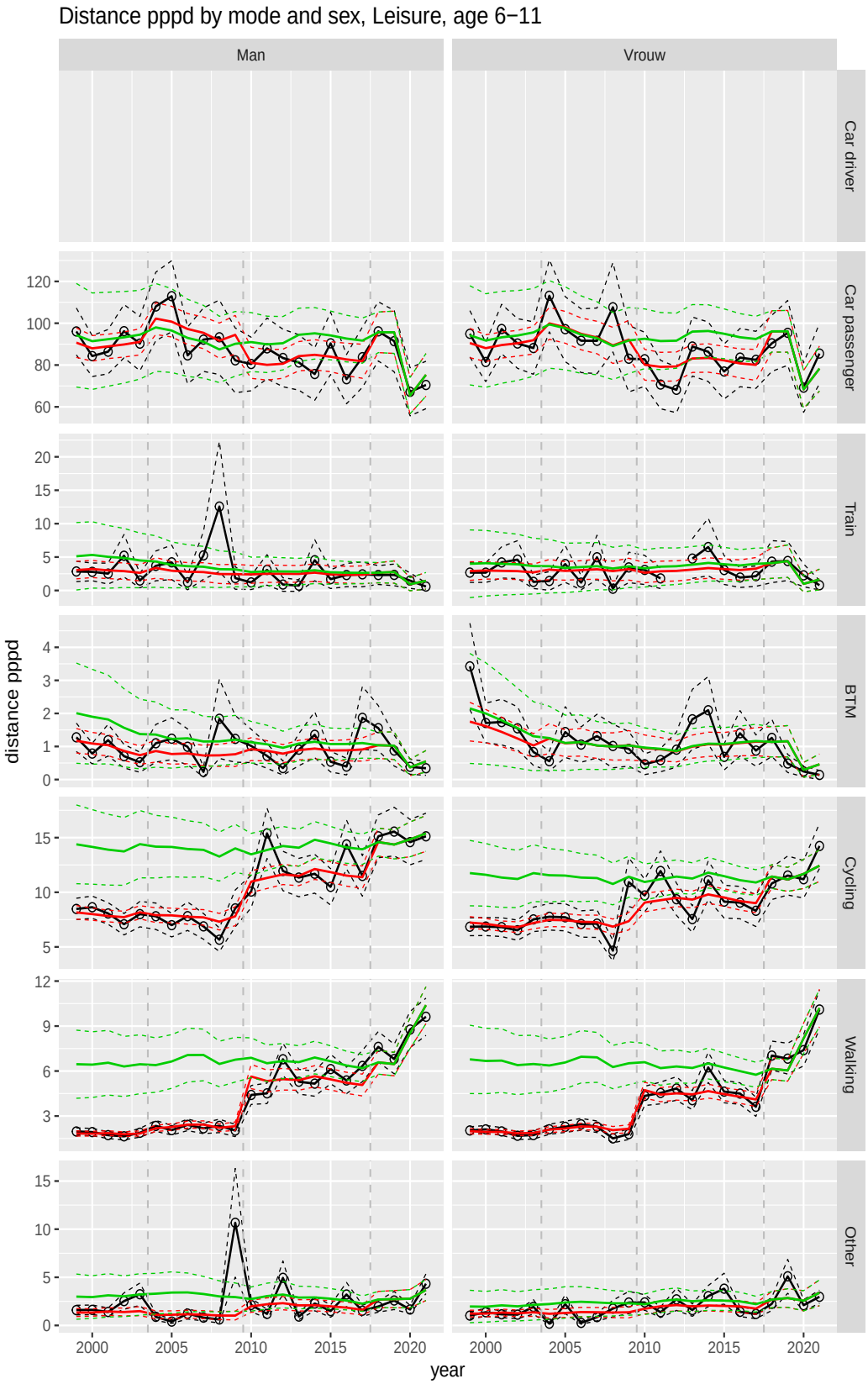

Figure A.230 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

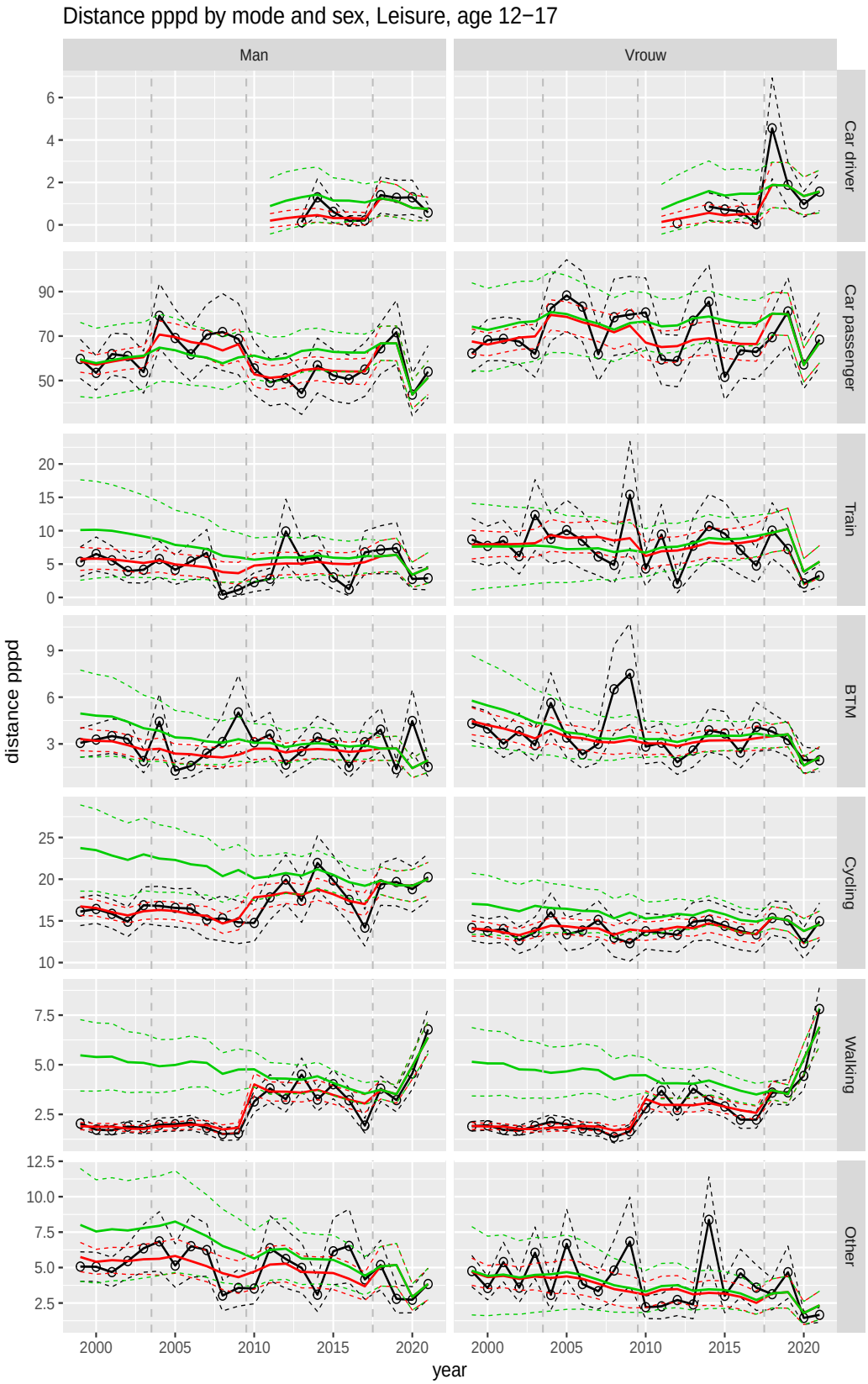

Figure A.231 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

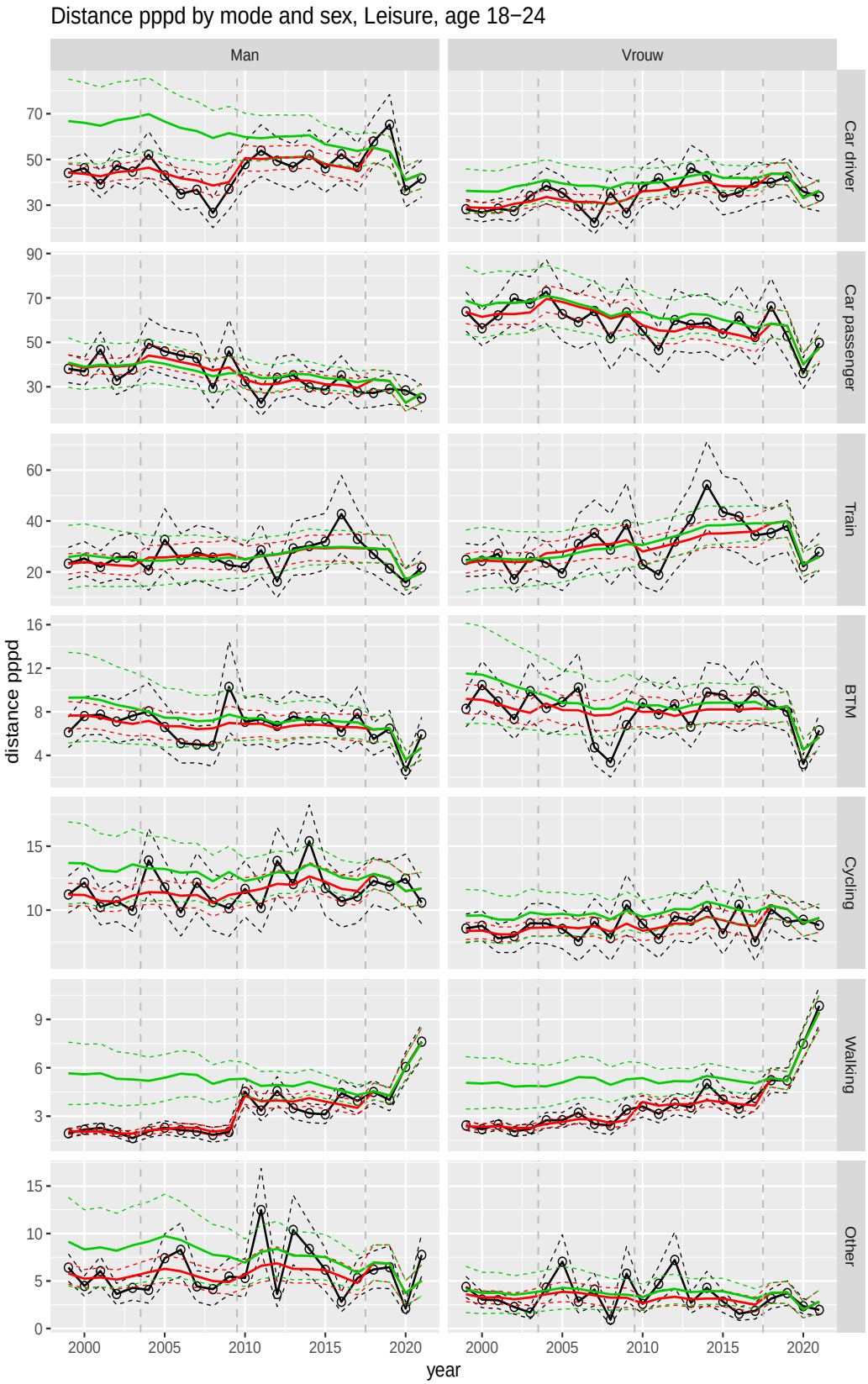

Figure A.232 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.233 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

