The purpose of this piece is to explain the derivation of Eq. 19 to Eq. 21 on p. 6265 of Xie (2013).
Xie, Y. 2013. "Population Heterogeneity and Causal Inference." *Proceedings of the National Academy of Sciences* 110(16):6262–68.

All equations with numbers are from the original article.

$$ITE(t,t') = \frac{E(Y|T=t') - E(Y|T=t)}{F(t') - F(t)} \tag{17}$$

This holds because ITE can be conceptualized as LATE in the IV framework hence can be estimated by a Wald estimator in the form of Eq. 17. In order to understand this, imagine the dataset being expanded so that each data point is represented twice, at time t and t', respectively.

Potential outcome	Potential outcome	Treatment status	T (IV)
$Y_{i=1}^{1}(t=0)$	$Y_{i=1}^{0}(t=0)$	0	1
$Y_{i=2}^{1}(t=0)$	$Y_{i=2}^{0}(t=0)$	0	1
$Y_{i=3}^{1}(t=0)$	$Y_{i=3}^{0}(t=0)$	1	1
$Y_{i=1}^{1}(t=1)$	$Y_{i=1}^{0}(t=1)$	0	2
$Y_{i=2}^{1}(t=1)$	$Y_{i=2}^{0}(t=1)$	1	2
$Y_{i=3}^{1}(t=1)$	$Y_{i=3}^{0}(t=1)$	1	2

If the potential outcomes are time-constant, then ignorability of T ($E(Y_i^1 \middle| T=t) = E(Y_i^1 \middle| T=t')$ & $E(Y_i^0 \middle| T=t) = E(Y_i^0 \middle| T=t')$) is trivially fulfilled. And T only affects the observed mean $E(Y)$ via treatment expansion. Hence the Wald estimator gives a valid estimate of LATE/ITE.

Taking limit for Eq. 17 and get Eq. 18,

$$\begin{aligned} ITE(t) &= \lim_{F(t')-F(t)\to 0} \frac{E(Y|T=t') - E(Y|T=t)}{F(t') - F(t)} \\ &= \frac{dE[Y(t)]}{dF(t)} \end{aligned} \tag{18}$$

Now I prove Eq. 19.

$$\frac{1}{F(t)}\int_0^t ITE(u)dF(u) \tag{19}$$

$$= \frac{1}{F(t)}\int_0^t \frac{dE[Y(t)]}{dF(t)} dF(u) = \frac{1}{F(t)}\int_0^t dE[Y(t)] = \frac{1}{F(t)}\int_0^t \frac{dE[Y(t)]}{du} du$$

The last equality is by the substitution rule of integration.
As $E[Y(t)]$ is an antiderivative of $\frac{dE[Y(t)]}{du}$, by fundamental theorem of calculus, we get

$$\frac{1}{F(t)}\int_{0}^{t}\frac{dE[Y(t)]}{du}du = \frac{E[Y(t)] - E[Y(0)]}{F(t)}$$

Because of the way $F(t)$ is defined (the cumulative proportion), $F(0) = 0$. Thus,

$$\frac{E[Y(t)] - E[Y(0)]}{F(t)} = \frac{E[Y(t)] - E[Y(0)]}{F(t) - F(0)}$$

, which is the Wald estimator for the LATE:

$$E(Y^1 - Y^0 | D(t) = 1, D(0) = 0)$$

As $D(0) = 0$ is universally true, we get

$$E(Y^1 - Y^0 | D(t) = 1)$$

, which is just the treatment effect on the treated (TT) at time t.

Eq. 20 and Eq. 21 can be proved very similarly. For Eq. 21, simply note that

$$\int_{0}^{\infty} ITE(u)dF(u) = \int_{0}^{\infty}\frac{dE[Y(t)]}{du}du = E[Y(\infty)] - E[Y(0)]$$

When $\mathrm{t} = \infty$, all units are treated, and when $\mathrm{t} = 0$, no unit is treated. Hence, $E[Y(\infty)] - E[Y(0)] = E(Y^1 - Y^0)$, which is the ATE.