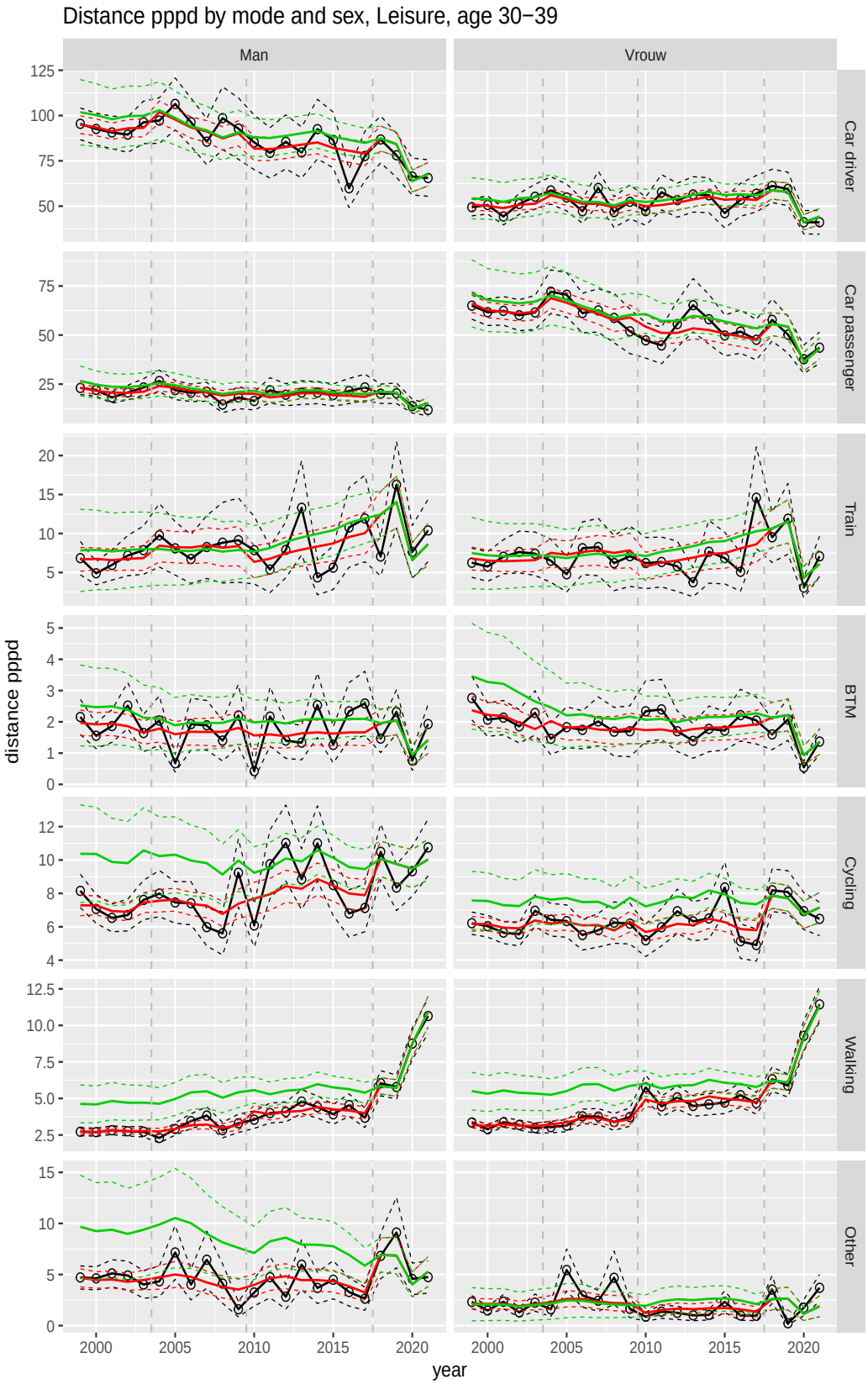

Figure A.234 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

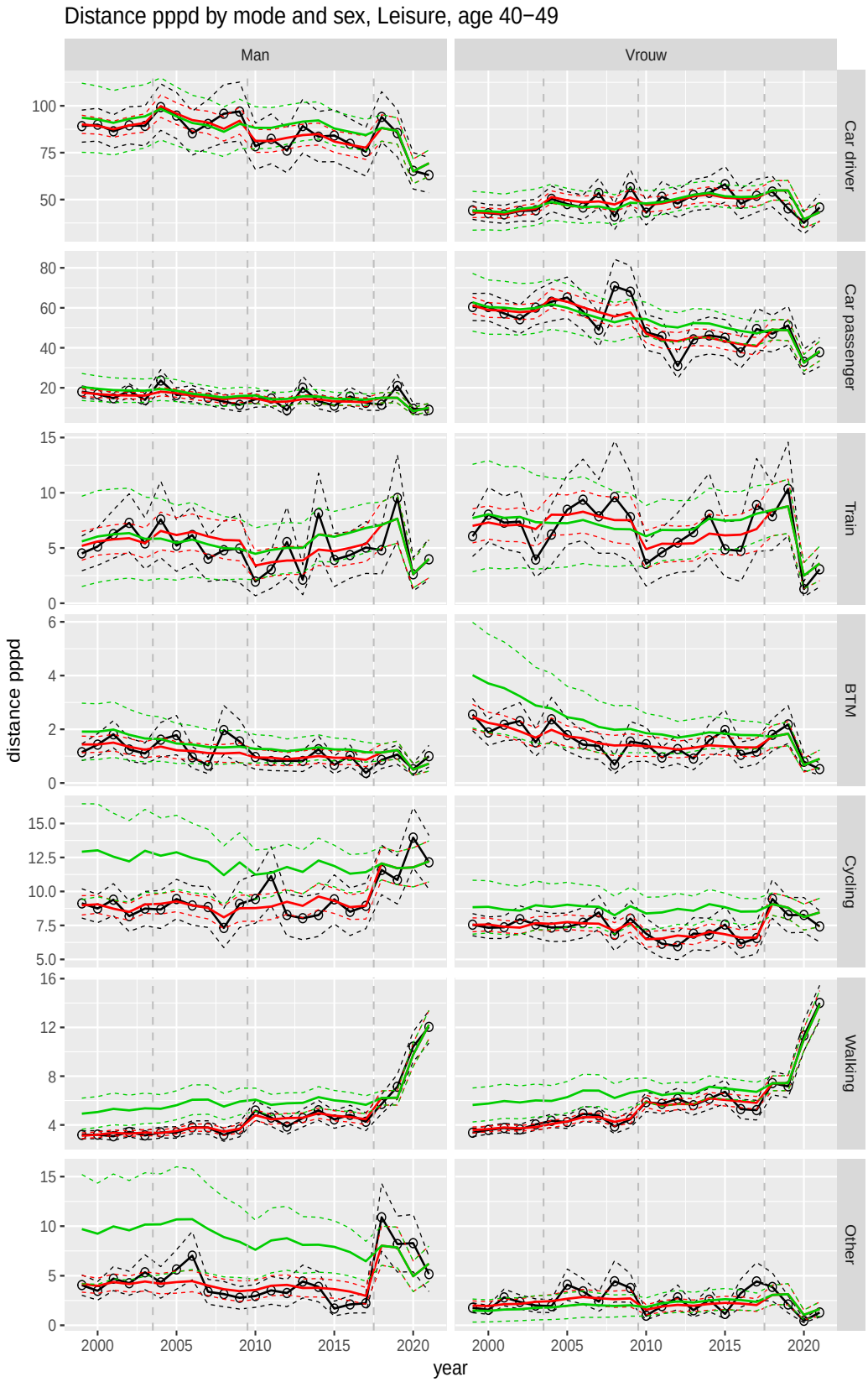

Figure A.235 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

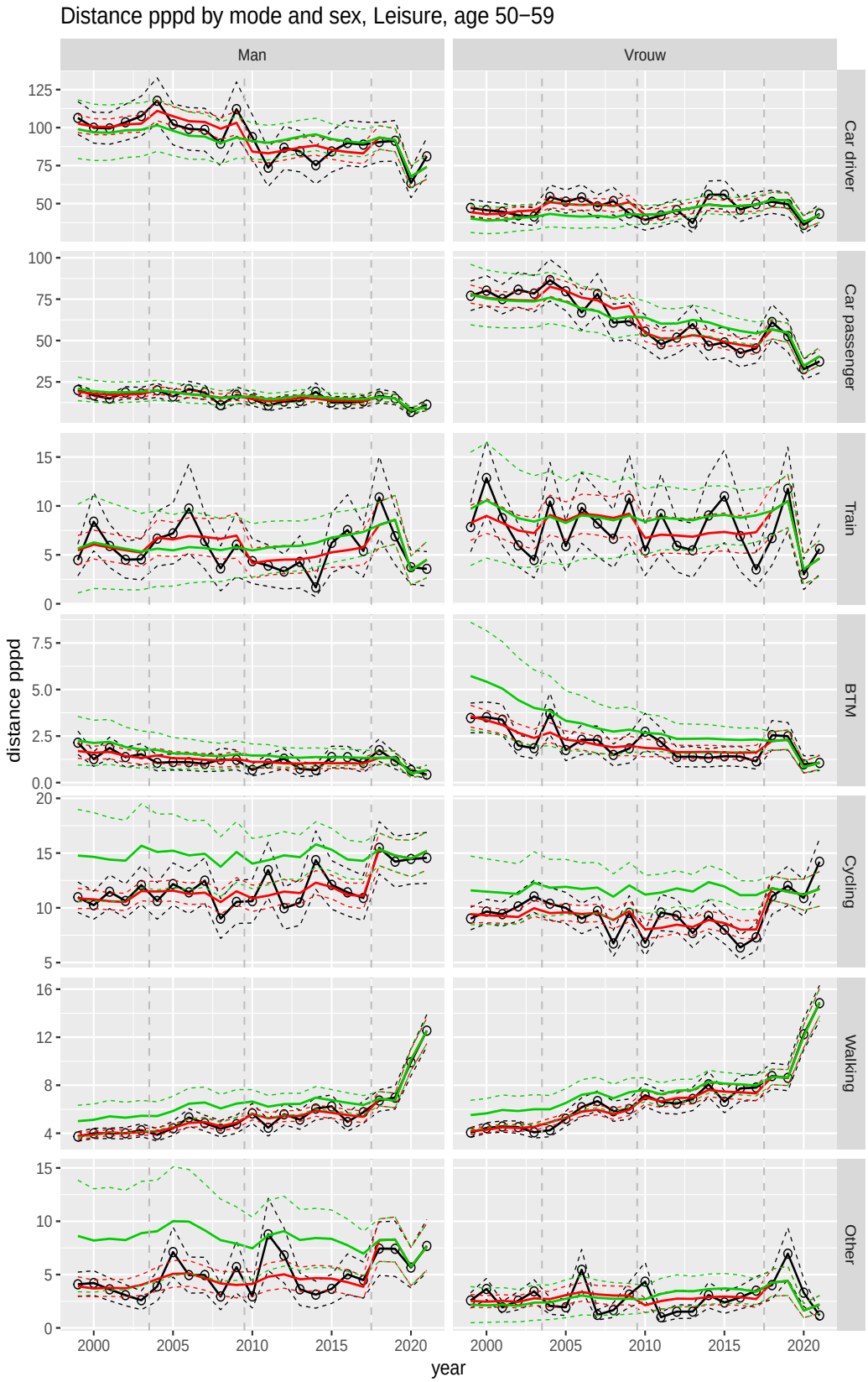

Figure A.236 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

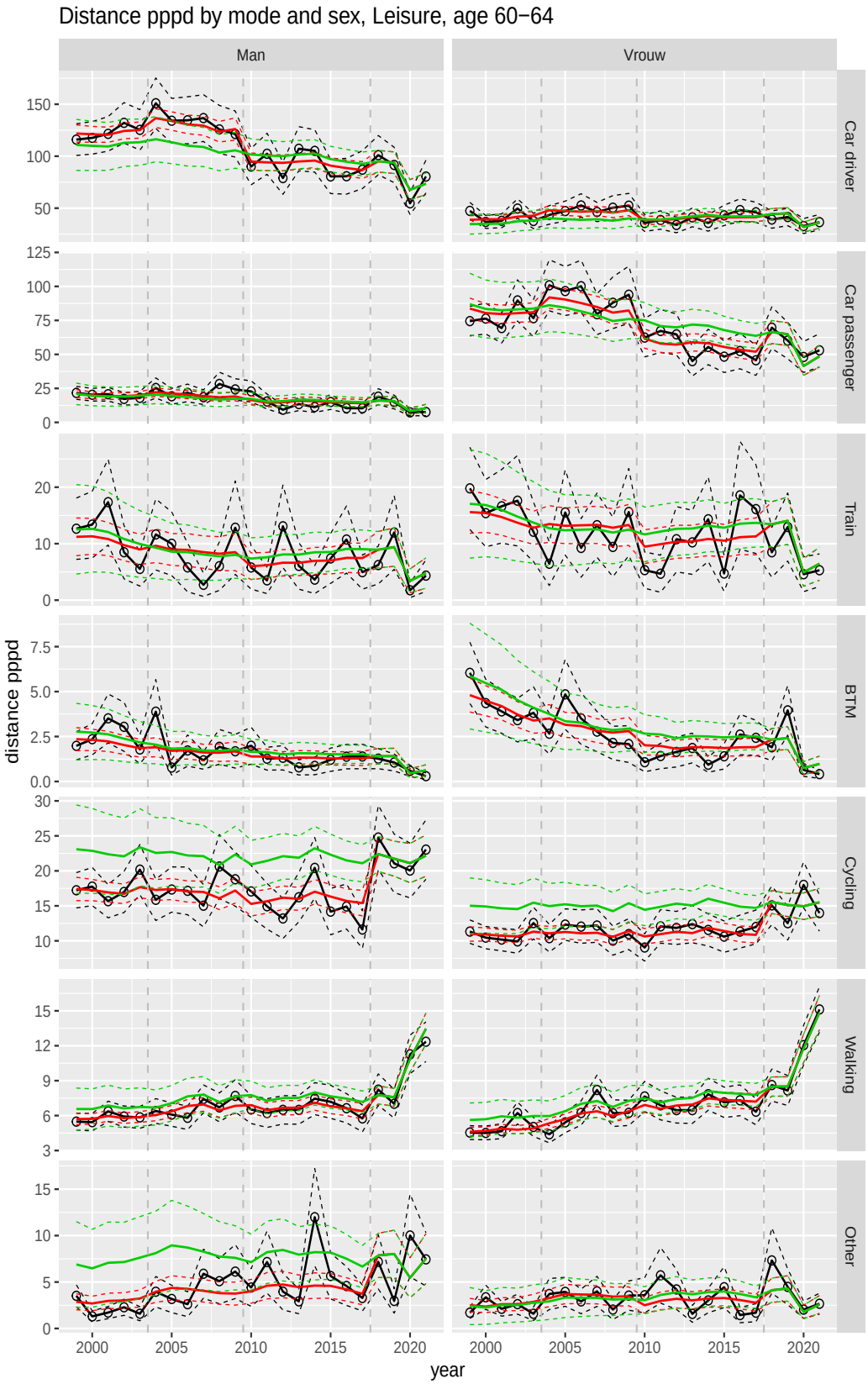

Figure A.237 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

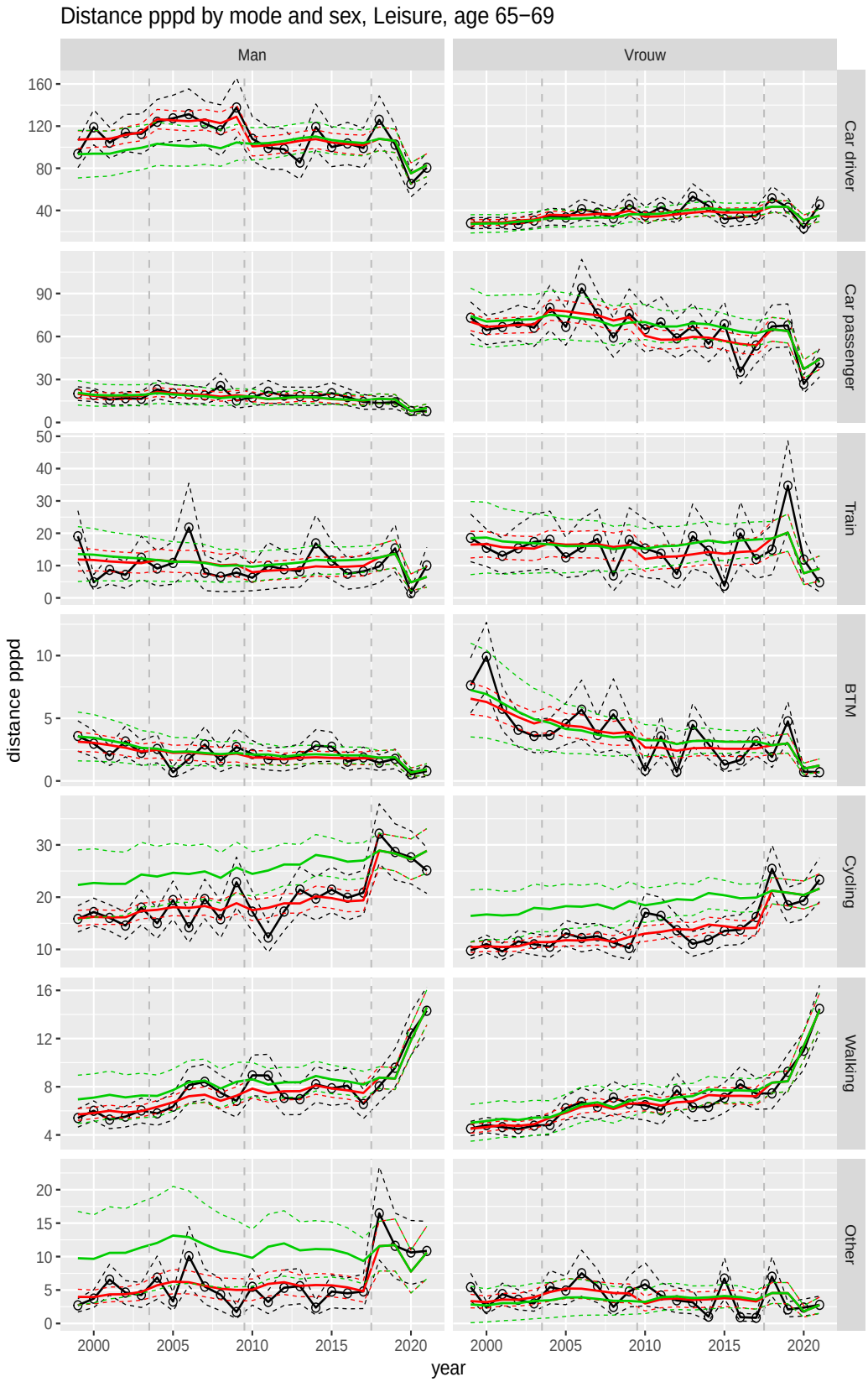

Figure A.238 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

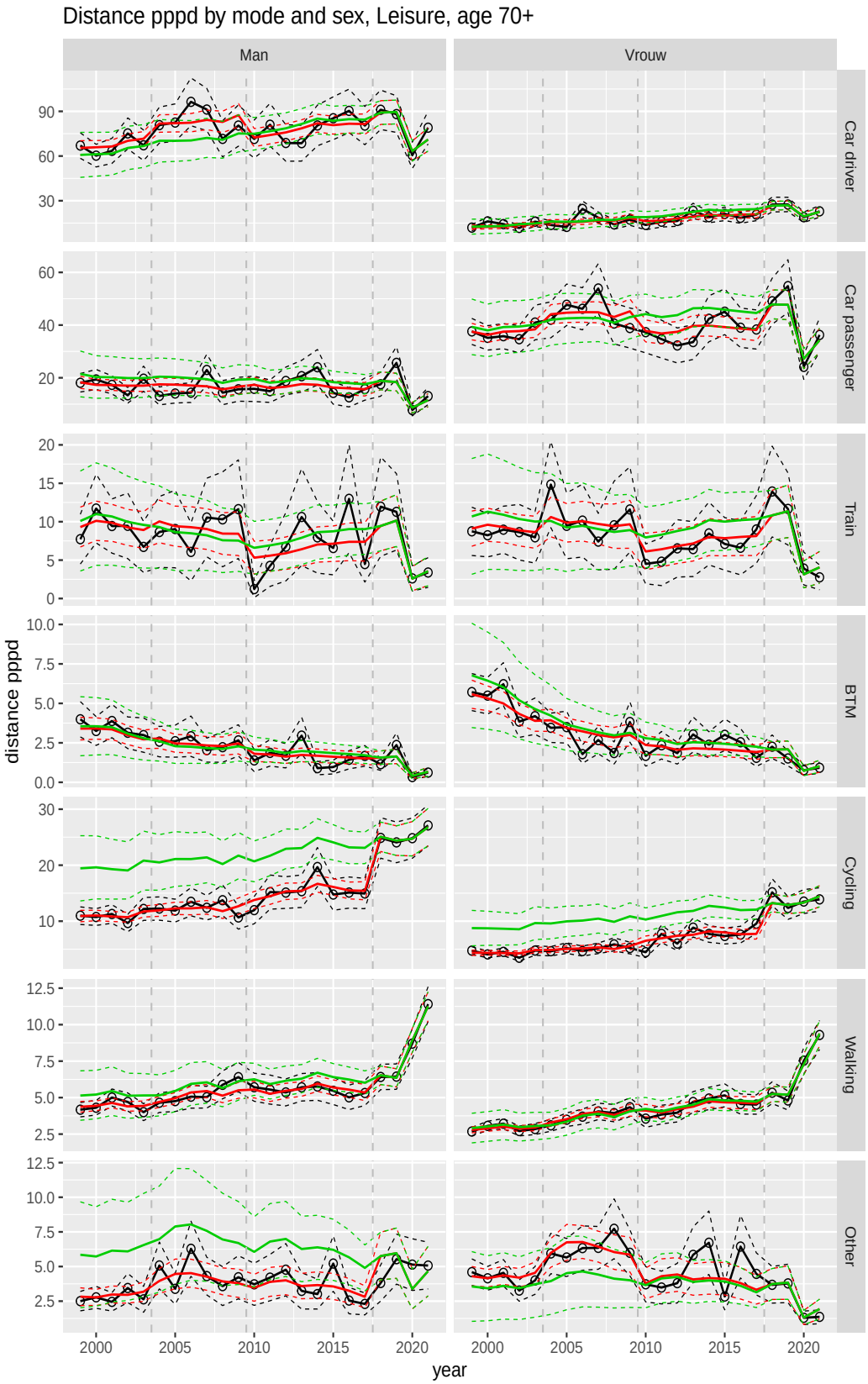

Figure A.239 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

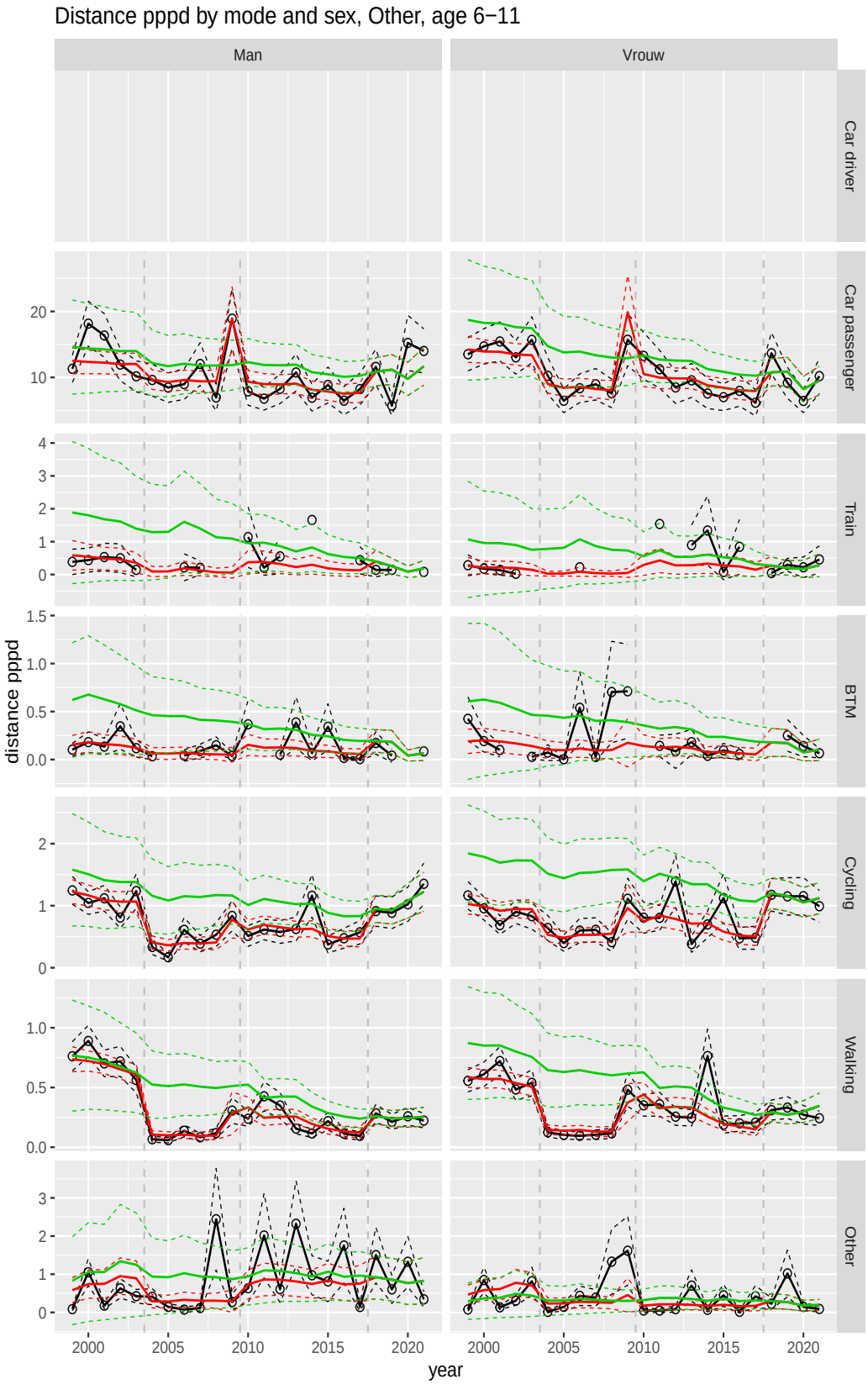

Figure A.240 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

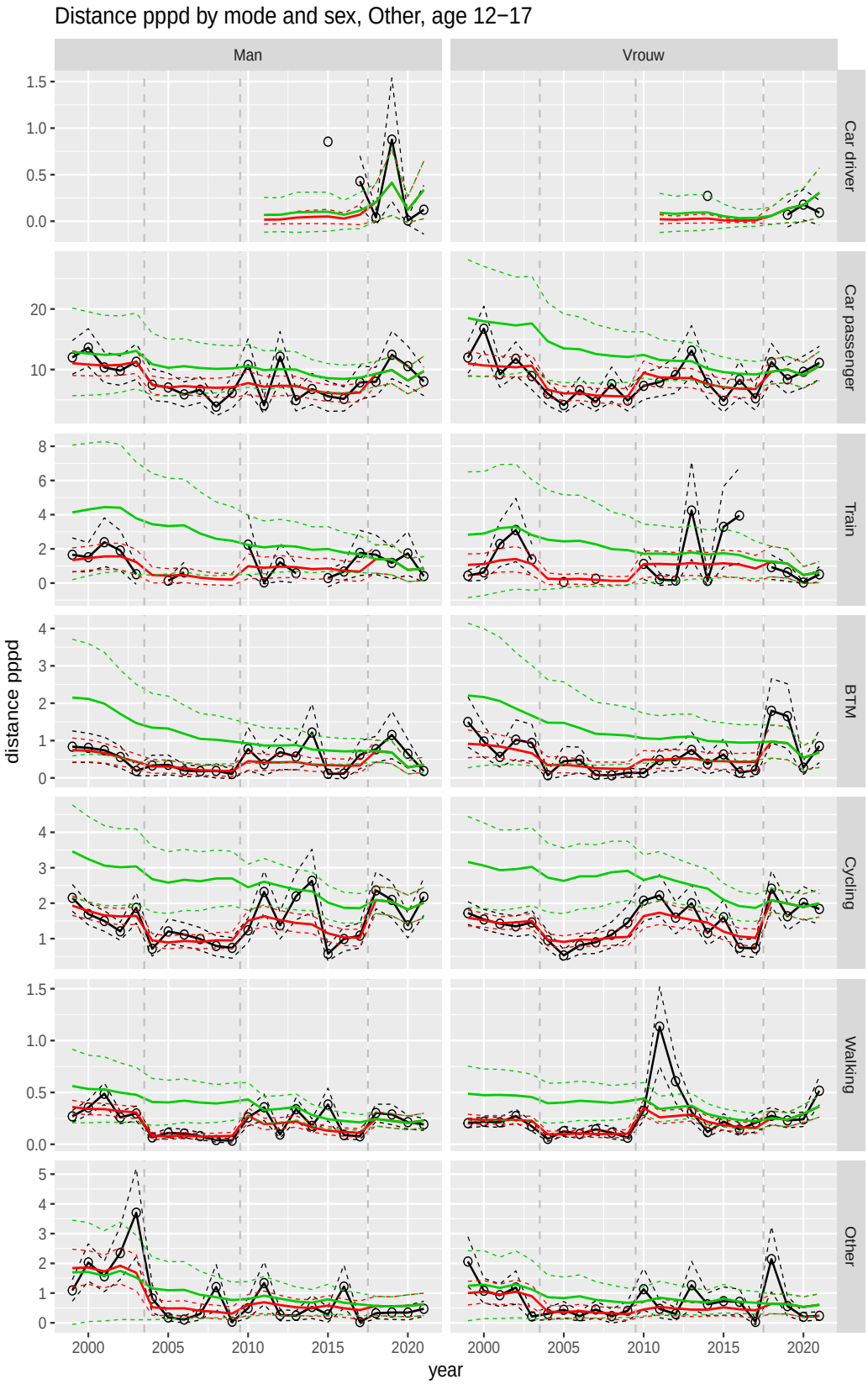

Figure A.241 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

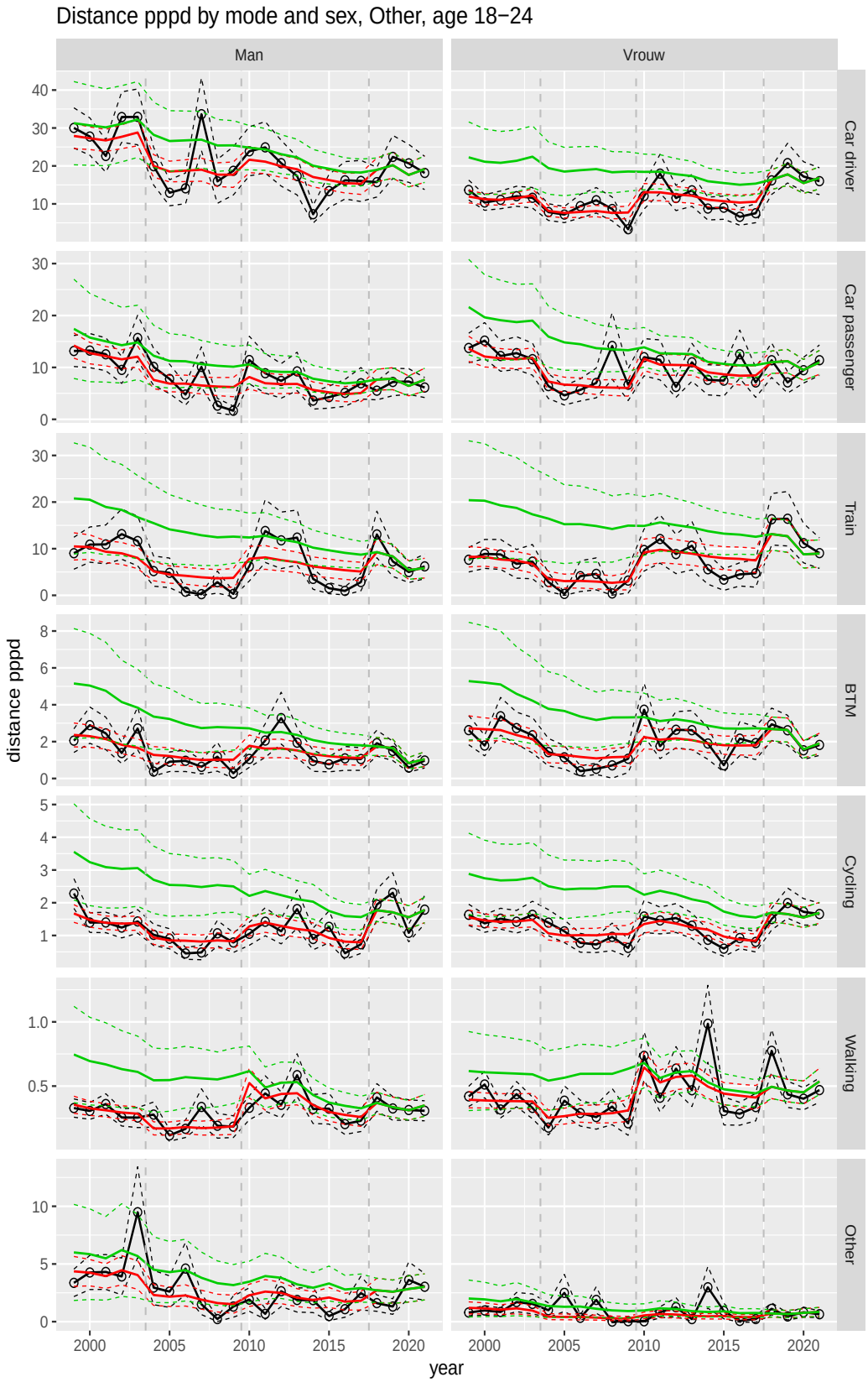

Figure A.242 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

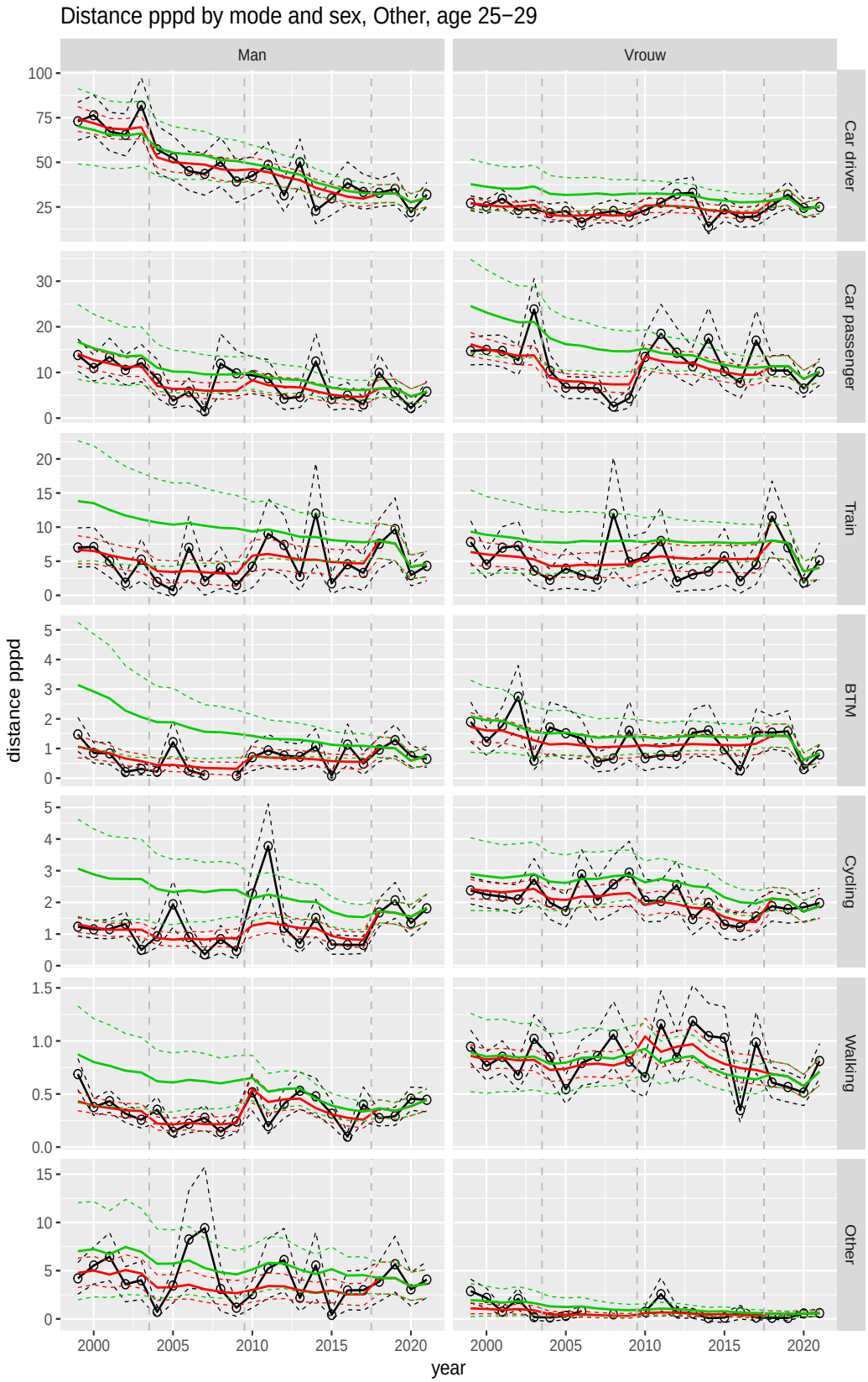

Figure A.243 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

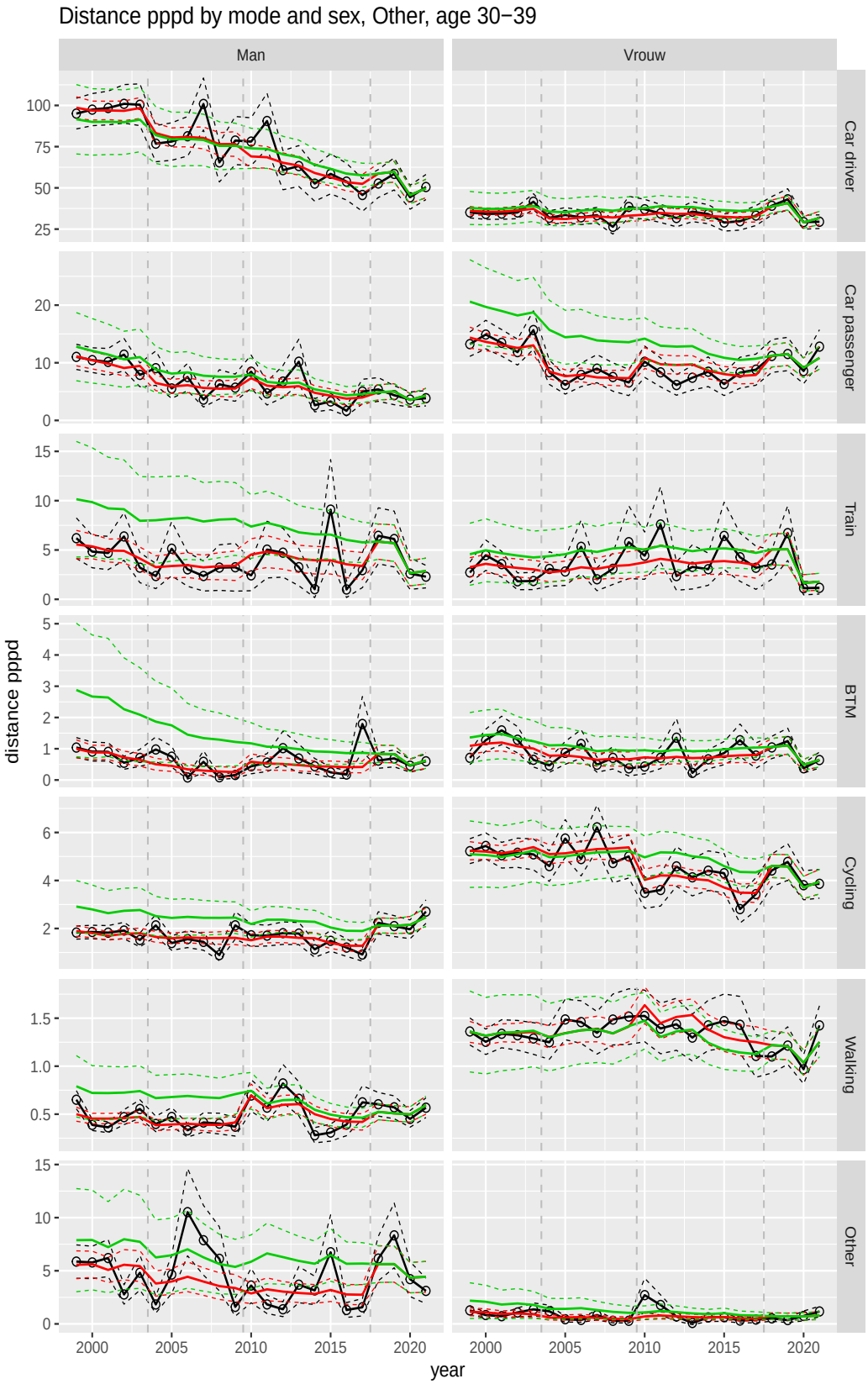

Figure A.244 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

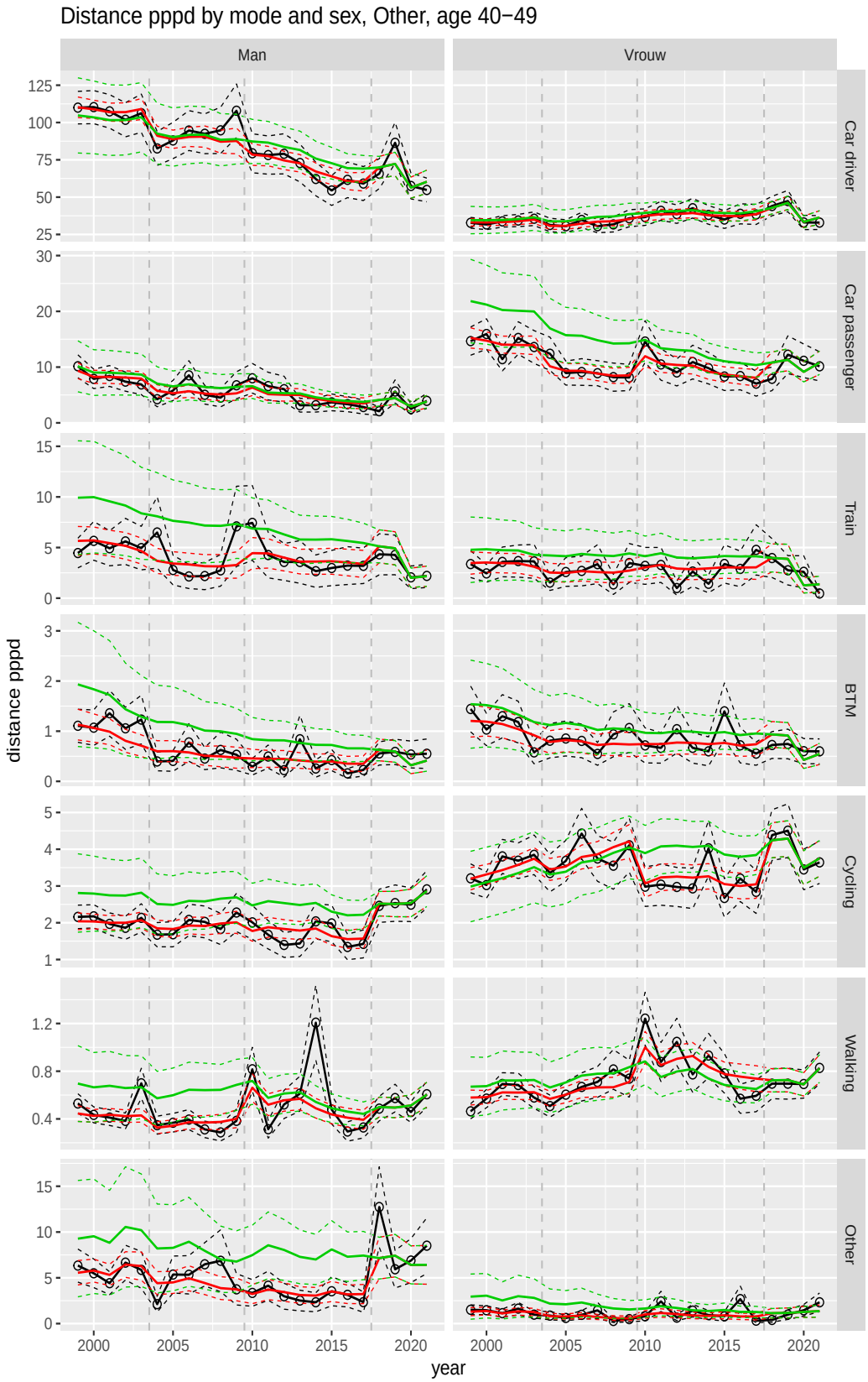

Figure A.245 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

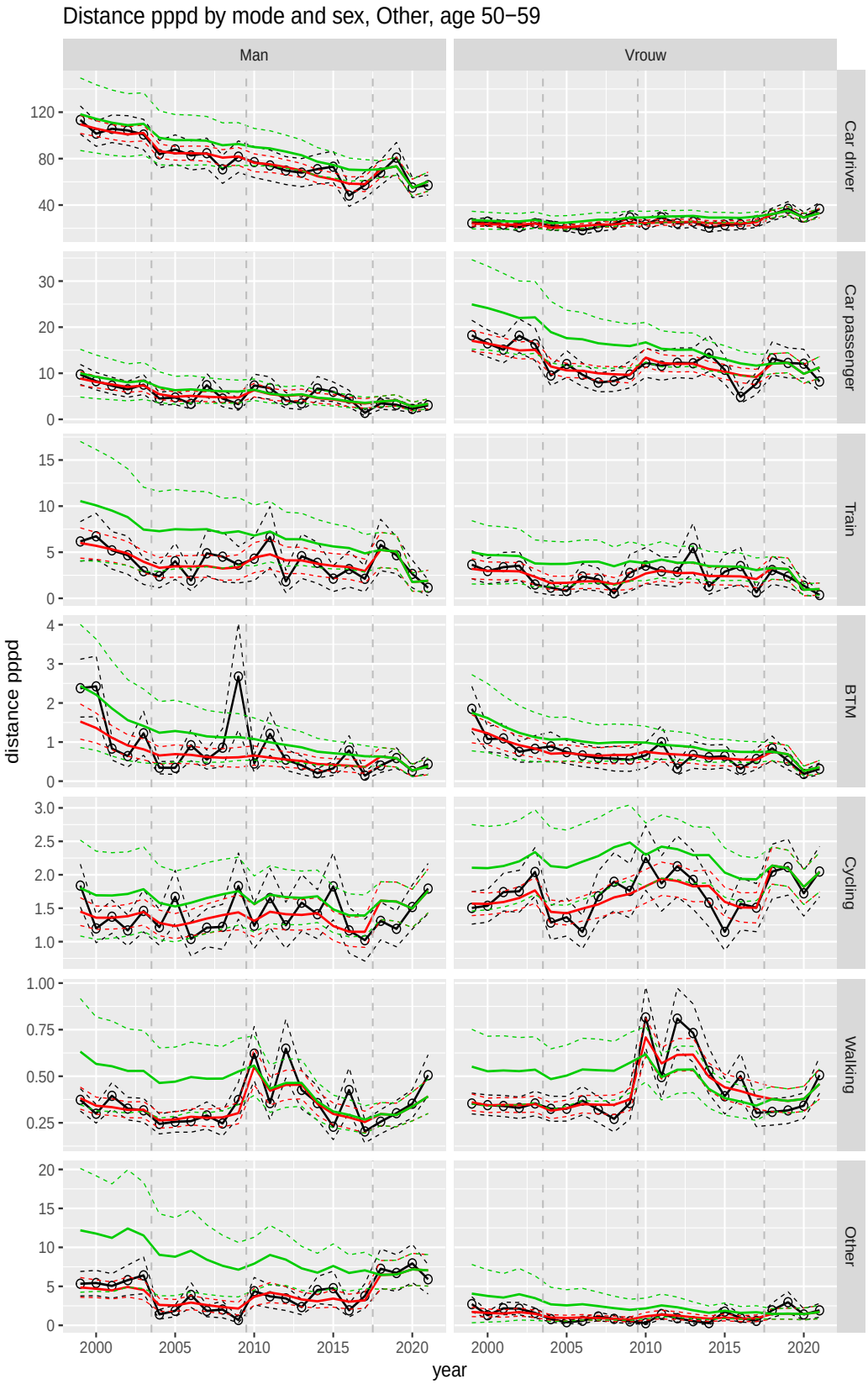

Figure A.246 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

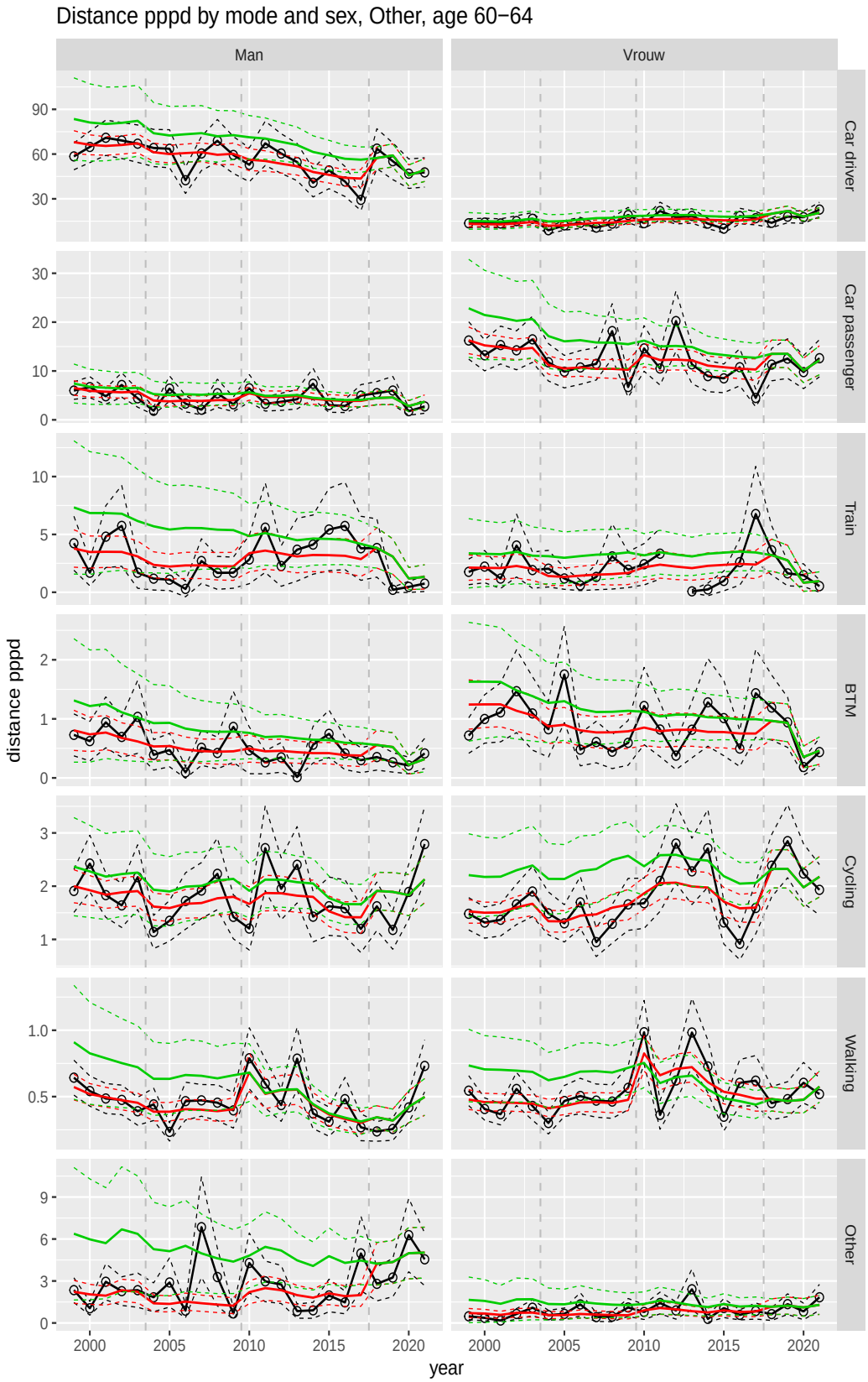

Figure A.247 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

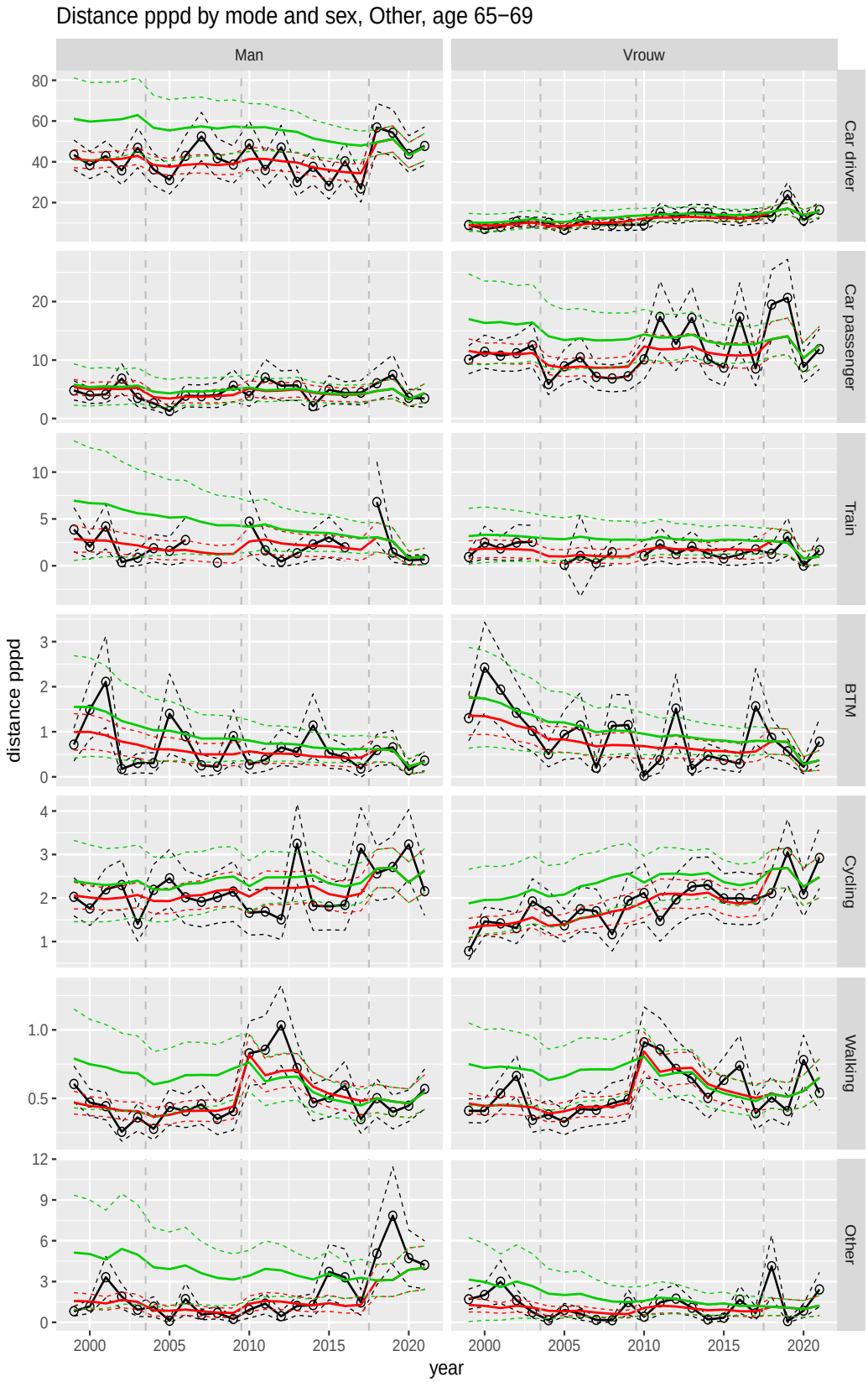

Figure A.248 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

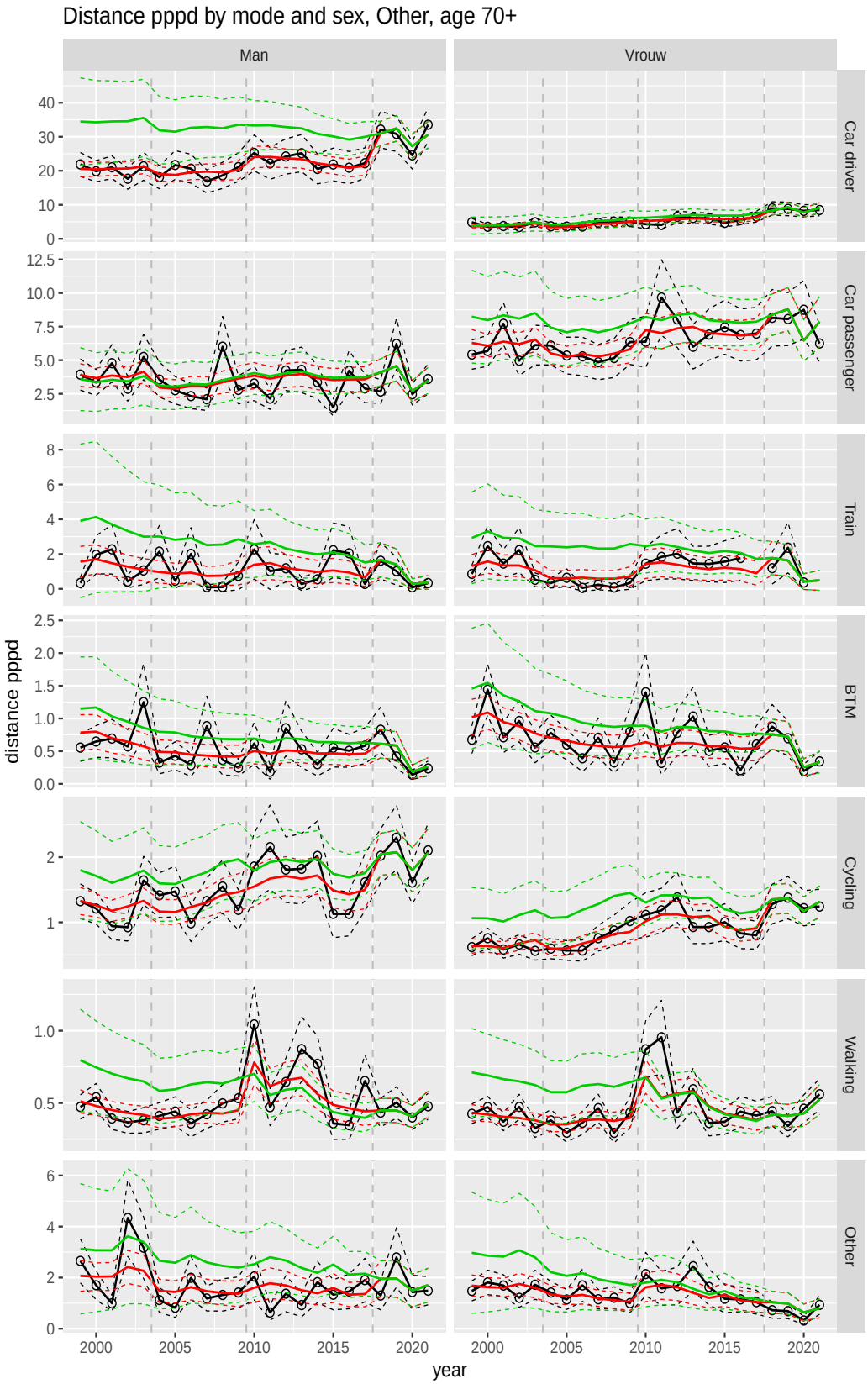

Figure A.249 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

B Plots revision analysis

B.1 Overall level revisions

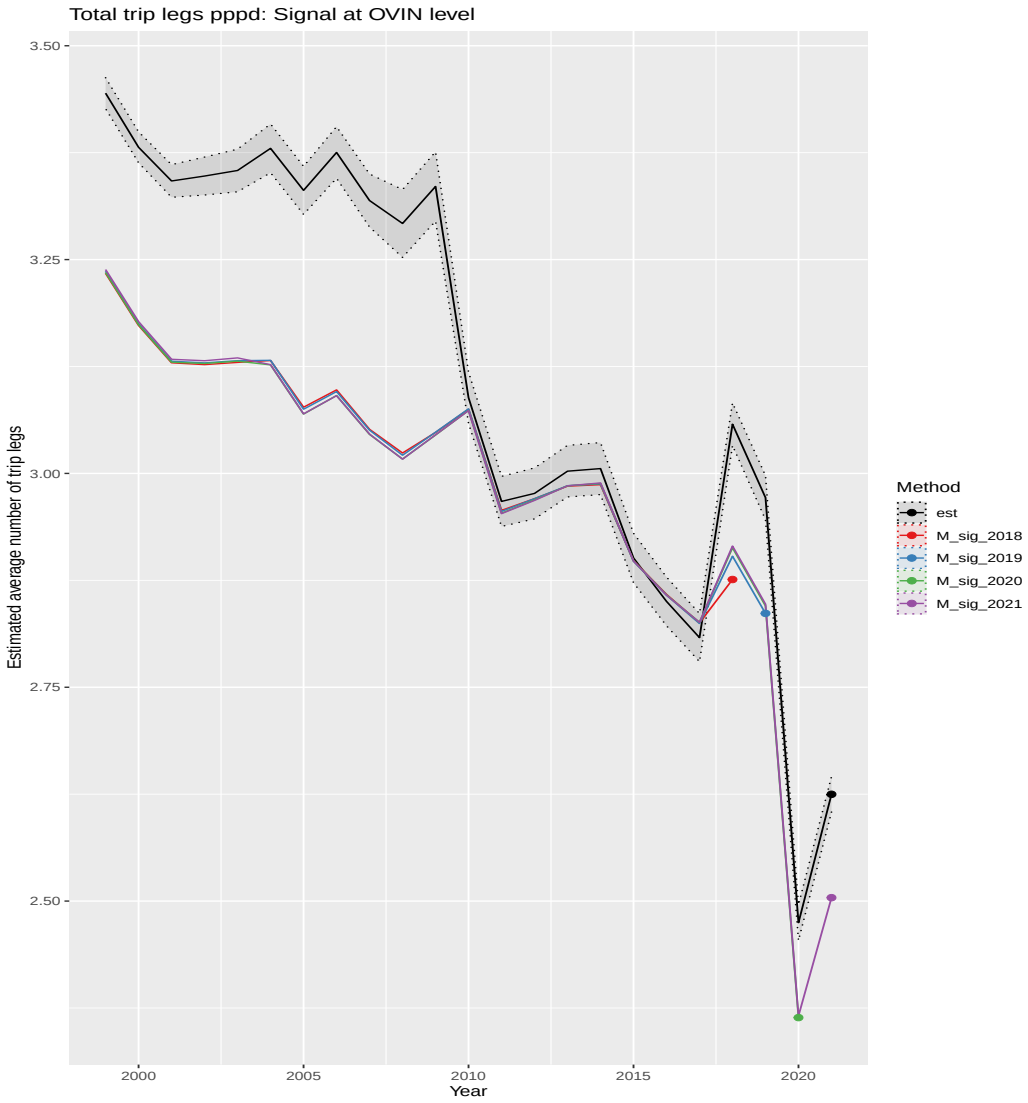

Figure B.1 Revision analysis number of trip legs overall level and trend estimates at OViN level.

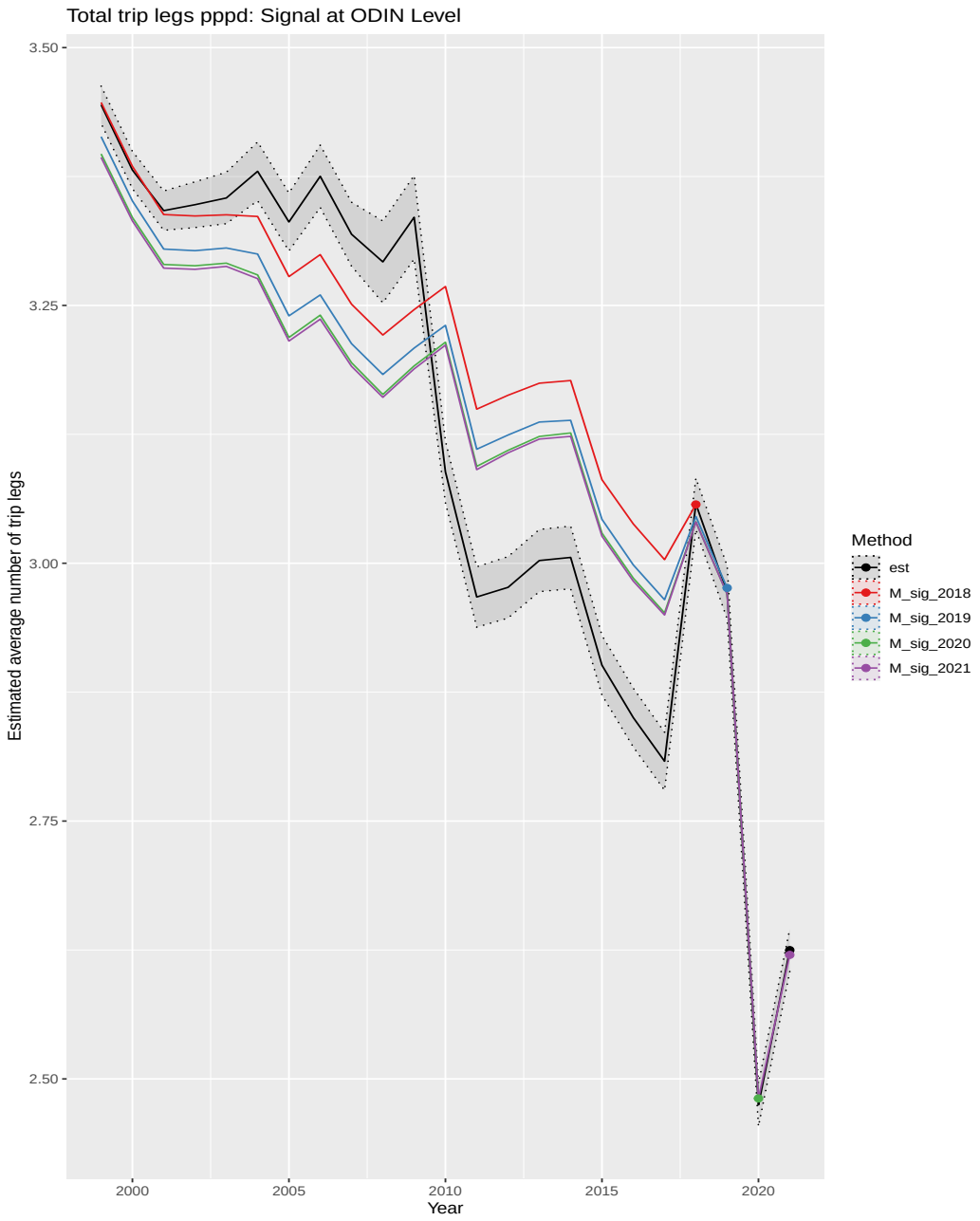

Figure B.2 Revision analysis number of trip legs overall level and trend estimates at ODiN level.

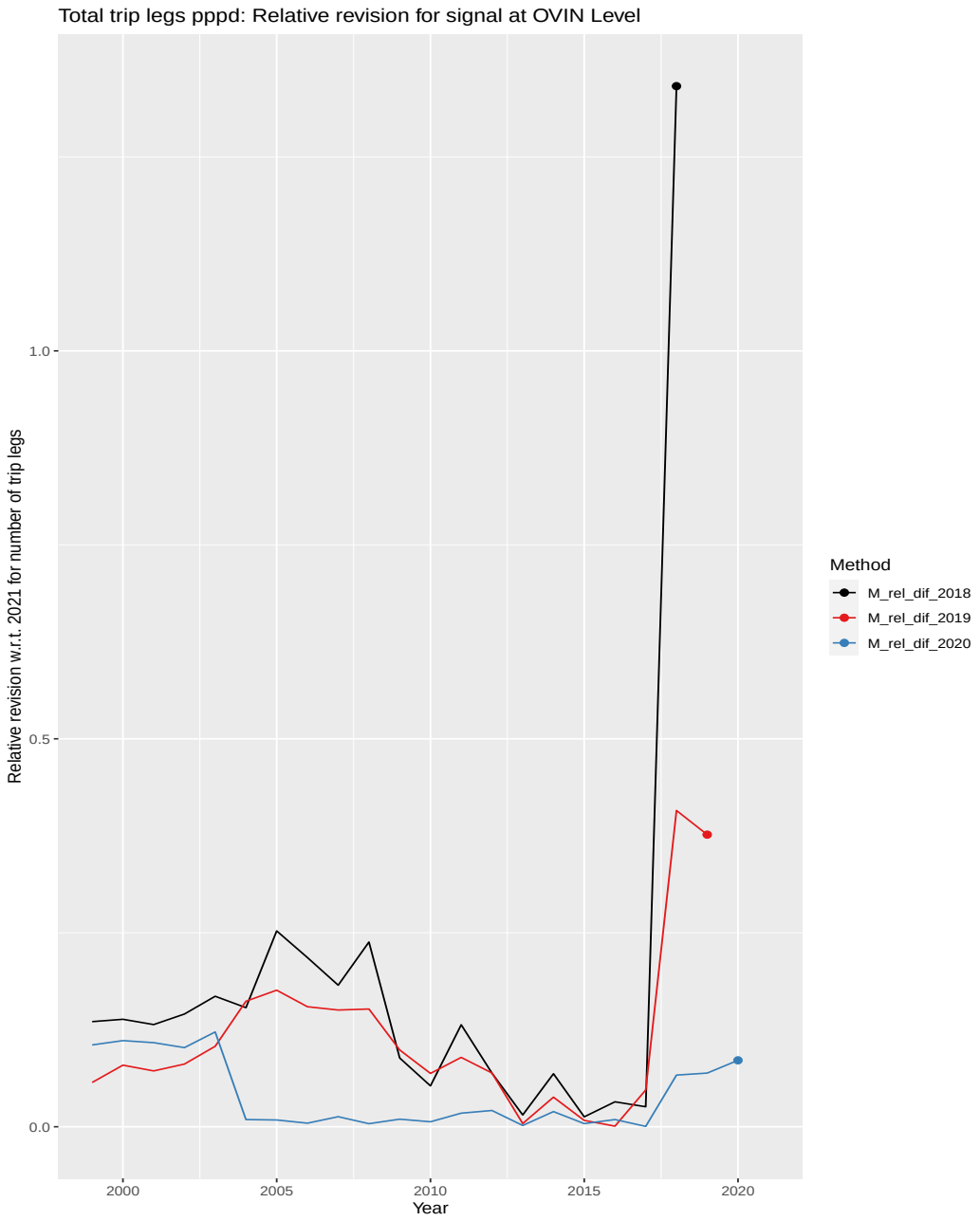

Figure B.3 Relative revision number of trip legs overall level and trend estimates at OViN level.

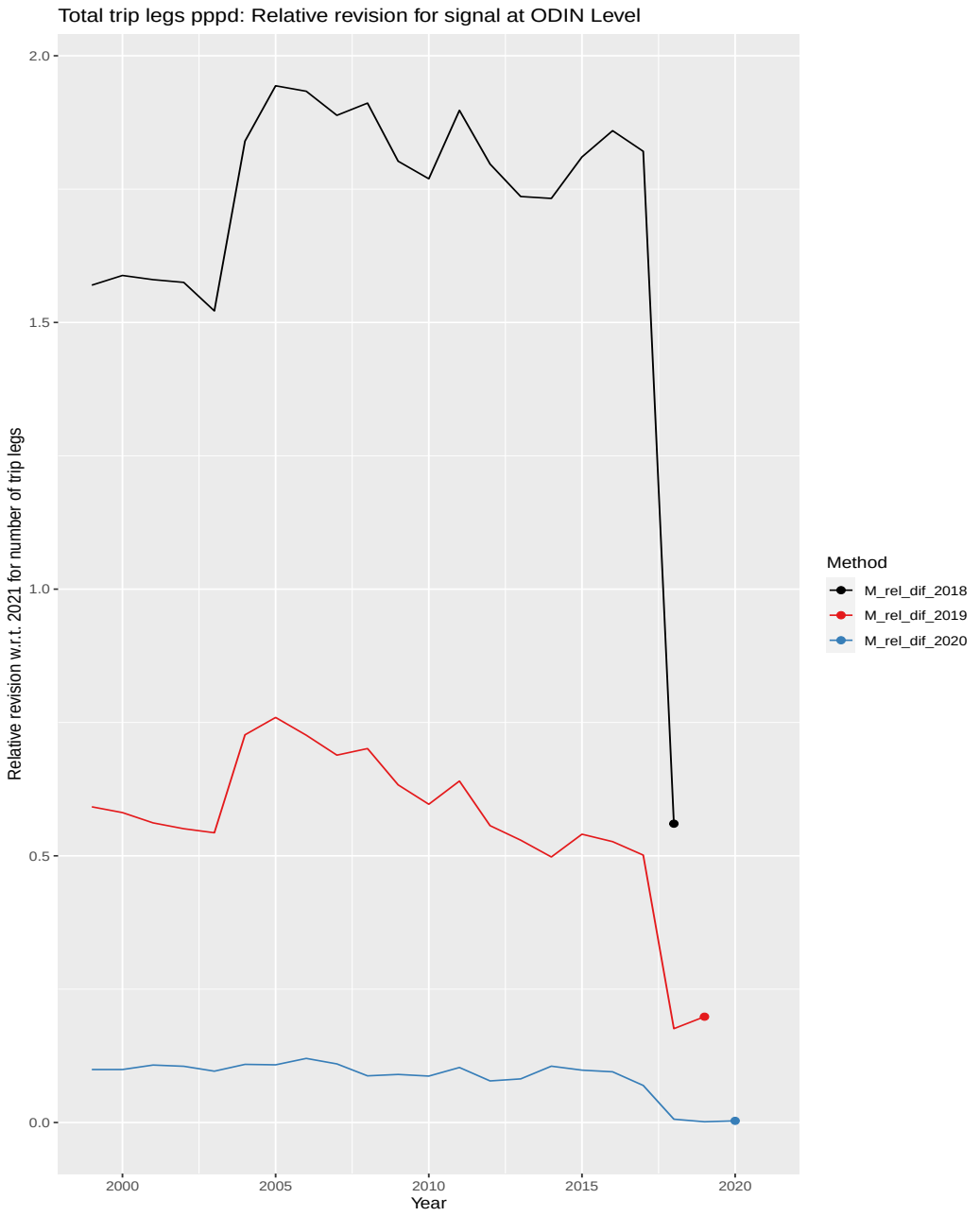

Figure B.4 Relative revision number of trip legs overall level and trend estimates at ODiN level.

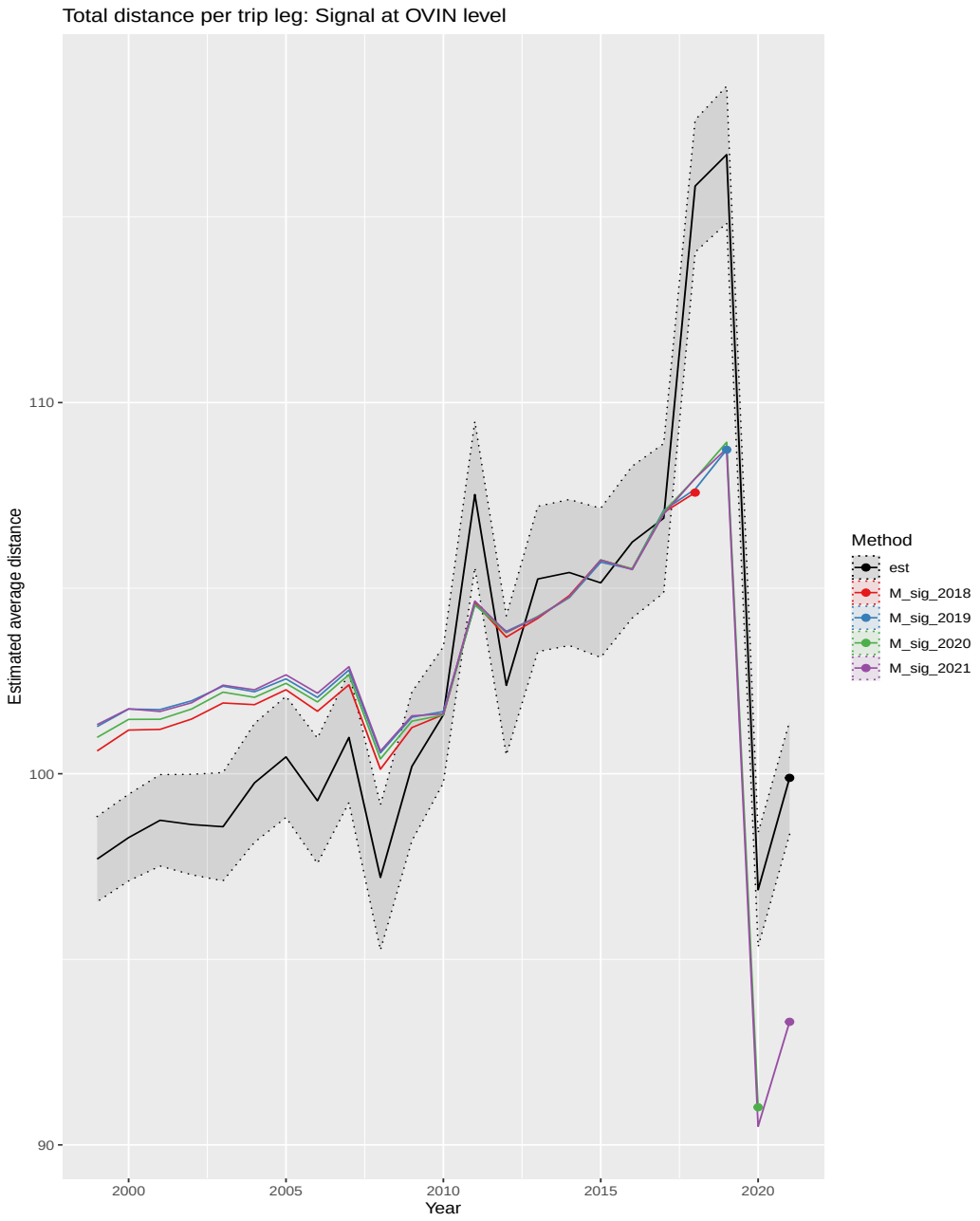

Figure B.5 Revision analysis distance per trip leg overall level and trend estimates at OViN level.

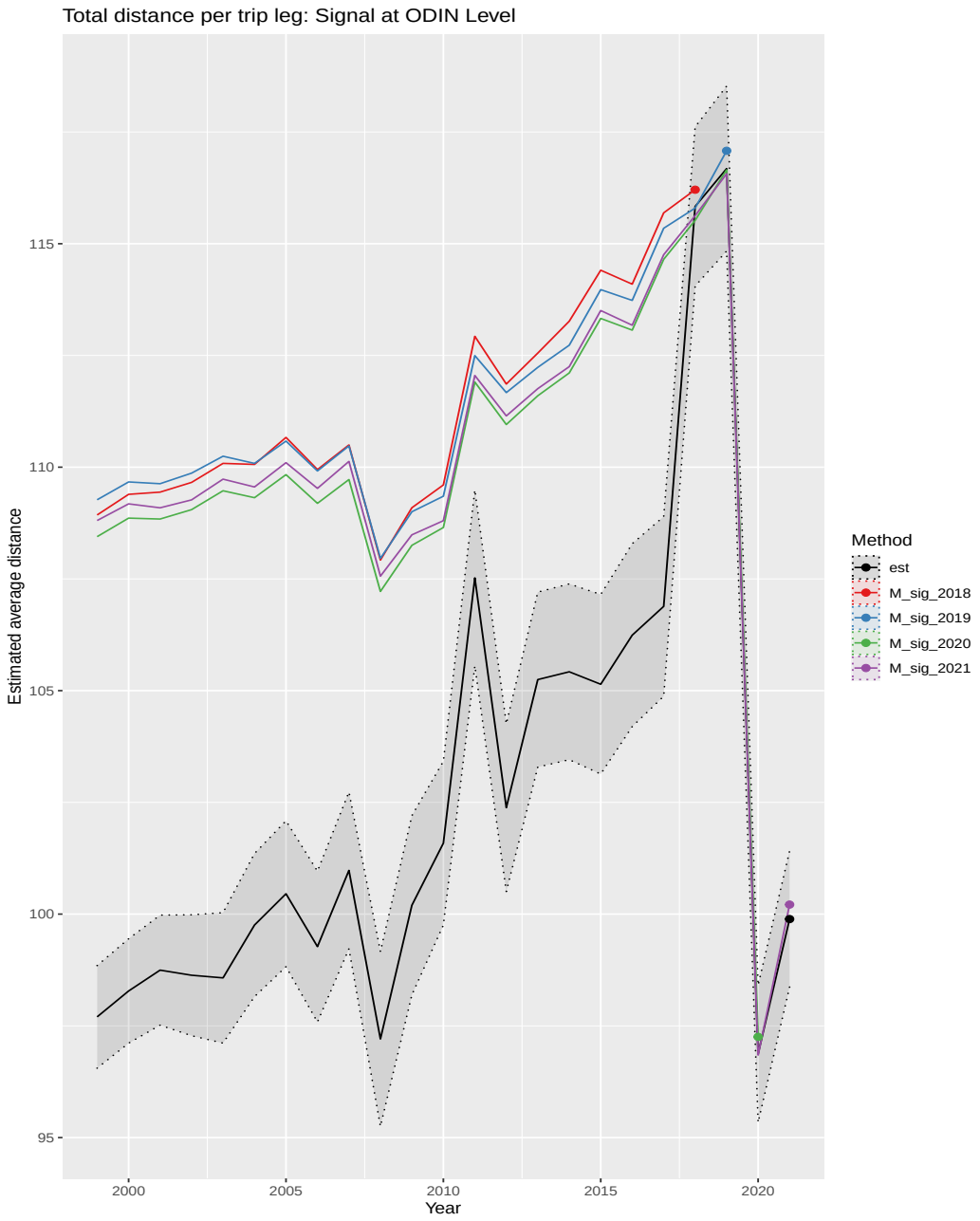

Figure B.6 Revision analysis distance per trip leg overall level and trend estimates at ODiN level.

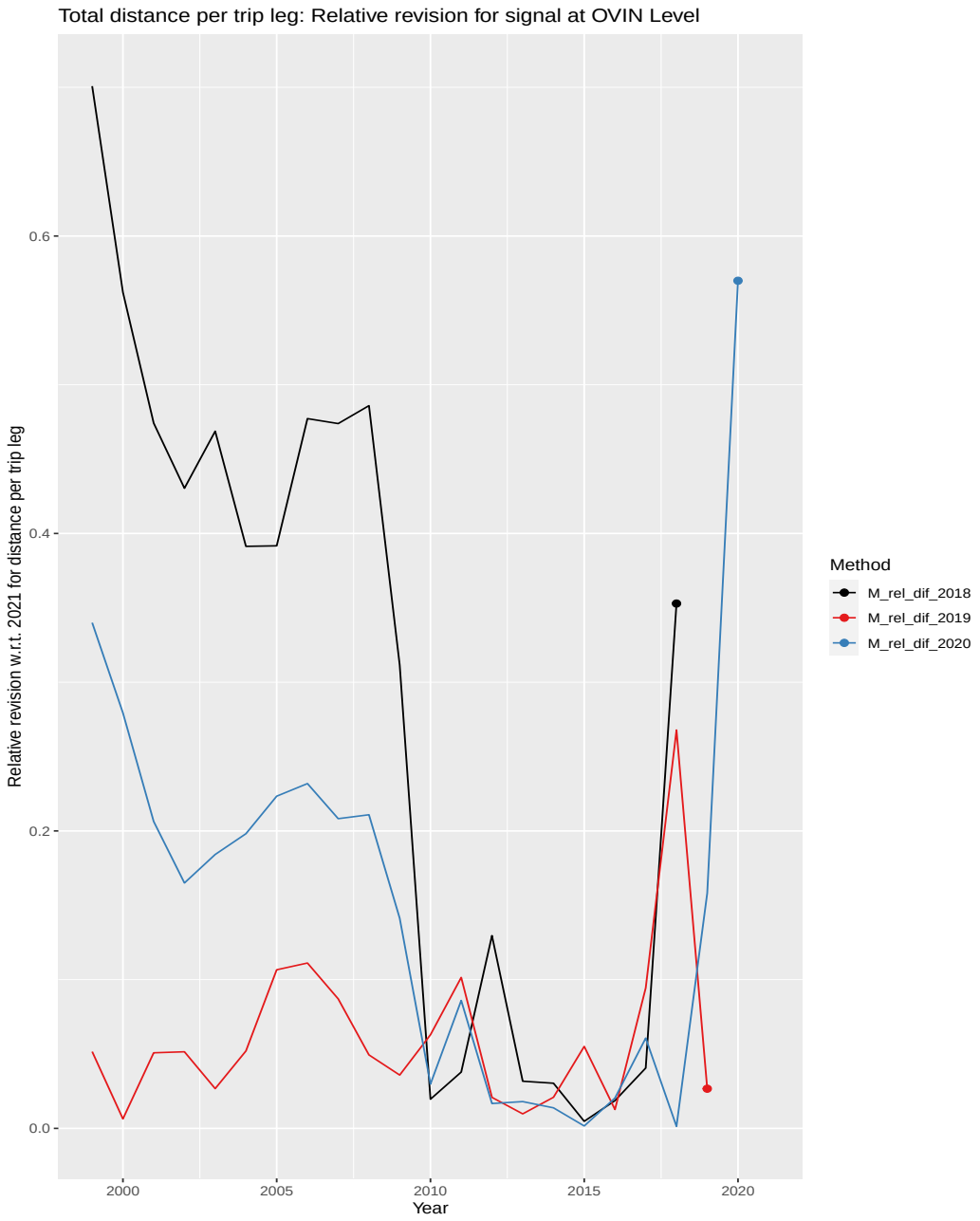

Figure B.7 Relative revision distance per trip leg overall level and trend estimates at OViN level.

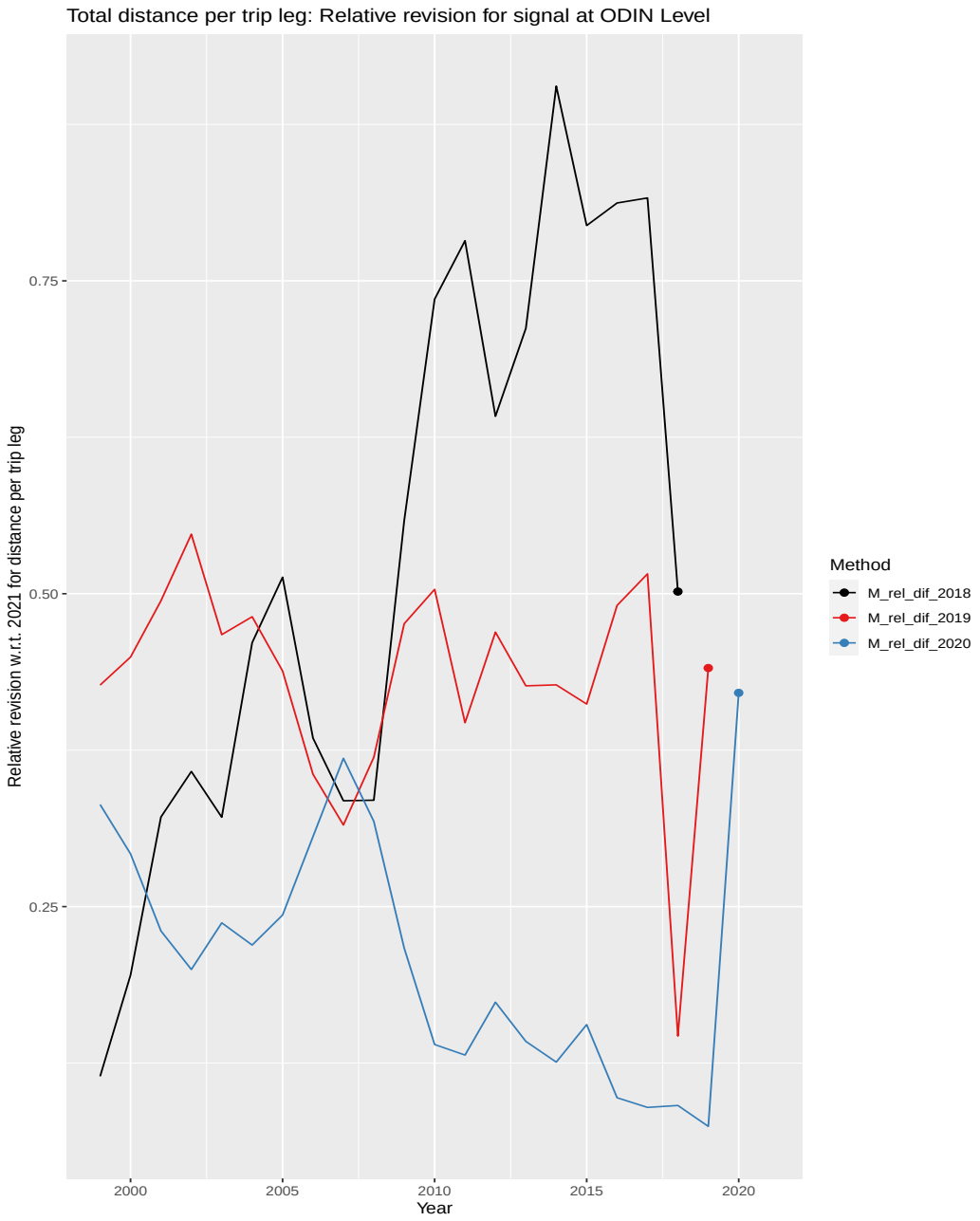

Figure B.8 Relative revision distance per trip leg overall level and trend estimates at ODiN level.

B.2 Revisions by transportation mode

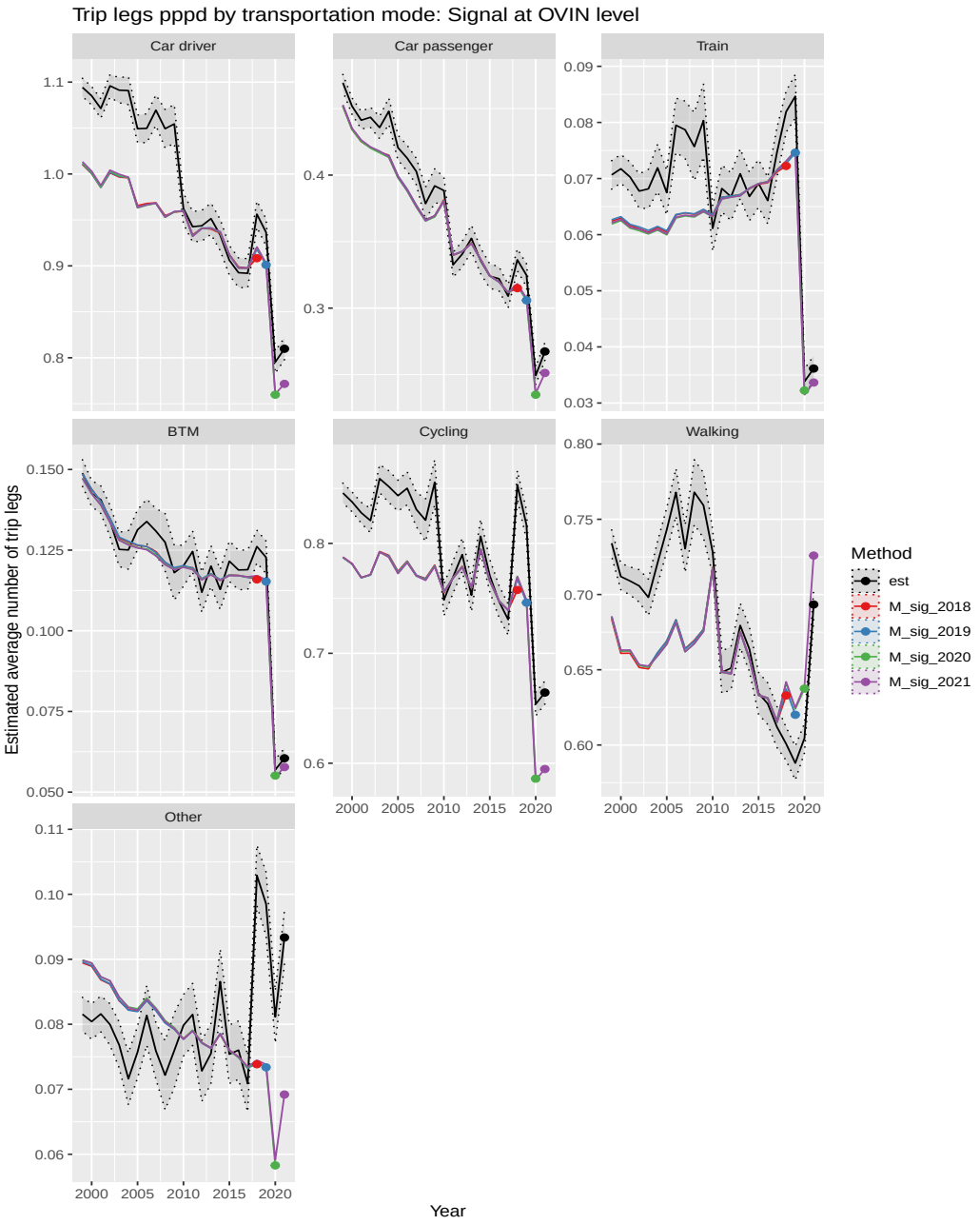

Figure B.9 Revision analysis number of trip legs for transportation modes and trend estimates at OViN level.

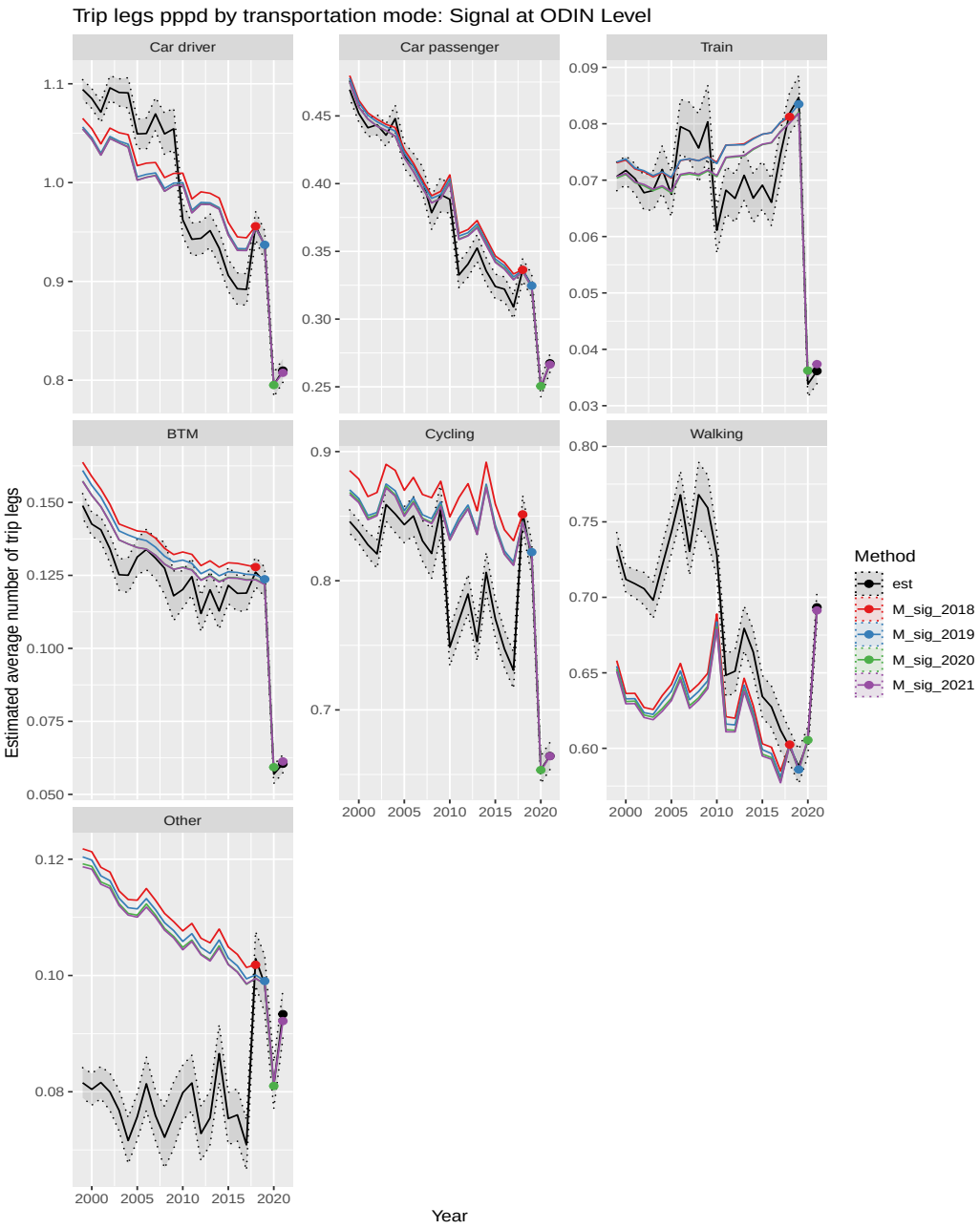

Figure B.10 Revision analysis number of trip legs for transportation modes and trend estimates at ODiN level.

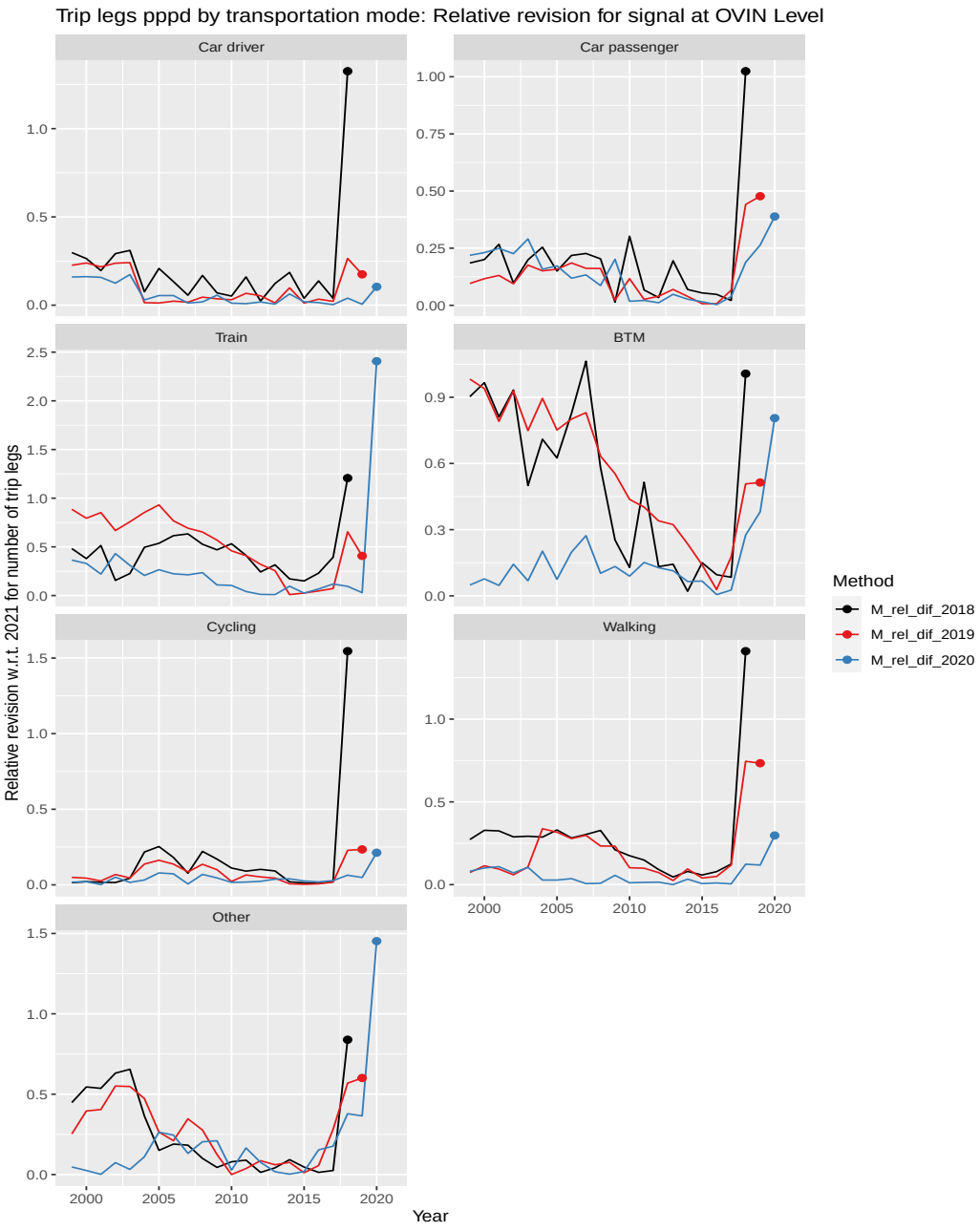

Figure B.11 Relative revision number of trip legs for transportation modes and trend estimates at OViN level.

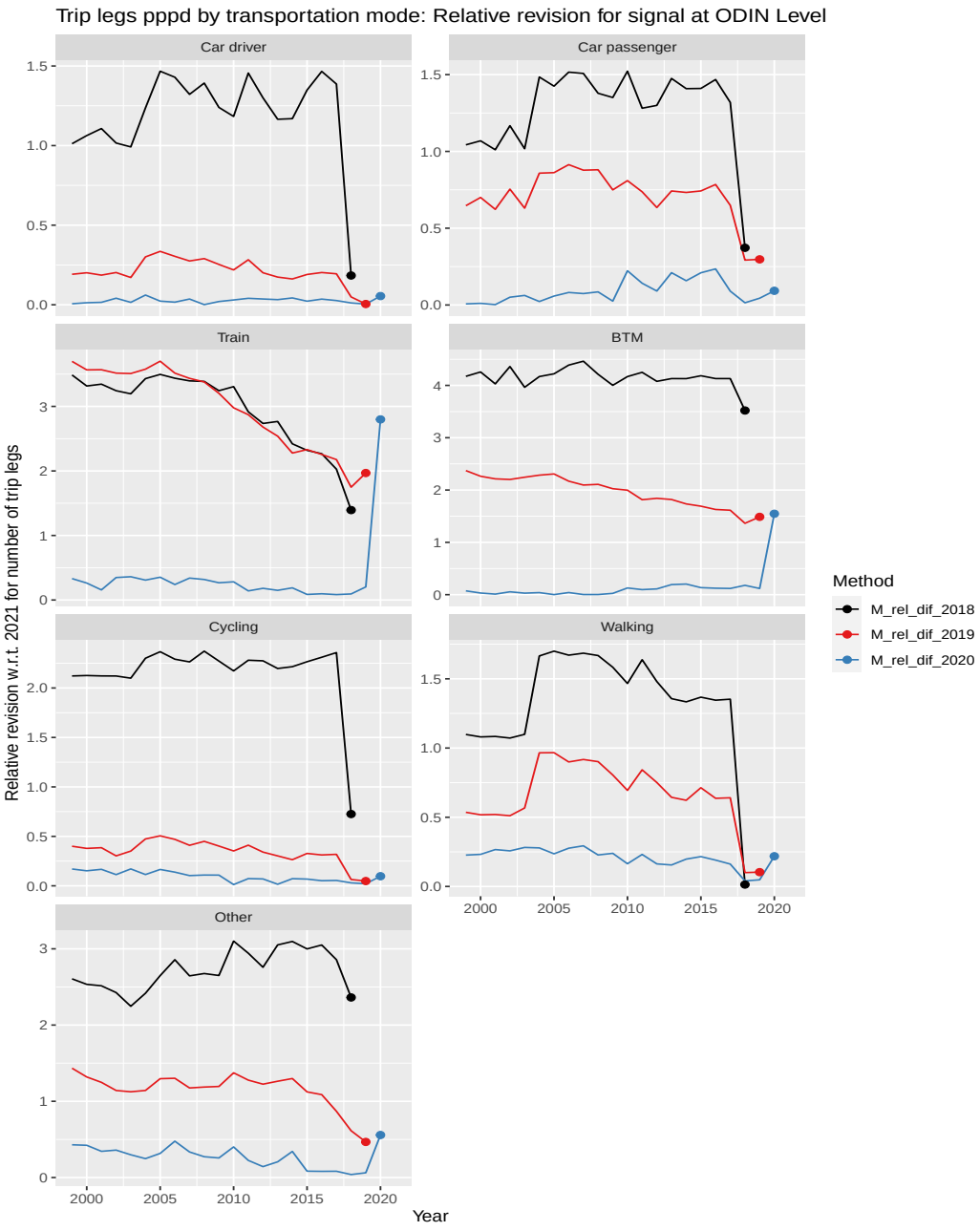

Figure B.12 Relative revision number of trip legs for transportation modes and trend estimates at ODiN level.

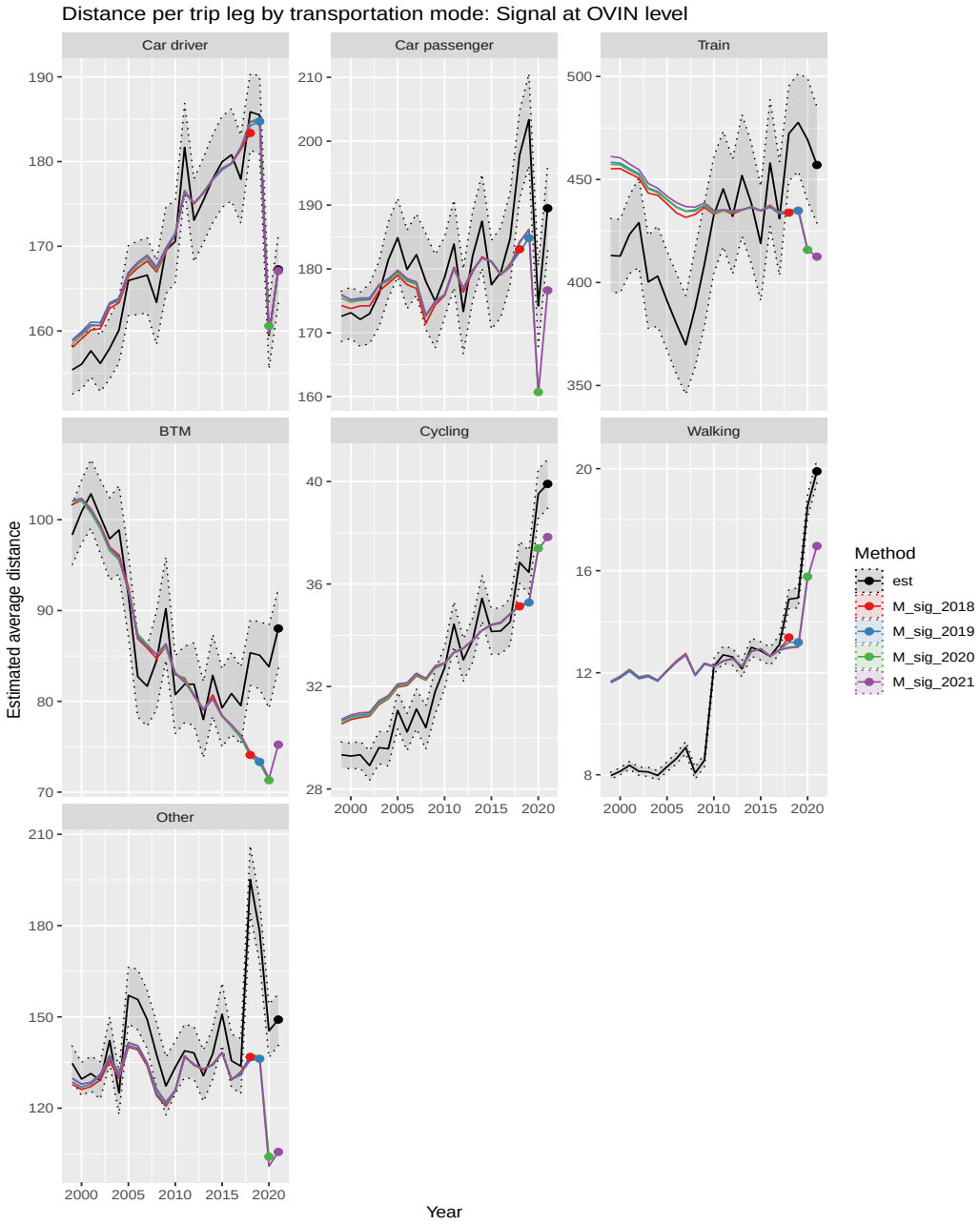

Figure B.13 Revision analysis distance per trip leg for transportation modes and trend estimates at OViN level.

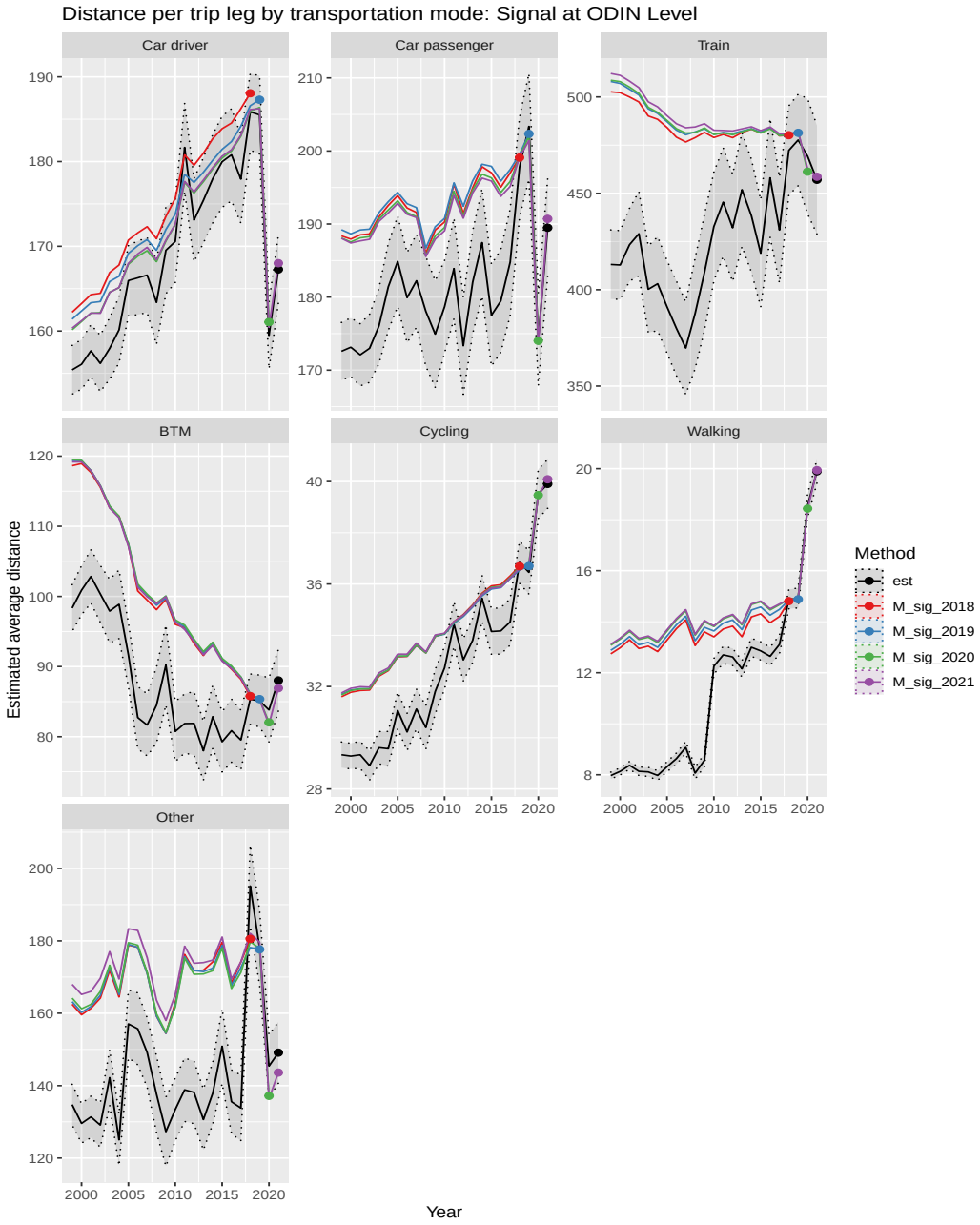

Figure B.14 Revision analysis distance per trip leg for transportation modes and trend estimates at ODiN level.

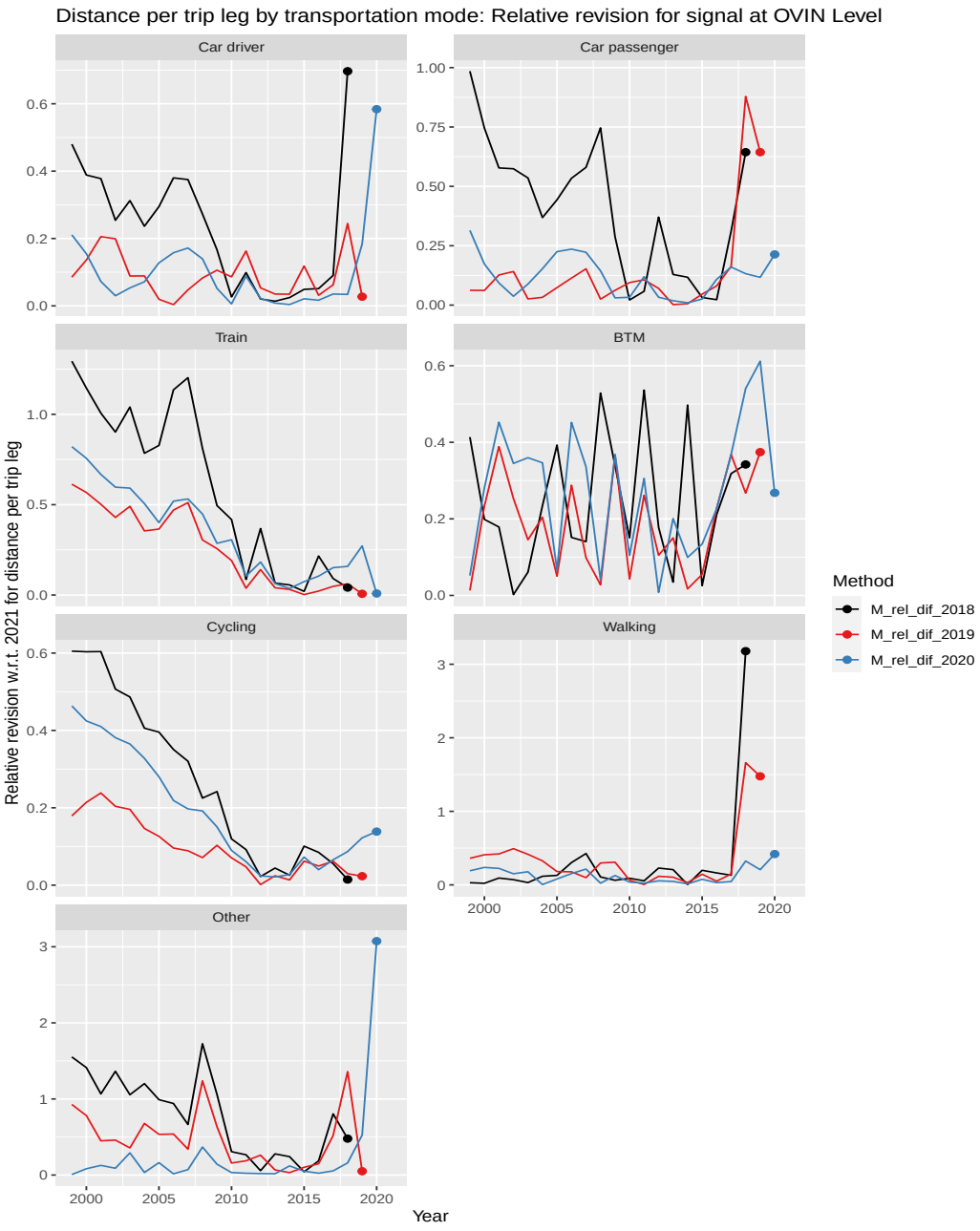

Figure B.15 Relative revision distance per trip leg for transportation modes and trend estimates at OViN level.

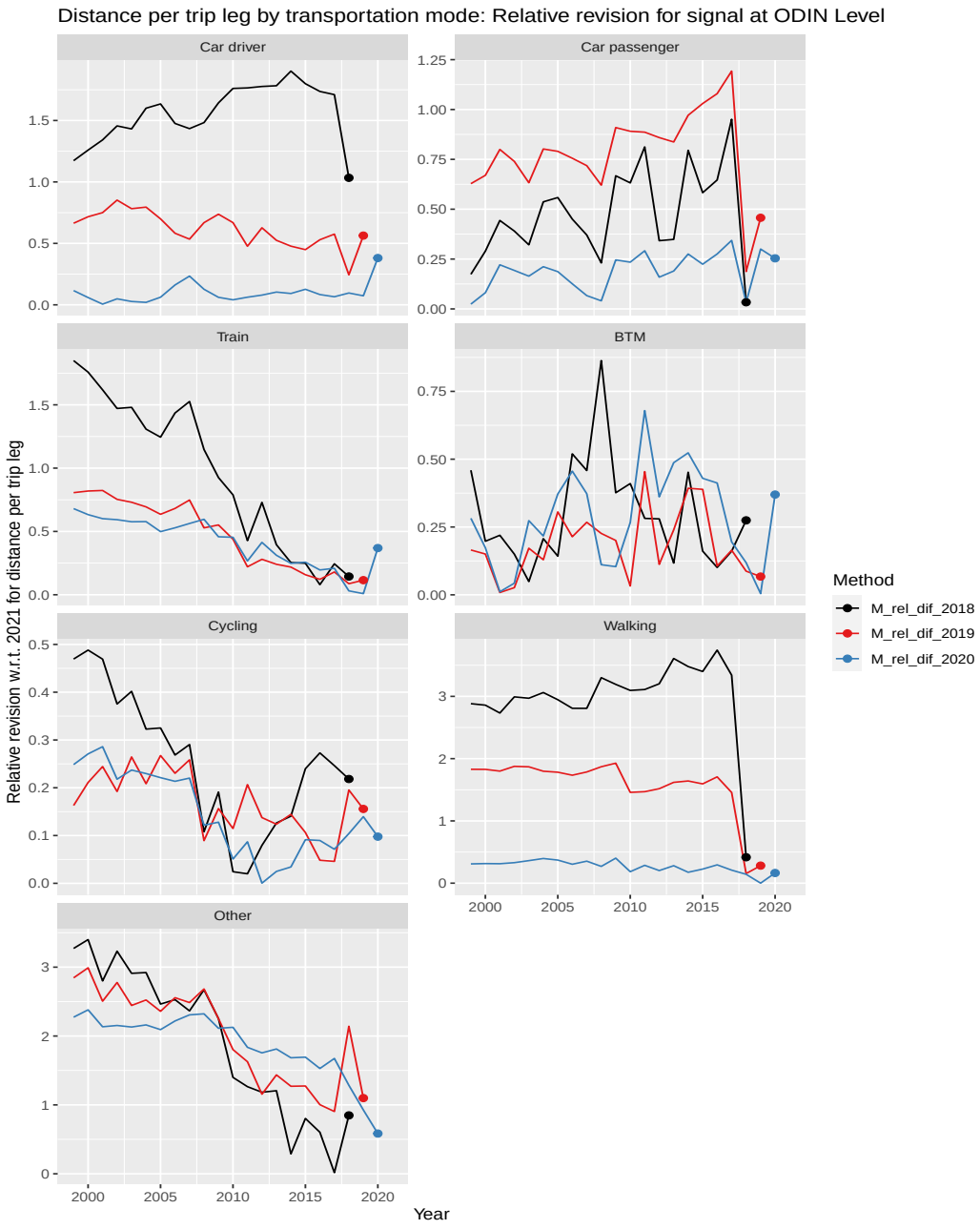

Figure B.16 Relative revision distance per trip leg for transportation modes and trend estimates at ODiN level.

B.3 Revisions by purpose

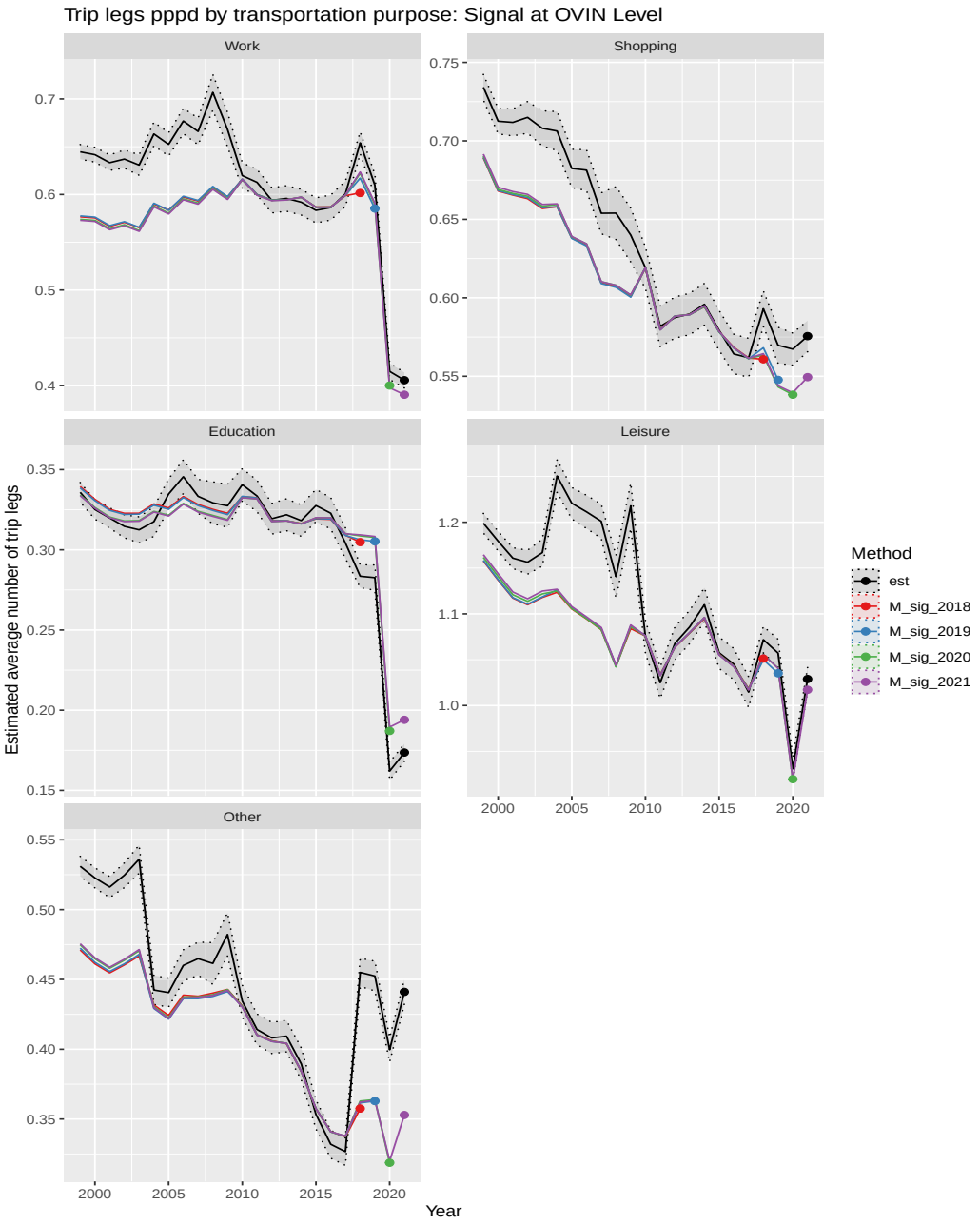

Figure B.17 Revision analysis number of trip legs for transportation purposes and trend estimates at OViN level.

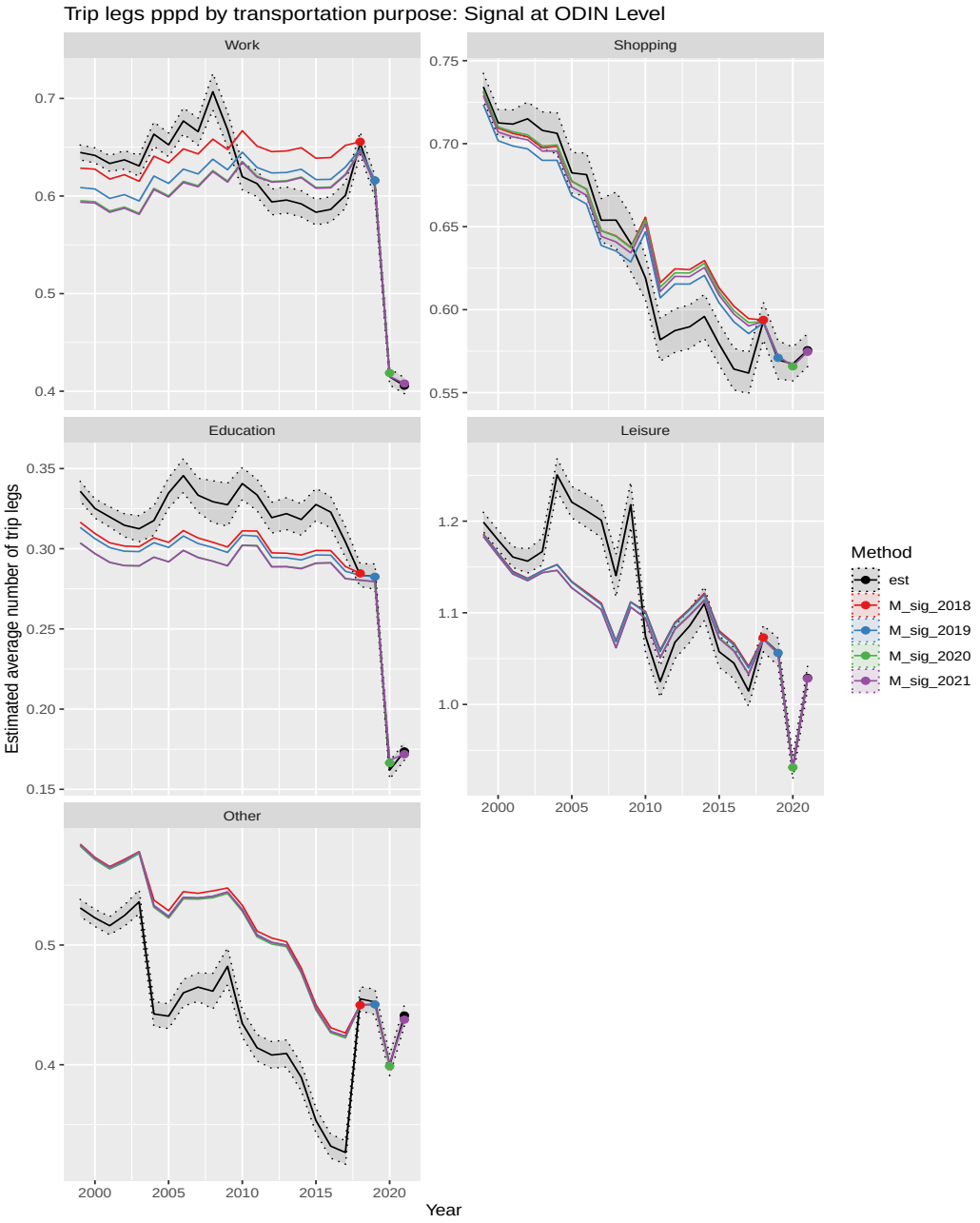

Figure B.18 Revision analysis number of trip legs for transportation purposes and trend estimates at ODiN level.

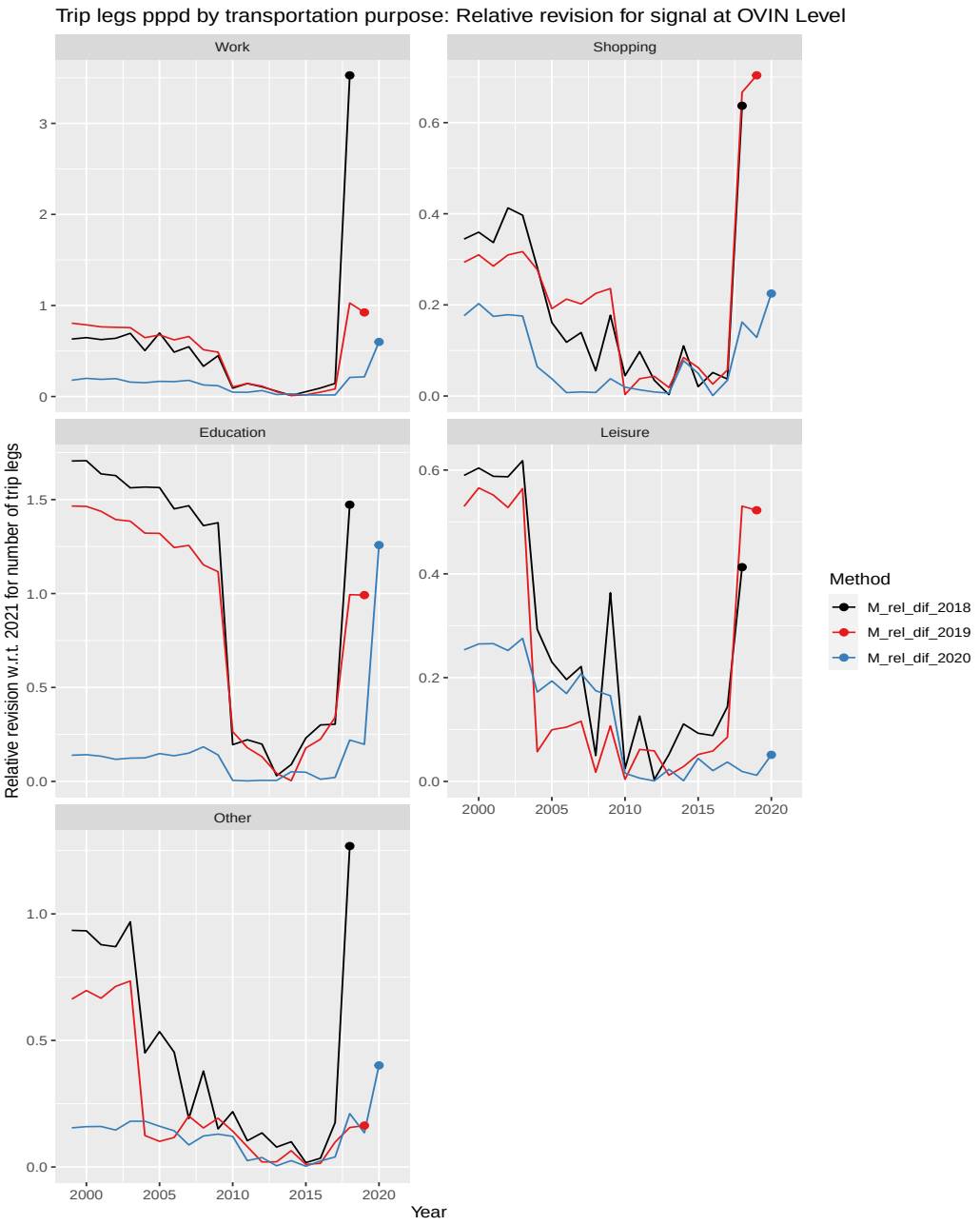

Figure B.19 Relative revision number of trip legs for transportation purposes and trend estimates at OViN level.

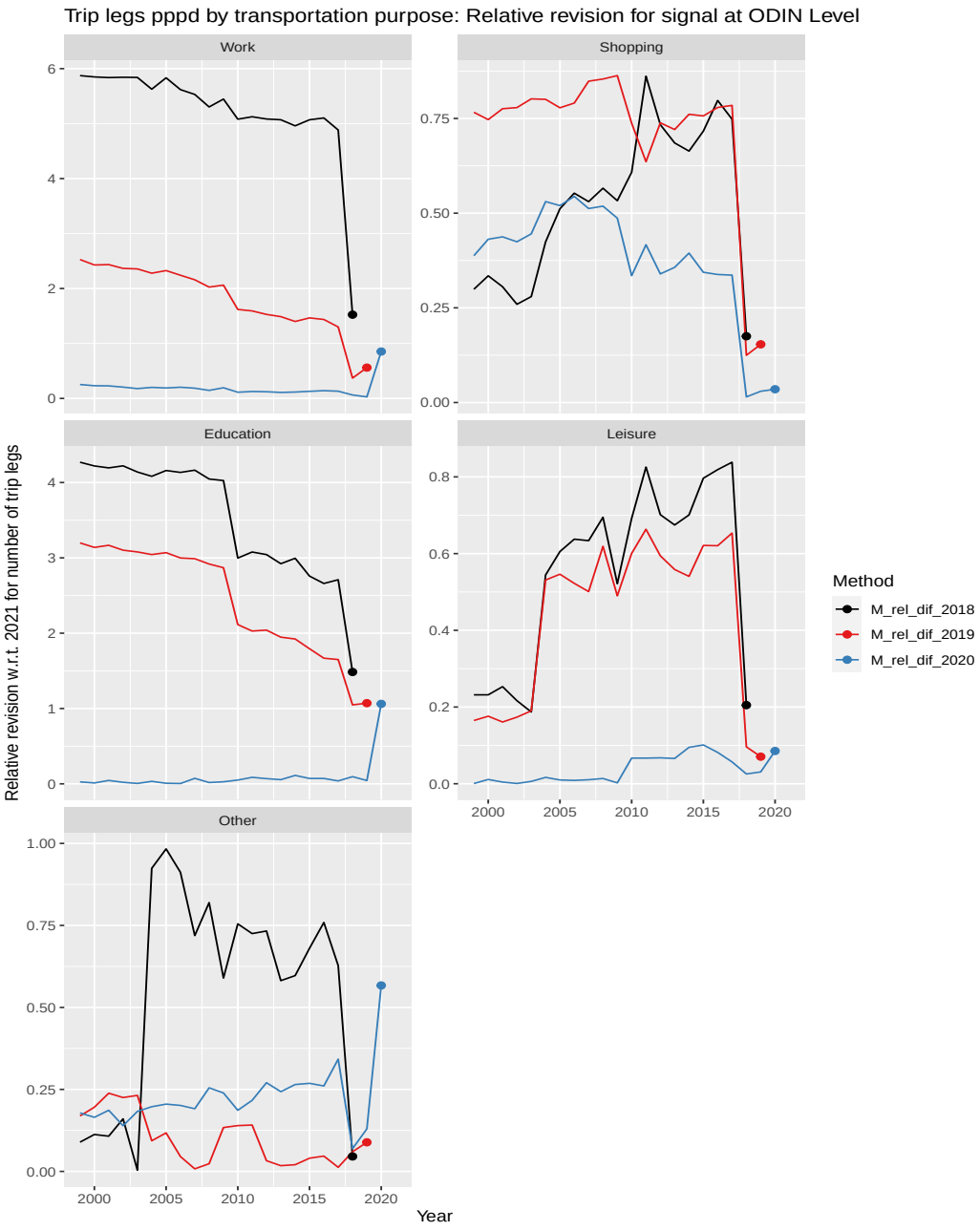

Figure B.20 Relative revision number of trip legs for transportation purposes and trend estimates at ODiN level.

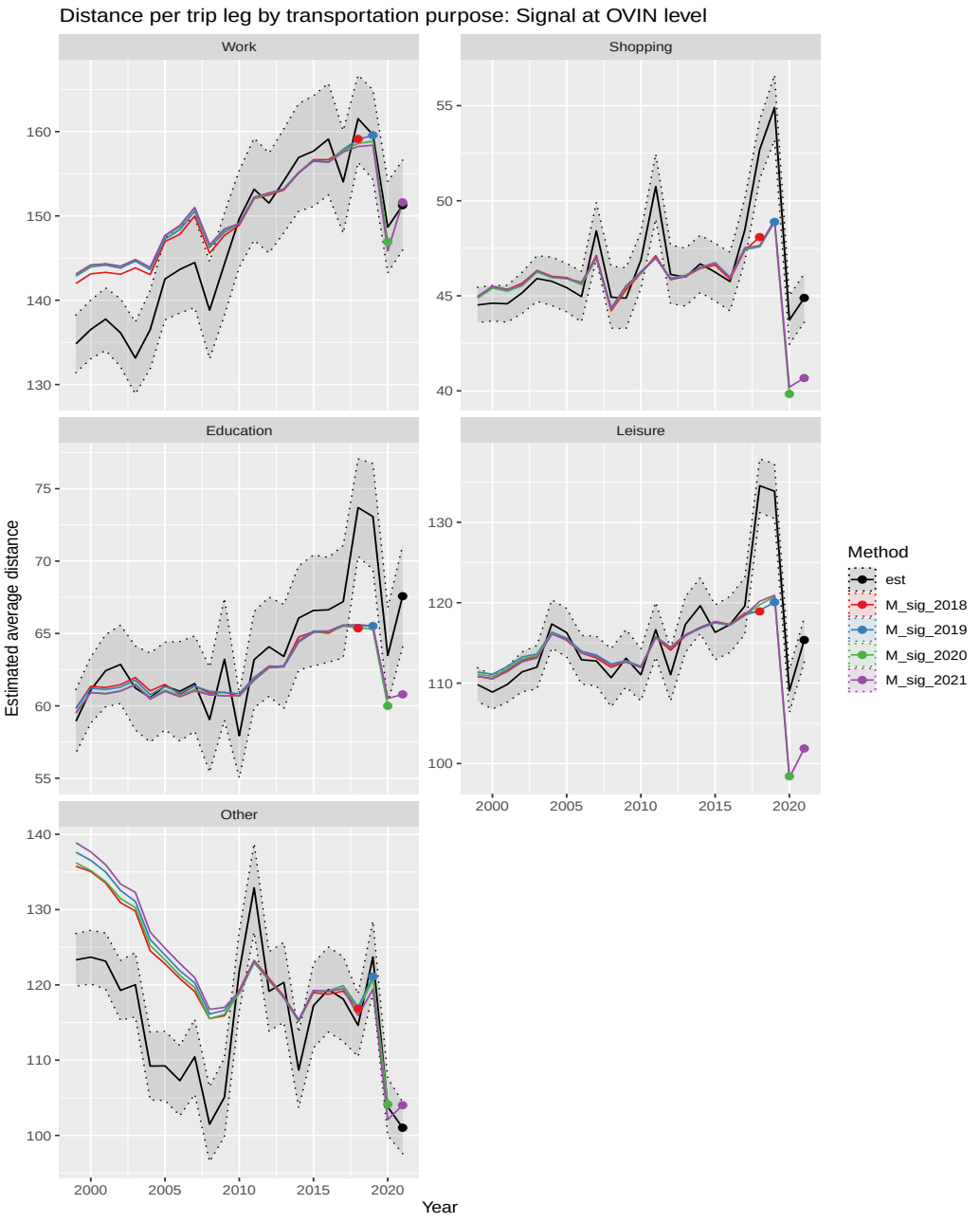

Figure B.21 Revision analysis distance per trip leg for transportation purposes and trend estimates at OViN level.

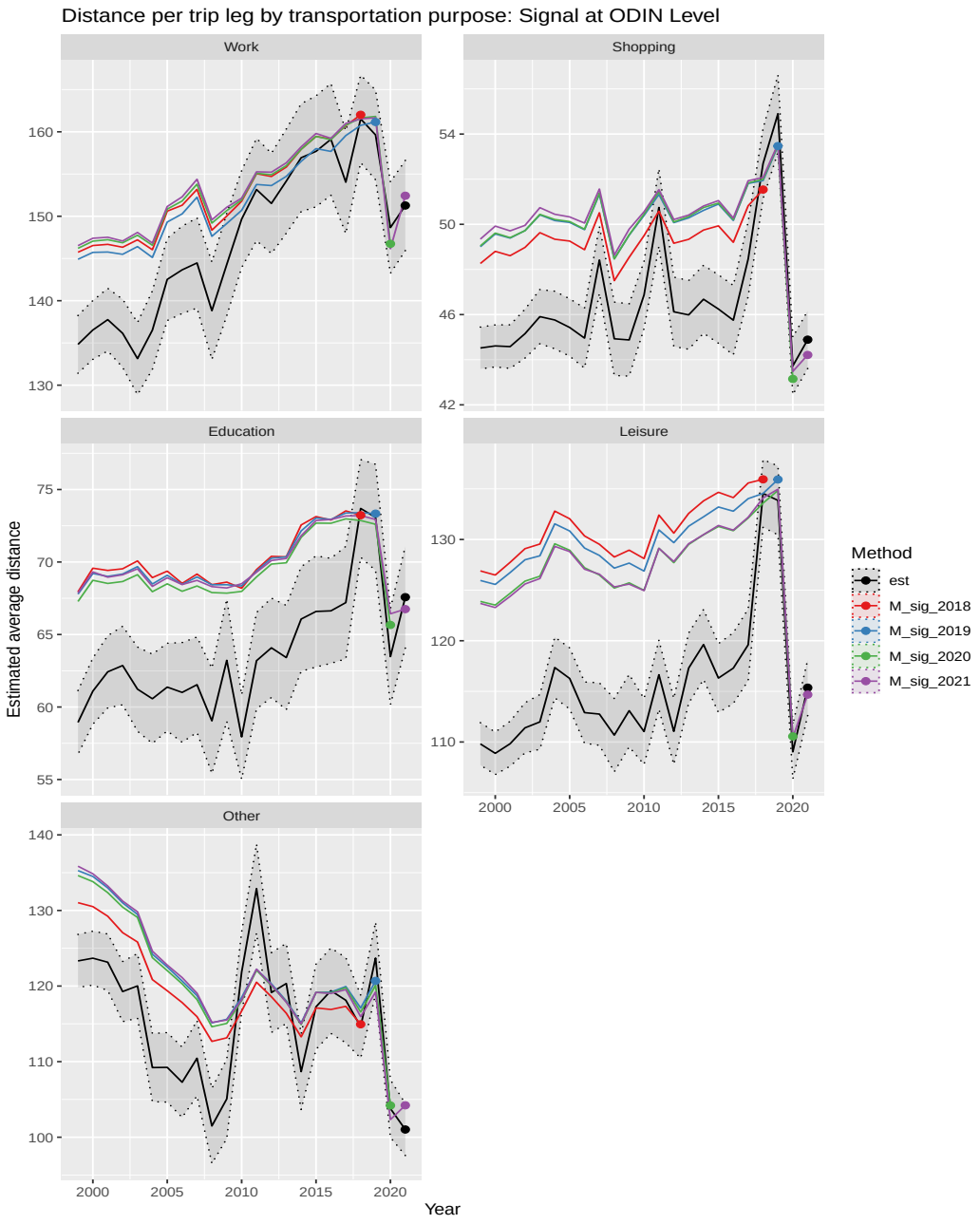

Figure B.22 Revision analysis distance per trip leg for transportation purposes and trend estimates at ODiN level.

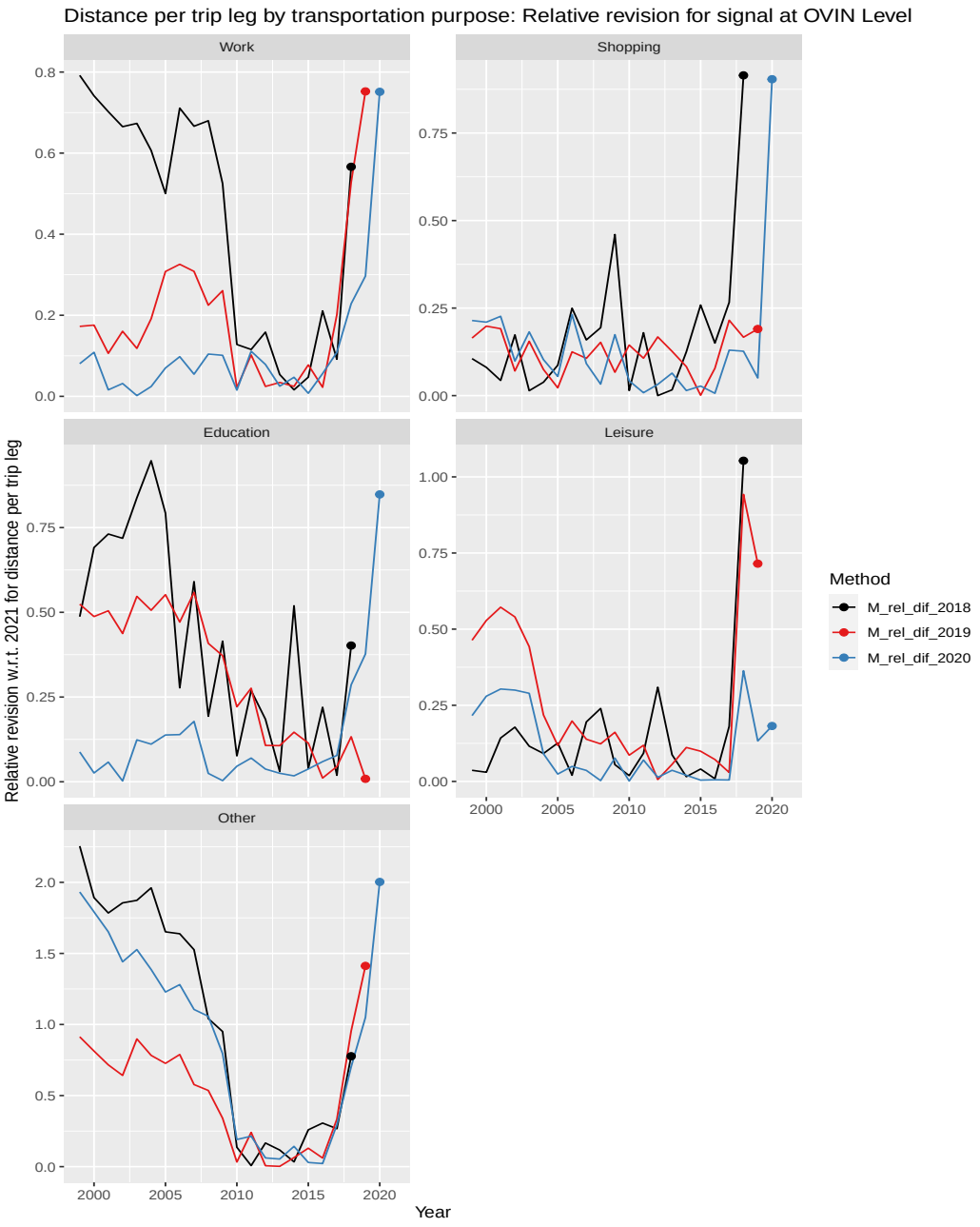

Figure B.23 Relative revision distance per trip leg for transportation purposes and trend estimates at OViN level.

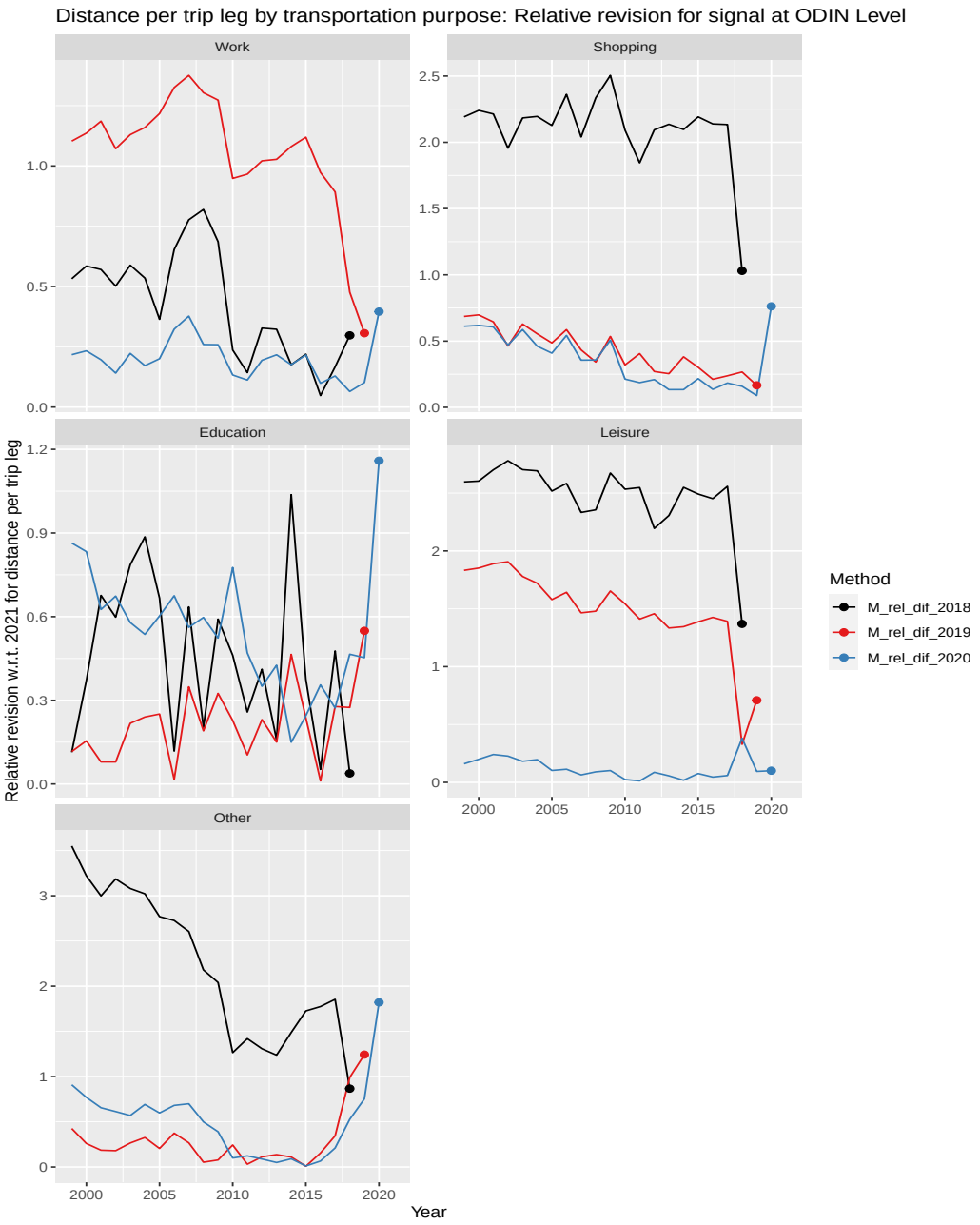

Figure B.24 Relative revision distance per trip leg for transportation purposes and trend estimates at ODiN level.

Colophon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Grafimedia

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information